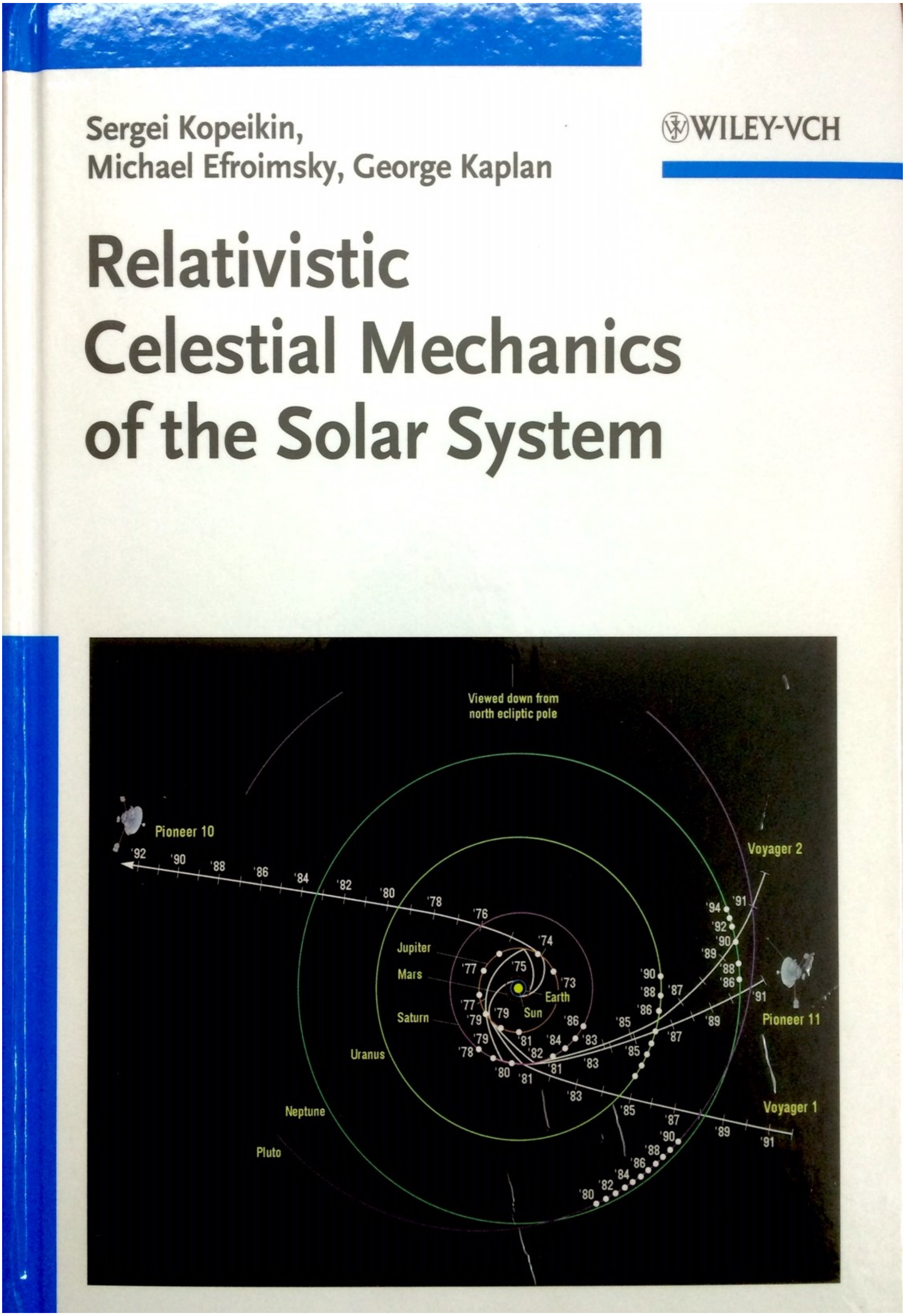
Sergei Kopeikin,
Michael Efroimsky, George Kaplan
WILEY-VCH
Relativistic
Celestial Mechanics
of the Solar System
Viewed down from
north ecliptic pole
Pioneer 10
Voyager 2
Pioneer 11
Voyager 1
Jupiter
Mars
Saturn
Uranus
Neptune
Pluto
Earth
Sun

The Authors

Prof. Sergei Kopeikin
University of Missouri
Department of Physics and Astronomy
Columbia, Missouri, USA
KopeikinS@missouri.edu

Dr. Michael Efroimsky
US Naval Observatory
3450 Massachusetts Ave NW
Washington, DC, USA
michael.efroimsky@usno.navy.mil

Dr. George Kaplan
Consultant to US Naval Observatory
3450 Massachusetts Ave NW
Washington, DC, USA
gk@gkaplan.us

Cover Picture

Pioneer 10 artwork, Ames Research Center/Nasa Center, 2006

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data:
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.d-nb.de.

Typesetting le-tex publishing services GmbH, Leipzig
Cover Design Grafik-Design Schulz, Fußgönheim
Printing and Binding Fabulous Printers Pte Ltd, Singapore

Printed in Singapore
Printed on acid-free paper

ISBN Print 978-3-527-40856-6

ISBN oBook 978-3-527-63456-9
ISBN ePub 978-3-527-63457-6
ISBN ePDF 978-3-527-63458-3
ISBN Mobi 978-3-527-63459-0

Contents

Preface

The general theory of relativity was developed by Einstein a century ago, and since then it has become the standard theory of gravity, especially important to the fields of fundamental astronomy, astrophysics, cosmology, and experimental gravitational physics. Today, the application of general relativity is also essential for many practical purposes involving astrometry, navigation, geodesy, and time synchronization. Numerous experiments have successfully tested general relativity to a remarkable level of precision. Exploring relativistic gravity in the solar system now involves a variety of high-accuracy techniques, such as very long baseline radio interferometry, pulsar timing, spacecraft Doppler tracking, planetary radio ranging, lunar laser ranging, the global positioning system (GPS), torsion balances and atomic clocks.

Over the last few decades, various groups within the International Astronomical Union have been active in exploring the application of the general theory of relativity to the modeling and interpretation of high-accuracy astronomical observations in the solar system and beyond. A Working Group on Relativity in Celestial Mechanics and Astrometry was formed in 1994 to define and implement a relativistic theory of reference frames and time scales. This task was successfully completed with the adoption of a series of resolutions on astronomical reference systems, time scales, and Earth rotation models by 24th General Assembly of the IAU, held in Manchester, UK, in 2000. However, these resolutions form only a framework for the practical application of relativity theory, and there have been continuing questions on the details of the proper application of relativity theory to many common astronomical problems. To ensure that these questions are properly addressed, the 26th General Assembly of the IAU, held in Prague in August 2006, established the IAU Commission 52, "Relativity in Fundamental Astronomy". The general scientific goals of the new commission are:

- clarify the geometrical and dynamical concepts of fundamental astronomy within a relativistic framework,
- provide adequate mathematical and physical formulations to be used in fundamental astronomy,
- deepen the understanding of relativity among astronomers and students of astronomy,
- promote research needed to accomplish these tasks.

The present book is intended to make a theoretical contribution to the efforts undertaken by this Commission. The first three chapters of the book review the foundations of celestial mechanics, as well as those of special and general relativity. Subsequent chapters discuss the theoretical and experimental principles of applied relativity in the solar system. The book is written for graduate students and researchers working in the area of gravitational physics and its applications in modern astronomy. Chapters 1 to 3 were written by Michael Efroimsky and Sergei Kopeikin, chapters 4 to 8 by Sergei Kopeikin, and Chapter 9 by George Kaplan. Sergei Kopeikin also edited the overall text.

It hardly needs to be said that the Newtonian celestial mechanics is a very broad area. In Chapter 1, we have concentrated on derivation of the basic equations, on explanation of the perturbed two-body problem in terms of osculating and non-osculating elements, and on discussion of the gauge freedom in the six-dimensional configuration space of the orbital parameters. The gauge freedom of the configuration space has many similarities to the gauge freedom of solutions of the Einstein field equations in general theory of relativity. It makes an important element of the Newtonian theory of gravity, which is often ignored in the books on classic celestial mechanics.

Special relativity is discussed in Chapter 2. While our treatment is in many aspects similar to the other books on special relativity, we make a special emphasis on explanation of the Lorentz and Poincaré transformations, and on the appropriate transformation properties of geometric objects like vectors and tensors – such as the velocity, acceleration, force, electromagnetic field, etc.

Chapter 3 is devoted to general relativity. It explains the main ideas of the tensor calculus on curved manifolds, the theory of the affine connection and parallel transport, and the mathematical and physical foundations of Einstein's approach to gravity. Into this Chapter, we have also included topics which are not well-covered in standard books on general relativity: namely, the variational analysis on manifolds and the multipolar expansion of gravitational radiation.

Chapter 4 introduces a detailed theory of relativistic reference frames and time scales in an N-body system comprised of massive, extended bodies – like our own solar system. Here, we go beyond general relativity and base our analysis on the scalar-tensor theory of gravity. This allows us to extend the domain of applicability of the IAU resolutions on relativistic reference frames, which in their original form were applicable only in the framework of general relativity. We explain the principles of construction of reference frames, and explore their relationship with the solutions of the gravitational field equations. We also discuss the post-Newtonian multipole moments of the gravitational field, from the viewpoint of global and local coordinates.

Chapter 5 discusses the principles of derivation of transformations between reference frames in relativistic celestial mechanics. The standard parameterized post-Newtonian (PPN) formalism by K. Nordtevdt and C. Will operates with a single coordinate frame covering the entire N-body system, but it is insufficient for discussion of more subtle relativistic effects showing up in orbital and rotational motion of extended bodies. Consideration of such effects require, besides the global frame, the introduction of a set of local frames needed to treat properly each body and its internal structure and dynamics. The entire set of global and local frames allows us to to discover and eliminate spurious coordinate effects that have no physical meaning. The basic mathematical technique used in our theoretical treatment is based on matching of asymptotic post-Newtonian expansions of the solutions of the gravity field equations.

In Chapter 6, we discuss the principles of relativistic celestial mechanics of massive bodies and particles. We focus on derivation of the post-Newtonian equations of orbital and rotational motion of an extended body possessing multipolar moments. These moments couple with the tidal gravitational fields of other bodies, which makes the motion of the body under consideration very complicated. Simplification is possible if the body can be assumed spherically symmetric. We discuss the conditions under which this simplification can be afforded, and derive the equations of motion of spherically-symmetric bodies. These equations are solved in the case of the two-body problem, and we demonstrate the rich nature of the possible coordinate presentations of such a solution.

The relativistic celestial mechanics of light particles (photons) propagating in a time-dependent gravitational field of an N-body system is addressed in Chapter 7. This is a primary subject of relativistic astrometry, which became especially important for analysis of space observations from the Hipparcos satellite in the early 1990s. New astrometric space missions, orders of magnitude more accurate than Hipparcos, such as Gaia, SIM, JASMINE, etc., will require even more complete developments. Additionally, relativistic effects play an important role in other areas of modern astronomy, such as pulsar timing, very long baseline radio interferometry, cosmological gravitational lensing, etc. High-precision measurements of gravitational light bending in the solar system are among the most crucial experimental tests of the general theory of relativity. Einstein predicted that the amount of light bending by the Sun is twice that given by a Newtonian theory of gravity. This prediction has been confirmed with a relative precision about 0.01%. Measurements of light bending by major planets of the solar system allow us to test the dynamical characteristics of spacetime and draw conclusions about the ultimate speed of gravity, as well as explore the so-called gravitomagnetic phenomena.

Chapter 8 deals with the theoretical principles and methods of the high-precision gravimetry and geodesy, based on the framework of general relativity. A gravitational field and the properties of geocentric and topocentric reference frames are described by the metric tensor, which is obtained from the Einstein equations with the help of post-Newtonian iterations. By matching the asymptotic, post-Newtonian expansions of the metric tensor in geocentric and topocentric coordinates, we derive the relationship between the reference frames, and relativistic corrections to the Earth's force of gravity and its gradient. Two definitions of a relativistic geoid are discussed, and we prove that these geoids coincide under the condition of a constant rigid-body rotation of the Earth. We consider, as a model of the Earth's matter, the notion of the relativistic level surface of a self-gravitating perfect fluid. We discover that, under conditions of constant rigid rotation of the fluid and hydrostatic behavior of tides, the post-Newtonian equation of the level surface is the same as that of the relativistic geoid. In the conclusions to this chapter, a relativistic generalization of the Clairaut's equation is obtained.

Chapter 9 is a practical guide to the relativistic resolutions of the IAU, with enough background information to place these resolutions into the context of the late 20th century positional astronomy. These resolutions involve the definitions of reference systems, time scales, and Earth rotation models; and some of the resolutions are quite detailed. Although the recommended Earth rotation models have not been developed *ab initio* within the relativistic framework presented in the other resolutions (in that regard, there still exist some difficult problems to solve), their relativistic terms are accurate enough for all the current and near-future observational techniques. At that level, the Earth rotation models are consistent with the general relativity framework recommended by the IAU and considered in this book. The chapter presents practical algorithms for implementing the recommended models.

The appendices to the book contain a list of astronomical constants and the original text of the relevant IAU resolutions adopted by the IAU General Assemblies in 1997, 2000, 2006, and 2009.

Numerous colleagues have contributed to this book in one way or or another. It is a pleasure for us to acknowledge the enlightening discussions, which one or more of the authors had, on different occasions, with Victor A. Brumberg of the Institute of Applied Astronomy (St. Petersburg, Russia); Tianyi Huang and Yi Xie of Nanjing University (China); Edward B. Fomalont of the National Radio Astronomical Observatory (USA); Valeri V. Makarov, William J. Tangren, and James L. Hilton of the U.S. Naval Observatory; Gerhard Schäfer of the Institute of Theoretical Physics (Jena, Germany); Clifford M. Will of Washington University (St. Louis, USA); Ignazio Ciufolini of the Universitá del Salento and INFN Sezione di Lecce (Italy); and Patrick Wallace, retired from Her Majesty's Nautical Almanac Office (UK).

We also would like to thank Richard G. French of Wellesley College (Massachusetts, USA); Michael Soffel and Sergei Klioner of the Technical University of Dresden; Bahram Mashhoon of the University of Missouri-Columbia; John D. Anderson, retired from the Jet Propulsion Laboratory (USA); the late Giacomo Giampieri, also of JPL; Michael Kramer, Axel Jessner, and Norbert Wex of the Max-Planck-Institut für Radioastronomie (Bonn, Germany); Alexander F. Zakharov of the Institute of Theoretical and Experimental Physics (Moscow, Russia); the late Yuri P. Ilyasov from Astro Space Center of Russian Academy of Science; Michael V. Sazhin, Vladimir A. Zharov, and Igor Yu. Vlasov of the Sternberg Astronomical Institute (Moscow, Russia); and Vladimir B. Braginsky of Moscow State University (Russia) for their remarks and comments, all of which helped us to formulate properly the theoretical concepts and other material presented in this book.

The discussions among the members of the IAU Working Group on Relativity in Celestial Mechanics and Astrometry, as well as those within the Working Group on Nomenclature for Fundamental Astronomy, have also been quite valuable and have contributed to what is presented here. The numerous scientific papers written by Nicole Capitaine of Paris Observatory and her collaborators have been essential references. Victor Slabinski and Dennis D. McCarthy of the U.S. Naval Observatory, P. Kenneth Seidelmann of the University of Virginia, Catherine Y. Hohenkerk of Her Majesty's Nautical Almanac Office, and E. Myles Standish, retired from the Jet Propulsion Laboratory, reviewed early drafts of the material that became Chapter 9, and made many substantial suggestions for improvement.

We were, of course, influenced by many other textbooks available in this field. We would like to pay particular tribute to:

C. W. Misner, K. S. Thorne & J. A. Wheeler "Gravitation",
V. A. Brumberg "Essential Relativistic Celestial Mechanics",
B. F. Schutz "Geometrical Methods of Mathematical Physics".
M. H. Soffel "Relativity in Celestial Mechanics, Astrometry and Geodesy",
C. M. Will "Theory and Experiment in Gravitational Physics",

There are many other books and influential papers that are important as well, which are referenced in the relevant parts of the present book.

None of our aforementioned colleagues is responsible for any remaining errors or omissions in this book, for which, of course, the authors bear full responsibility.

Last, and by no means least, Michael Efroimsky and George Kaplan wish to thank John A. Bangert of the U.S. Naval Observatory for the administrative support, which he so kindly provided to the project in all of its stages. Sergei Kopeikin is sincerely grateful to the Research Council of the University of Missouri-Columbia for generous financial support (grants RL-08-027, URC-08-062B, SRF-09-012) that was essential for successful completion of the book.

— Sergei Kopeikin, Michael Efroimsky, George Kaplan —

General Notations

Greek indices $\alpha, \beta, \gamma, ...$ run from 0 to 3 and mark spacetime components of four-dimensional objects. Roman indices $i, j, k, ...$ run from 1 to 3 and denote components of three-dimensional objects (zero component belongs to time). Repeated indices mean the Einstein summation rule with respect to corresponding indices, for instance, $A^\alpha B_\alpha = A^0 B_0 + A^1 B_1 + A^2 B_2 + A^3 B_3$, $T^k{}_k = T^1{}_1 + T^2{}_2 + T^3{}_3$, etc.

Minkowski metric has signature $\eta_{\alpha\beta} = \mathrm{diag}(-1, +1, +1, +1)$. Kronecker symbol (the unit matrix) is denoted $\delta_{ij} = \mathrm{diag}(1, 1, 1)$. Levi-Civita fully-antisymmetric symbol is ε_{ijk} such that $\varepsilon_{123} = +1$. Kronecker symbol is used to rise and lower Roman indices. Complete metric tensor $g_{\alpha\beta}$ is used to rise and lower the Greek indices in exact tensor equations whereas the Minkowski metric $\eta_{\alpha\beta}$ is employed for rising and lowering indices in equations of the post-Newtonian and post-Minkowskian approximations.

Round brackets surrounding a group of Roman indices mean full symmetrization with a corresponding normalizing coefficient, for example,

$$A_{(ij)} \equiv \frac{1}{2!}\left(A_{ij} + A_{ji}\right) , A_{(ijk)} \equiv \frac{1}{3!}\left(A_{ijk} + A_{jki} + A_{kij} + A_{ikj} + A_{kji} + A_{jik}\right) ,$$

and so on. Square brackets around a group of Roman indices denote anti-symmetrization with a corresponding normalizing coefficient, that is

$$A_{[ij]} \equiv \frac{1}{2!}\left(A_{ij} - A_{ji}\right) , A_{[ijk]} \equiv \frac{1}{3!}\left(A_{ijk} + A_{jki} + A_{kij} - A_{ikj} - A_{kji} - A_{jik}\right)$$

, etc. Angular brackets surrounding a group of Roman indices denote the symmetric trace-free (STF) part of the corresponding three-dimensional object, for instance,

$$A_{<ij>} = A_{(ij)} - \frac{1}{3}\delta_{ij}A_{kk} , A_{<ijk>} = A_{(ijk)} - \frac{1}{5}\delta_{ij}A_{kpp} - \frac{1}{5}\delta_{jk}A_{ipp} - \frac{1}{5}\delta_{ik}A_{jpp} ,$$

and the general definition of STF tensor is discussed in section 3.10.2.

We also use multi-index notations, for example, $A_L \equiv A_{i_1 i_2 ... i_l}$, $B_{P-1} \equiv B_{i_1 i_2 ... i_{p-1}}$, $D_{<L>} = D_{<i_1 i_2 ... i_l>}$. Contraction over multi-indices is understood as follows, $A_L Q^L = A_{i_1 i_2 ... i_l} Q^{i_1 i_2 ... i_l}$, $P_{aL-1} T^{bL-1} = P_{a i_1 i_2 ... i_{l-1}} T^{b i_1 i_2 ... i_{l-1}}$, and so on. The sign ∂ in front of indices denotes a partial derivative with respect to a corresponding coordinate which is taken as many times as the number of indices following the ∂, for example, $\partial_\alpha \phi = \partial\phi/\partial x^\alpha$, $\partial_{\alpha\beta}\phi = \partial^2\phi/\partial x^\alpha \partial x^b$, where $\partial_0 \phi = c^{-1}\partial\phi/\partial t$, $\partial_i \phi = \partial\phi/\partial x^i$, and, similarly, $\nabla_\beta T^\alpha$ denotes a covariant derivative. The partial derivatives will be also denoted sometimes with a comma, for example, $F_{,\alpha} \equiv \partial F/\partial x^\alpha$, etc. L-order partial derivative with respect to spatial coordinates is denoted $\partial_L = \partial_{i_1 i_2 ... i_l} = \partial_{i_1} ... \partial_{i_l}$. Other conventions are introduced and explained as they appear in the text of the book. Particular symbols for various mathematical objects are given below.

Mathematical Symbols Used in the Book.

$\boldsymbol{e}_\alpha$ vector basis on manifold

$\tilde{\omega}^\alpha$ covector basis on manifold

$\Lambda^\alpha{}_{\beta'}$, $\Lambda^{\alpha'}{}_\beta$ the matrix of transformation from one basis on manifold to another

$g_{\mu\nu}$	physical (Jordan-Fierz frame) metric tensor
$\tilde{g}_{\mu\nu}$	conformal (Einstein frame) metric tensor
g	the determinant of $g_{\mu\nu}$
$\tilde{g}$	the determinant of $\tilde{g}_{\mu\nu}$
$\eta_{\mu\nu}$	the Minkowski (flat) metric tensor
$\mathfrak{G}^{\beta}_{\alpha\gamma}$	the affine connection
$\mathfrak{K}_{\alpha\beta\gamma}$	the contortion tensor
$\mathfrak{R}_{\alpha\beta\gamma}$	the Ricci rotation coeffieicnts
$\mathfrak{D}_{\alpha\beta\gamma}$	the deviation tensor
$T^{\alpha}_{\beta\gamma}$	the torsion tensor
$Q^{\alpha}_{\beta\gamma}$	the nonmetricity tensor
$\Gamma^{\alpha}_{\mu\nu}$	the Christoffel symbol
$R_{\mu\nu}$	the Ricci tensor
R	the Ricci scalar
$\tilde{R}_{\mu\nu}$	the conformal Ricci tensor
$T_{\mu\nu}$	the energy-momentum tensor of matter
$T = T^{\alpha}{}_{\alpha}$	the trace of the energy-momentum tensor
$t^{\alpha\beta}$	the canonical pseudotensor of gravitational field
$t^{\alpha\beta}_{LL}$	the pseudotensor of Landau and Lifshitz
$\Lambda^{\alpha\beta}$	the effective tensor of matter and gravitational field
ϕ	the scalar field
ϕ_0	the background value of the scalar field ϕ
ζ	the dimensionless perturbation of the scalar field
$\theta(\phi)$	the coupling function of the scalar field
$\Box_{\rm g}$	the Laplace-Beltrami operator
$\Box$	the D'Alembert operator in the Minkowski spacetime
ρ	the density of matter in the comoving frame
ρ^*	the invariant (Fock) density of matter
Π	the internal energy of matter in the comoving frame
$\pi^{\mu\nu}$	the tensor of (anisotropic) stresses of matter
u^{α}	the four-velocity of matter
v^{i}	the 3-dimensional velocity of matter in the global frame
ω	the asymptotic value of the coupling function $\theta(\phi)$
ω'	the asymptotic value of the derivative of the coupling function $\theta(\phi)$
c	the ultimate speed of general and special theories of relativity
ϵ	a small dimensional parameter, $\epsilon = 1/c$
$h_{\mu\nu}$	the metric tensor perturbation, $g_{\mu\nu} - \eta_{\mu\nu}$
$\overset{(n)}{h}_{\mu\nu}$	the metric tensor perturbation of order ϵ^n in the post-Newtonian expansion of the metric tensor
N	a shorthand notation for $\overset{(2)}{h}_{00}$
L	a shorthand notation for $\overset{(4)}{h}_{00}$
N_i	a shorthand notation for $\overset{(1)}{h}_{0i}$
L_i	a shorthand notation for $\overset{(3)}{h}_{0i}$

H_{ij}	a shorthand notation for $\overset{(2)}{h}_{ij}$
H	a shorthand notation for $\overset{(2)}{h}_{kk}$
$\tilde{N},\ \tilde{L}$	shorthand notations for perturbations of conformal metric $\tilde{g}_{\mu\nu}$
γ	the 'space-curvature' PPN parameter
β	the 'non-linearity' PPN parameter
η	the Nordtvedt parameter, $\eta = 4\beta - \gamma - 3$
G	the observed value of the universal gravitational constant
$\mathcal{G}$	the bare value of the universal gravitational constant
$x^\alpha = (x^0, x^i)$	the global coordinates with $x^0 = ct$ and $x^i \equiv \boldsymbol{x}$
$w^\alpha = (w^0, w^i)$	the local coordinates with $w^0 = cu$ and $w^i \equiv \boldsymbol{w}$
U	the Newtonian gravitational potential in the global frame
$U^{(A)}$	the Newtonian gravitational potential of body A in the global frame
U_i	a vector potential in the global frame
$U_i^{(A)}$	a vector potential of body A in the global frame
$\chi,\ \Phi_1, \dots, \Phi_4$	various special gravitational potentials in the global frame
$V,\ V^i$	potentials of the physical metric in the global frame
$\sigma,\ \sigma^i$	the *active* mass and current-mass densities in the global frame
$I_{<L>}$	the *active* mass multipole moments in the global frame
$S_{<L>}$	the *active* spin multipole moments in the global frame
$\bar{V}$	potential of the scalar field in the global frame
$\bar{\sigma}$	scalar mass density in the global frame
$\bar{I}_{<L>}$	scalar mass multipole moments in the global frame
$\tilde{V}$	gravitational potential of the conformal metric in the global frame
$\tilde{\sigma}$	the *conformal* mass density in the global frame
$\tilde{I}_{<L>}$	the *conformal* mass multipole moments in the global frame
$\mathbb{M}$	conserved mass of an isolated system
$\mathbb{P}^i$	conserved linear momentum of an isolated system
$\mathbb{S}^i$	conserved angular momentum of an isolated system
$\mathbb{D}^i$	integral of the center of mass of an isolated system
$\hat{A}$	symbols with the hat stand for quantities in the local frame
(B)	sub-index referring to the body and standing for the internal solution in the local frame
(E)	sub-index referring to the external with respect to (B) bodies and standing for the external solution in the local frame
(C)	sub-index standing for the coupling part of the solution in the local frame
P_L	external STF multipole moments of the scalar field
Q_L	external STF gravitoelectric multipole moments of the metric tensor
C_L	external STF gravitomagnetic multipole moments of the metric tensor
$Z_L,\ S_L$	other sets of STF multipole moments entering the general solution for the space-time part of the external local metric
$Y_L, B_L, D_L, E_L, F_L, G_L$	STF multipole moments entering the general solution for the space-space part of the external local metric
$\mathcal{V}_i,\ \Omega_i$	linear and angular velocities of kinematic motion of the local frame
v^i	3-dimensional velocity of matter in the local frame

$\mathcal{I}_L$ active STF mass multipole moments of the body in the local frame
σ_B active mass density of body B in the local frame
$\bar{\mathcal{I}}_L$ scalar STF mass multipole moments of the body in the local frame
$\bar{\sigma}_B$ scalar mass density of body B in the local frame
$\tilde{\mathcal{I}}_L$ conformal STF mass multipole moments of the body in the local frame
$\tilde{\sigma}_B$ conformal mass density of body B in the local frame
σ^i_B current mass density of body B in the local frame
S_L spin STF multipole moments of the body in the local frame
ξ^0, ξ^i Relativistic corrections in the post-Newtonian transformation of time and space coordinates
x^i_B, v^i_B, a^i_B position, velocity and acceleration of the body's center of mass with respect to the global frame
R^i_B $x^i - x^i_B(t)$, i.e. the spacial coordinates taken with respect to the center of mass of body B in the global frame
$\mathcal{A}$, $\mathcal{B}_{<L>}$ functions appearing in the relativistic transformation of time
$\mathcal{D}_{<L>}$, $\mathcal{F}_{<L>}$, $\mathcal{E}_{<L>}$ functions appearing in the relativistic transformation of spacial coordinates
$\Lambda^\beta{}_\alpha$ matrix of transformation between local and global coordinate bases in the post-Newtonian approximation scheme
$\gimel^\beta{}_\alpha$ matrix of the inverse transformation between local and global coordinate bases in the post-Newtonian approximation scheme
$\mathfrak{B}$, $\mathfrak{D}$, $\mathfrak{B}^i$, $\mathfrak{P}^i$, $\mathfrak{R}^i{}_j$ the terms in the post-Newtonian expansion of the matrix of transformation $\Lambda^\beta{}_\alpha$
$\bar{U}$, $\bar{U}^i$, etc. external gravitational potentials
$\bar{U}_{,L}(\boldsymbol{x}_B)$, $\bar{U}^i{}_{,L}(\boldsymbol{x}_B)$ l-th spatial derivative of an external potential taken at the center of mass of body B
$\mathcal{U}^{(B)}$ PN correction in the formula of matching of the local Newtonian potential
F^{ik} the matrix of relativistic precession of local coordinates with respect to global coordinates
$\mathcal{M}_*$, $\mathcal{J}^i_*$, $\mathcal{P}^i_*$ Newtonian-type mass, center of mass, and linear momentum of the body in the local frame
M general relativistic PN mass of the body in the local frame
$\mathcal{M}$ active mass of the body in the local frame
$\tilde{\mathcal{M}}$ conformal mass of the body in the local frame
$\mathcal{I}^{(2)}$ rotational moment of inertia of the body in the local frame
$\mathcal{N}^L$ a set of STF multipole moments
$\mathcal{P}^i$ PN linear momentum of the body in the local frame
$\Delta\dot{\mathcal{P}}^i$ scalar-tensor PN correction to $\dot{\mathcal{P}}^i$
$\tilde{\mathcal{M}}_{ij}$ conformal anisotropic mass of the body in the local frame
, $\Delta\mathbb{F}^i_N$, $\mathbb{F}^i_{pN}$, $\Delta\mathbb{F}^i_{pN}$ gravitational forces in the expression for Q_i
$\mathcal{S}^i$ the bare post-Newtonian definition of the angular momentum (spin) of a body
$\mathcal{T}^i$ the post-Newtonian torque in equations of rotational motion
$\Delta\mathcal{T}^i$ the post-Newtonian correction to the torque $\mathcal{T}^i$
$\Delta\mathcal{S}^i$ the post-Newtonian correction to the bare spin $\mathcal{S}^i$
$\mathcal{R}^i$ velocity-dependent multipole moments
$\mathcal{S}^i_+$ the (measured) post-Newtonian spin of the body

r radial space coordinate in the body's local frame, $r = |\boldsymbol{w}|$
Ω^j_B angular velocity of rigid rotation of the body B referred to its local frame
$I^{(2l)}_B$ l-th rotational moment of inertia of the body B
$\mathbb{I}^L_B$ multipole moments of the multipolar expansion of the Newtonian potential in the global coordinates
R_B $|\boldsymbol{R}_B|$, where $\boldsymbol{R}_B = \boldsymbol{x} - \boldsymbol{x}_B$
R^i_{BC} $x^i_C - x^i_B$
F^i_N, F^i_{EIH}, F^i_S, F^i_{IGR}, δF^i_{IGR} forces from the equation of motion of spherically-symmetric bodies
$\mathfrak{M}_B$ Nordtvedt's gravitational mass of the body B

Abbreviations and Symbols Frequently Used in Astronomy

α right ascension,
δ declination,
$\Delta\psi$ nutation in [ecliptic] longitude (usually expressed in arcseconds),
$\Delta\epsilon$ nutation in obliquity (usually expressed in arcseconds),
ϵ mean obliquity of date,
ϵ' true obliquity of date ($= \epsilon + \Delta\epsilon$),
ϵ_0 mean obliquity of J2000.0,
θ Earth Rotation Angle,
μas microarcecond ($= 10^{-6}$ arcsecond $\approx 4.8{\times}10^{-12}$ radian),
σ a non-rotating origin or, specifically, the Celestial Intermediate Origin,
$\boldsymbol{\sigma}$ unit vector toward a non-rotating origin or, specifically, the Celestial Intermediate Origin,
Υ the equinox,
$\boldsymbol{\Upsilon}$ unit vector toward the equinox,
$''$ arcsecond ($= 1/3600$ degree $\approx 4.8{\times}10^{-6}$ radian),
AU astronomical unit,
B frame bias matrix,
BCRS Barycentric Celestial Reference System[1],
BIPM Bureau International des Poids et Mesures,
C matrix for transformation from GCRS to E_σ,
cen century, specifically, the Julian century of 36525 days of 86400 seconds,
CIO Celestial Intermediate Origin[2],
CIP Celestial Intermediate Pole,
CIRS See E_σ
E_Υ instantaneous (true) equator and equinox of date,
E_σ Celestial Intermediate Reference System (CIRS),
E_ϖ Terrestrial Intermediate Reference System (TIRS),
$\mathcal{E}_\Upsilon$ equation of the equinoxes,
$\mathcal{E}_o$ equation of the origins,

1) "Barycentric" always refers to the solar system barycenter, the center of mass of all bodies in the solar system.
2) The abbreviation CIO was used throughout much of the 20th century to designate the Conventional International Origin, the reference point for the measurement of polar motion.

ESA European Space Agency,
FKn n^{th} Fundamental Catalog (Astronomisches Rechen-Institut, Heidelberg),
GAST Greenwich apparent sidereal time,
GCRS Geocentric Celestial Reference System,
GMST Greenwich mean sidereal time,
GPS Global Positioning System,
HCRF Hipparcos Celestial Reference Frame,
IAG International Association of Geodesy,
IAU International Astronomical Union,
ICRF International Celestial Reference Frame,
ICRS International Celestial Reference System,
IERS International Earth Rotation and Reference System Service,
ITRF International Terrestrial Reference Frame,
ITRS International Terrestrial Reference System,
IUGG International Union of Geodesy and Geophysics,
J2000.0 the epoch 2000 January 1, 12^{h} TT (JD 2451545.0 TT) at the geocenter[3)],
JD Julian date (time scale used should be specified),
JPL Jet Propulsion Laboratory,
mas milliarcsecond (= 10^{-3} arcsecond $\approx 4.8\times10^{-9}$ radian),
N nutation matrix (for transformation from mean to true system of date),
n unit vector toward the CIP (celestial pole),
NOVAS Naval Observatory Vector Astrometry Subroutines (software),
P precession matrix (for transformation from J2000.0 system to mean system of date)
W "wobble" (polar motion) matrix (for transformation from ITRS to E_ϖ),
$\mathbf{R}_1(\phi)$ rotation matrix to transform column 3-vectors from one cartesian coordinate system to another. Final system is formed by rotating original system about its own x-axis by angle ϕ (counterclockwise as viewed from the +x direction):

$$\mathbf{R}_1(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{pmatrix}$$

$\mathbf{R}_2(\phi)$ rotation matrix to transform column 3-vectors from one cartesian coordinate system to another. Final system is formed by rotating original system about its own y-axis by angle ϕ (counterclockwise as viewed from the +y direction):

$$\mathbf{R}_2(\phi) = \begin{pmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{pmatrix}$$

$\mathbf{R}_3(\phi)$ rotation matrix to transform column 3-vectors from one cartesian coordinate system to another. Final system is formed by rotating original system about its own

3) "J2000.0 system" is shorthand for the celestial reference system defined by the mean dynamical equator and equinox of J2000.0.

z-axis by angle ϕ (counterclockwise as viewed from the +z direction):

$$\mathbf{R}_3(\phi) = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

s CIO locator: the difference between two arcs on the celestial sphere, providing the direction toward the CIO,
SI Système International d'Unités (International System of Units),
SOFA Standards of Fundamental Astronomy (software),
T unless otherwise specified, time[4] in Julian centuries (36525 days of 86400 seconds) from JD 2451545.0 (2000 Jan 1.5),
$\mathrm{T}_{\mathrm{eph}}$ time argument of JPL planetary and lunar ephemerides,
TAI International Atomic Time,
TCB Barycentric Coordinate Time,
TCG Geocentric Coordinate Time,
TDB Barycentric Dynamical Time,
TIO Terrestrial Intermediate Origin [5],
TIRS See E_{ϖ},
TT Terrestrial Time,
UCAC USNO CCD Astrographic Catalog,
USNO U.S. Naval Observatory,
UT1 Universal Time (affected by variations in length of day),
UTC Coordinated Universal Time (an atomic time scale),
VLBI very long baseline [radio] interferometry,
WGS 84 World Geodetic System 1984,
$\left.\begin{matrix} X \\ Y \\ Z \end{matrix}\right\}$ components of $\mathbf{n}_{\mathrm{GCRS}}$, unit vector toward the CIP with respect to the GCRS,
$\left.\begin{matrix} x_p \\ y_p \end{matrix}\right\}$ standard polar motion parameters, defining location of the CIP in the ITRS.

4) The time scale used should be specified, otherwise TT is understood.
5) The fundamental reference points referred to here as the Celestial Intermediate Origin (CIO) and the Terrestrial Intermediate Origin (TIO) were called, respectively, the Celestial Ephemeris Origin (CEO) and the Terrestrial Ephemeris Origin (TEO) in the IAU resolutions of 2000. The IAU Working Group on Nomenclature for Fundamental Astronomy [Capitaine *et al.*, 2007] has recommended the change of nomenclature with no change in the definitions. The new terminology is already in use in *The Astronomical Almanac* [2010] and in IERS documents. It is used throughout this book, except in the verbatim text of the IAU resolutions in appendix C.

1
Newtonian Celestial Mechanics

1.1 Prolegomena. Classical Mechanics in a Nutshell

1.1.1 Kepler's Laws

By trying numerous fits on a large volume of data collected earlier by Tycho Brahe and his assistants, Kepler realized in early 1605 that the orbit of Mars is not at all a circle, as he had expected, but is an ellipse with the Sun occupying one of its foci. This accomplishment of Kepler-astronomer was an affliction to Kepler-theologist, as it jeopardized his cherished theory of "celestial polyhedra" inscribed and circumscribed by spherical orbs, a *Mysterium Cosmographicum* theory according to which the planets were supposed to describe circles. For theological reasons, Kepler never relinquished the polyhedral-spherist cosmogony. Years later, in 1621, he re-worked the *Mysterium Cosmographicum* model in an attempt to reconcile it with elliptic trajectories.

Although the emergence of ellipses challenged Kepler's belief in the impeccable harmony of the celestial spheres, he put the scientific truth first, and included the new result into his book "*Astronomia Nova*". Begun in early 1600, the treatise saw press only in 1609 because of four-year-long legal disputes over the use of the late Tycho Brahe's observations, the property of Tycho's heirs. The most cited paragraphs of that 390-page-long volume are Kepler's first and second laws of planetary movement. In the modern formulation, the laws will read:

- ◇ The planets move in ellipses with the Sun at one focus.
- ◇ A vector directed from the Sun to a planet sweeps out equal areas in equal times.

These celebrated conjectures should not overshadow another revolutionary statement pioneered in "*Astronomia Nova*" - the hypothesis that the Sun is not stationary in space but describes a trajectory across the stars. Pioneering this idea, Kepler-astronomer again came into a conflict with Kepler-theologian. The heliocentric views of Kepler rested on a religious basis. Kepler was convinced that the universe was an image of God, with the Sun corresponding to the Father, the "stellar sphere" to the Son, and the enclosed space to the Holy Spirit. Kepler's hypothesis that the Sun could travel relative to the stars indicates how his scientific insight was overpowering his theological constructions.

Another famous book by Kepler, "*Harmonices Mundi*", saw light a decade later, in 1619. In the final volume of that treatise, Kepler publicized his finding that the ratio of the cubed semi-major axis to the squared orbital period is the same for all planets. In modern terms, this, third law of Kepler is usually formulated as:

- ◇ The cube of a planet's semi-major axis is proportional to the square of its orbital period: $a^3 \sim T^2$.

This way, for a planet with period T_1 and semi-major axis a_1, and a planet with period T_2 and semi-major axis a_2, the following relation takes place: $(a_1/a_2)^3 = (T_1/T_2)^2$.

1.1.2
Fundamental Laws of Motion: from Descartes, Newton, and Leibniz to Poincaré and Einstein

The next milestone contribution to the science of mechanics was offered a quarter-century later by René Descartes.

Basing his reasoning on the scholastic argument that nothing moves by virtue of its own nature towards its opposite or towards its own destruction, Descartes [1644] in his "*Principles of Philosophy*" came up with three laws of bodily motion. The first of those stated "that each thing, as far as is in its power, always remains in the same state; and that consequently, when it is once moved, it always continues to move". The second law held that "all movement is, of itself, along straight lines". The third law was an attempt to describe colliding bodies and to introduce a conserving quantity.

The first two laws of Descartes, together, constitute what is currently termed *the law of inertia* or the first law of Newton. Indeed, the wording of the law of inertia, suggested by Newton [1760] in his "*Principia*", was an exact equivalent to the first and second laws by Descartes.

First Law: *Every body persists in its state of being at rest or of moving uniformly straight forward, except insofar as it is compelled to change its state by force impressed.*

There are two parts to this statement - one which predicts the behavior of stationary objects and the other which predicts the behavior of moving objects. Mathematical formulation of the first law of Newton demands introduction of new concepts, the absolute time t and the absolute space endowed with a special class of coordinates, $\boldsymbol{x} = (x^i) = (x^1, x^2, x^3)$, introduced in space - the so-called, inertial reference frames. Motion of the body in space is described by vector function $\boldsymbol{x}(t)$. The first law of Newton simply states that velocity of the body $\boldsymbol{v} = d\boldsymbol{x}/dt$ is nil or remains constant in the inertial reference frame, if the body is not subject to the action of a net force. In both cases, acceleration, $\boldsymbol{a} = d\boldsymbol{v}/dt$, of a freely moving or static body vanishes in the inertial frame,

$$\boldsymbol{a} = 0\ . \tag{1.1}$$

Equation (1.1) makes it evident that there is a multitude of the inertial frames moving with respect to each other with constant velocities.

This *law of inertia*, however, marked the point, beyond which Newton's thought sharply diverged from Descartes' heritage. In formulating the rules of dynamics, Newton succeeded where Descartes had failed. The third law of Descartes, while marking one of the first attempts to locate an invariant or unchanging feature of bodily interactions, was just short of what is now called the momentum conservation. Newton, on his part, proposed two other laws, the law of impulse and the law of reciprocal actions, which must be put together in order to ensure the conservation of body's momentum

$$\boldsymbol{p} = m\boldsymbol{v}\ , \tag{1.2}$$

where m is mass of the body, and $\boldsymbol{v}$ is its velocity. The law of impulse and the law of reciprocal actions are known as the second and third laws of Newton respectively. They are formulated as follows.

Second Law: *The time rate of change of body's linear momentum* $\boldsymbol{p}$ *is equal to the net force* $\boldsymbol{F}$ *exerted on the body,*

$$\frac{d\boldsymbol{p}}{dt} = \mathcal{F} \,. \tag{1.3}$$

The second law is valid in any frame of reference that is written in an invariant form that is valid in arbitrary frames of reference both inertial and non-inertial. Therefore, the force $\mathcal{F}$ splits algebraically in two parts - the force of inertia, $\boldsymbol{F}_{\rm in}$, and the external force, $\boldsymbol{F}$, so that

$$\mathcal{F} = \boldsymbol{F} + \boldsymbol{F}_{\rm in} \,. \tag{1.4}$$

The force of inertia $\boldsymbol{F}_{\rm in}$ exists only in non-inertial (accelerated and/or rotating) frames and has a pure kinematic origin, while the external force $\mathcal{F}$ describes real physical interaction between the body under consideration with other bodies. The external force $\boldsymbol{F}$ determines the dynamical part of the body's motion that is not related to the choice of the reference frame. If mass of the body is conserved, $dm/dt = 0$, the second law of Newton can be written in a more familiar form,

$$m\boldsymbol{a} = \mathcal{F} \,, \tag{1.5}$$

which establishes a more simple than equation (1.3), relationship between the acceleration of the body, its constant mass, and the force applied.

By setting $\mathcal{F} = 0$ in Newton's law of impulse (1.5) written for a body of a constant mass, one would arrive to the conclusion, $\boldsymbol{a} = 0$, and would get an impression that the law of inertia (1.1) is a special case of the law of impulse: a vanishing net force yields a zero acceleration. It looks like the law of inertia is redundant, and can be derived instead of to be postulated. Newton surely observed the possibility of such a conclusion, but nevertheless chose to add the law of inertia as a separate statement. Moreover, he placed this law first. The only reason why he could have done so was his intention to single out a special class of forces - the kinematic forces of inertia - from the rest of genuine dynamical interactions, and to introduce a special class of inertial reference frames in which the forces of inertia $\boldsymbol{F}_{\rm in}$ vanish so that the second law of Newton is reduced to pure dynamical form

$$\frac{d\boldsymbol{p}}{dt} = \boldsymbol{F} \,. \tag{1.6}$$

This logic makes it clear that the first Newton's law is not a tautology following the second Newton's law but a crucial element of the entire theory introducing a special class of reference frames excluding the inertial forces. It may look simple now but it took Newton's successors centuries to arrive to the modern formulation of the law of inertia:

There exist reference frames, called inertial, such that a particle at rest or with constant velocity in one inertial frame will remain at rest or have constant velocity in all inertial frames, provided the net external force $\boldsymbol{F}$ acting on the particle is nil.

Crucial in this formulation is that it deliberately omits any mentioning of the absolute motion. This is because all inertial frames are effectively equivalent in the sense that the second law of Newton (1.6) is invariant (remains the same) irrespectively of the choice of the inertial frame.

Descartes-Newton's idea of inertia differs from the modern understanding of this phenomenon, in that they both regarded uniform motion and rest as different bodily states. Of a special interest is the position of Descartes who was partially *relationist* and partially *absolutist* [1]. On the one hand, he argued that space and matter are inseparable aspects of one phenomenon, and that motion is always the motion of bodies relative to one another. On the other hand, despite holding motion to be relational, Descartes also held there to be a privileged sense of motion ("true motion") over and above the merely relative motions. In distinction from Descartes, Newton was a pure *absolutist* whose system of views consistently stemmed from his belief that space (and, likewise, time) has existence of its own, independently of the bodies residing in it. The concept of *absolute space and time* proposed in "*Principia*", laid a foundation for a version of the æther theory developed by Newton in his "*Opticks*", a book in which he proposed a corpuscular theory of light. As the theory had trouble explaining refraction, Newton claimed that an "æthereal medium" was responsible for this effect. He then went further to suggest it might be responsible for other physical effects such as heat transfer.

The law of conservation of the linear momentum of a closed system of mutually-interacting bodies required from Newton to postulate his

Third Law: Whenever a particle exerts a force, $\boldsymbol{F}_{12}$, on another particle, the latter one *simultaneously* exerts a force, $\boldsymbol{F}_{21}$ on the former, with the same magnitude and in the opposite direction,

$$\boldsymbol{F}_{12} = -\boldsymbol{F}_{21} \ . \tag{1.7}$$

Be mindful that although the magnitude of the forces are equal, the accelerations of the bodies are not: the less massive body will have a greater acceleration due to Newton's second law (1.5). Let us apply the third law to a system of two interacting particles having instantaneous linear momenta, $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$, respectively. The second

1) Absolutism is a philosophical paradigm, according to which space and time are fundamental entities existing independently of matter. Relationism denies this paradigm so that space and time exists only as a supplementary mathematical tool to express relationships between the material bodies (and material fields). Relationism denies existence of privileged coordinates but may admit that some configurations of the bodies (fields) may have a privileged value for observer.

Newton's law for two particles written down in an inertial reference frame, is

$$\frac{d\boldsymbol{p}_1}{dt} = \boldsymbol{F}_{12}\,, \tag{1.8a}$$

$$\frac{d\boldsymbol{p}_2}{dt} = \boldsymbol{F}_{21}\,. \tag{1.8b}$$

Adding equations (1.8a) and (1.8b) together, and applying the third law of Newton (1.7), yields

$$\frac{d}{dt}(\boldsymbol{p}_1 + \boldsymbol{p}_2) = 0\,, \tag{1.9}$$

that is equivalent to the statement that the total linear momentum of the system, $\boldsymbol{p} = \boldsymbol{p}_1 + \boldsymbol{p}_2$ is constant in any inertial reference frame. One has to pay attention that this constancy of the overall linear momentum is preserved only in the inertial frames since in a non-inertial frame the overall momentum may be not conserved because the inertial forces, $\boldsymbol{F}_{\rm in}$, may not obey the third law of Newton. The law (1.9) can be extended to a system of N interacting particles.

An important property of the force postulated in the third law of Newton is that physical interaction between bodies is instantaneous. It was in a perfect agreement with experimental situation at the time of Newton. Much later, after development of electrodynamics, special and general theories of relativity, and other field theories, it became clear that there must be no instantaneous forces in nature. This does not undermine validity of the Newtonian mechanics which remains fully self-consistent and works very well in the limit of low velocities and small accelerations. Post-Newtonian celestial mechanics in the solar system can be treated in most cases on the premise of the instantaneous gravitational interaction. Only dedicated experiments require to include the finite speed of propagation of gravity to get theoretical predictions consistent with observations [2)].

The idea of absolute motion was challenged in that same XVII-th century by Leibniz, Huygens, and other *relationists*. Leibniz found the Newtonian notion of absolute space unacceptable, because two universes whose bodies occupied different absolute positions but identical relative positions would be indistinguishable from one another. Despite the objections, the indisputable authority of Newton kept the theory of æther afloat for more than two hundred years. As Ferraro [2007] put it: "*The controversy between relationists and absolutists quieted down in the following centuries, due to the success of the Newtonian science. Actually it remained in a latent state, because it would resurge at the end of the nineteenth century.*" The issue indeed returned into the center of attention in 1887, after the Michelson-Morley experiment challenged the possibility of æther's existence. The discussion continued and even spilled out into the XX-th century, with Poincaré and Einstein being on the opposite sides of the fence.

Although in his Saint Louis address of 1904, Poincaré came up with an early formulation of the relativity principle, he never granted this principle the fundamental status Einstein gave it in 1905 [Kobzarev, 1975]. Defending the idea of æther,

2) See section 7.9 for particular details and explanations.

Poincaré believed that some dynamical effects conspire to prevent us from observing it by mechanical or electromagnetic means. On the 11 of April 1912, three months before his death, Poincaré gave to the French Society of Physics a talk entitled "The Relations between Matter and Æther". This duality of Poincaré's concept of motion brings up strong parallels with Descartes.

Einstein, on his part, strictly followed the line of Leibniz, rejecting the absolute motion. Thus, he had no need to introduce æther in special relativity, because in special relativity this entity, in its classical meaning, was redundant. Years later, after general relativity was developed, Einstein [1920a] admitted in his Leiden's address that the general theory of relativity does not yet compel us to abandon æther. Einstein said that "*according to the general theory of relativity space is endowed with physical qualities (the metric potentials); in this sense, therefore, there exists an æther.*" At the same time, Einstein acknowledged that this general-relativistic meaning of the word "*differs widely from that of the æther of the mechanical undulatory theory of light*". At any rate, Einstein's interpretation deprives æther of its ability to define a reference frame.

Einstein's viewpoint was later corroborated by Dirac, on the grounds of hole theory of vacuum. According to Dirac [1951], vacuum is a substance of complex structure and therefore may be regarded as a physical medium, a kind of æther. However this medium is Lorentz-invariant and thus defines no preferred inertial frame of reference. This viewpoint has now become conventional in quantum physics [Lee, 1981]. The reason why the term *æther* is seldom applied to quantum vacuum in the modern literature is the necessity to avoid confusion with the old concept of æther, one associated with absolute motion.

1.1.3
Newton's Law of Gravity

Having formulated the three fundamental laws of motion, Newton went on in his *Principia* to explore a particular force, gravity. In this endeavor, Newton was getting inspiration from the works by Kepler and from correspondence with Hooke [3].

In 1666, Robert Hooke explained to the Royal Society his concept on what made the planets describe closed orbits about the Sun. According to Hooke, a force was needed, not to push a planet along from behind, but to pull it in towards a fixed gravitating center, so as to make the planet describe a closed curve instead of moving off along a straight tangent line into outer space. Circa 1680, in his correspondence with Newton, Hooke hypothesized that above ground level the gravity force changes inversely as the square of the distance from the Earth's center, and that below ground the force falls off as the center is approached. Hooke enquired as to what curve should be followed by a body subject to a central force obeying the inverse-square law [4]. Soon Newton proved that an orbit in the form of a conic section, with the

3) For historical analysis of the life and work of Newton and the contemporary scholars, see [Cohen and Smith, 2002].

4) Newton, in his letter to Halley of 20 June 1686, seeking to rebut Hooke's claim to have provided him

center of attraction located in one of the foci, necessarily implies an inverse-square attraction [5] - a result perfectly fitting Kepler's first law. We shall never know what made Newton procrastinate for almost five years with making his calculation public. For the first time, the discovery saw light in the *Principia* in 1687. In Newton's own words,

> "I deduced that the forces, which keep the planets in their orbs must [be] reciprocally as the squares of their distances from the centers about which they revolve: and thereby compared the force requisite to keep the Moon in her Orb with the force of gravity at the surface of the Earth; and found them answer pretty nearly."

In the modern notations, the law will read (see Figure 1.1)

$$\boldsymbol{F}_{12} = -G\frac{m_1 m_2}{r_{12}^3}\boldsymbol{r}_{12} \;, \tag{1.10}$$

where $\boldsymbol{F}_{12}$ is the gravitational force wherewith body 2 acts on body 1, $G = (6.67259 \pm 0.00030) \times 10^{-11}\ \mathrm{m^3 kg^{-1} s^{-2}}$ denotes the Newtonian gravity constant, m_1 and m_2 are the masses of the two interacting bodies located at positions with spatial coordinates $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ respectively, the vector

$$\boldsymbol{r}_{12} = \boldsymbol{x}_1 - \boldsymbol{x}_2 \tag{1.11}$$

is aimed from the second body to the first one, the quantity r_{12} being this vector's Euclidean magnitude $r_{12} = |\boldsymbol{r}_{12}|$. According to the third Newton law of reciprocal action, body 1 acts on body 2 with a force

$$\boldsymbol{F}_{21} = -\boldsymbol{F}_{12} = -G\frac{m_1 m_2}{r_{21}^3}\boldsymbol{r}_{21} \;, \tag{1.12}$$

where $\boldsymbol{r}_{21} = \boldsymbol{x}_2 - \boldsymbol{x}_1 = -\boldsymbol{r}_{12}$, and $r_{21} = |\boldsymbol{r}_{21}| = |\boldsymbol{r}_{12}| = r_{12}$.

Combining the gravity law with Newton's second law (the law of impulse), one gets

$$\frac{d}{dt}(m_1 \dot{\boldsymbol{x}}_1) = -G\frac{m_1 m_2}{r_{12}^3}\boldsymbol{r}_{12} \;, \tag{1.13}$$

$$\frac{d}{dt}(m_2 \dot{\boldsymbol{x}}_2) = -G\frac{m_2 m_1}{r_{21}^3}\boldsymbol{r}_{21} \;, \tag{1.14}$$

with overdot standing for an ordinary time derivative. If the masses of the bodies are constant[6] equations (1.13 - 1.14) can be simplified to the form usually employed in

originally with the idea of inverse-square gravity law, emphasized that the idea had been published before by Boulliau. In fact, Boulliau did not believe in a universal attraction force. At the same time, in his book he indeed argued that, had such a force existed, it would likely obey the inverse-square law [Boulliau, 1645].

5) Be aware that Newton proved that orbits being conics entail the inverse-square law. He did *not* prove that the inverse-square radial force results in orbits being conics [Weinstock, 1982].

6) This is not always true. For example, mass of the Sun changes due to the emission of the solar wind and radiation. In many cases, however, the mass loss is slow and can be neglected.

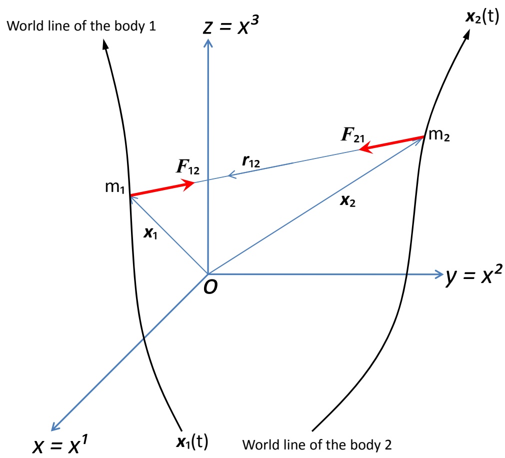

Figure 1.1 Newton's law of the universal gravitational attraction. The coordinates $x^i = (x, y, z)$ represent an arbitrary inertial reference frame with the origin O. This frame is assumed to be non-moving and its axes are non-rotating. Time t is absolute and parameterizes the worldlines of the bodies.

celestial mechanics,

$$m_1\ddot{\boldsymbol{x}}_1 = -G\frac{m_1m_2}{r_{12}^3}\boldsymbol{r}_{12}\ , \tag{1.15}$$

$$m_2\ddot{\boldsymbol{x}}_2 = -G\frac{m_2m_1}{r_{21}^3}\boldsymbol{r}_{21}\ . \tag{1.16}$$

Summing up equations (1.13) and (1.14), and integrating the result over the time, one arrives at the law of conservation of linear momentum of the gravitating two-body system,

$$m_1\dot{\boldsymbol{x}}_1 + m_2\dot{\boldsymbol{x}}_2 = \boldsymbol{P}\ , \tag{1.17}$$

where $\boldsymbol{P}$ is a constant vector of the linear momentum of the system. The *center of mass* of the two-body system is, by definition, a point given by the vector

$$\boldsymbol{X} = \frac{m_1\boldsymbol{x}_1 + m_2\boldsymbol{x}_2}{m_1 + m_2}\ . \tag{1.18}$$

Integration of the equation (1.17) with respect to time gives birth to a vector integral of motion of the center of mass,

$$M\boldsymbol{X} = \boldsymbol{D} + \boldsymbol{P}(t - t_0)\ . \tag{1.19}$$

Here

$$M = m_1 + m_2 \tag{1.20}$$

is a constant total mass of the two-body system,

$$\boldsymbol{D} = M\boldsymbol{X}_0\ , \tag{1.21}$$

and $\boldsymbol{X}_0$ is the constant position of the center of mass at the fiducial time, t_0, which is often called an *epoch* in dynamic astronomy. A constant vector

$$\boldsymbol{V} = \dot{\boldsymbol{X}} = \frac{\boldsymbol{P}}{M}\ , \tag{1.22}$$

is termed the *velocity of the center of mass*. Equation (1.19) tells us that the center of mass $\boldsymbol{X}$ of the two-body system moves uniformly along a straight line with the constant velocity $\boldsymbol{V}$.

Solving equations (1.11) and (1.18) elucidates that inertial coordinates of the bodies, $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, can be always represented as a sum of two vectors, $\boldsymbol{X}$ and $\boldsymbol{r}_{12}$,

$$\boldsymbol{x}_1 = \boldsymbol{X} + \frac{m_2}{m_1 + m_2}\boldsymbol{r}_{12}\ , \qquad \boldsymbol{x}_2 = \boldsymbol{X} - \frac{m_1}{m_1 + m_2}\boldsymbol{r}_{12}\ , \tag{1.23}$$

where $\boldsymbol{r}_{12}$ is a vector of a relative position of body 2 with respect to body 1. Substituting these equations to the equations of motion (1.15 and (1.16) and accounting for the conservation of the integral of the center of mass of the two-body system, $\ddot{\boldsymbol{X}} = 0$, one obtains the equations of relative motion

$$\mu\ddot{\boldsymbol{r}}_{12} = -G\frac{M\mu}{r_{12}^3}\boldsymbol{r}_{12}\ , \tag{1.24}$$

where

$$\mu = \frac{m_1 m_2}{M} , \tag{1.25}$$

is called the *reduced mass*. Equations of relative motion for vector $\boldsymbol{r}_{21}$ is obtained by exchanging the body indices, $1 \leftrightarrow 2$, but it does not provide us with a new information. Equations of relative motion (1.24) are naturally termed the equations of motion of the *reduced two-body problem*. They could be also derived by subtracting equation (1.16) from (1.15). It may make an impression that the conservation of the integral of the center of mass is not important in derivation of the equations of the relative motion. However, this point of view is mistaken since the integral of the center of mass were not exist, the equations of the relative motion would have extra terms associated with the force of inertia.

Inertial coordinates with the origin fixed at the center of mass of the gravitating system are named *barycentric*. In the barycentric frame of reference, the total momentum of the system is zero, $\boldsymbol{P} = 0$, while the position of the center of mass is constant and can be set to zero as well, $\boldsymbol{X}_0 = 0$. In this coordinate system, at any instant of time, one has, $\boldsymbol{X} = \boldsymbol{V} = 0$, as follows from the conservation of momentum. Hence, in this frame, equations (1.23) are simplified

$$\boldsymbol{x}_1 = \frac{m_2}{m_1 + m_2}\boldsymbol{r}_{12} , \qquad \boldsymbol{x}_2 = -\frac{m_1}{m_1 + m_2}\boldsymbol{r}_{12} , \tag{1.26}$$

and the problem of motion is reduced to solving of only one differential equation of the relative motion (1.24) that can be re-written as

$$\ddot{\boldsymbol{r}}_{12} = -G\frac{M}{r_{12}^3}\boldsymbol{r}_{12} , \tag{1.27}$$

since the reduced mass μ is canceled out. In the barycentric reference frame the problem of motion can be viewed as a problem of a single body with the reduced mass, μ, moving around a fixed center of gravity located at the barycenter of the two-body system and having a total mass, $M = m_1 + m_2$. In more general, N-body problem, a similar procedure of introducing the relative coordinates can be employed to reduce the problem to an $(N - 1)$-body setting.

1.2 The N-body Problem

Let us consider an isolated self-gravitating system consisting of a number of point-like massive bodies. In neglect of the gravitational pull from the Milky Way and the Hubble expansion of the universe, the solar system is a typical example but the consideration given in this chapter is applicable equally well to other isolated astronomical systems like a binary or multiple stellar system or a planetary system around other star. We shall enumerate the massive bodies with the capital letters $A, B, C, ...$ taking the values of $0, 1, 2, ..., N$, the index 0 being assigned to the primary body. Depending on a particular situation under consideration the primary can be either the Sun, or the Earth, or any other major planet.

1.2.1 Gravitational Potential

Let us begin from discussion of gravitational potential of a point-like mass m located at the origin of an inertial reference frame ($\boldsymbol{x} = (x^i) = (x^1, x^2, x^3)$). Gravitational force of the mass on a test particle of *unit* mass is given by expression

$$f^i = -\frac{Gm}{r^3}x^i \,, \tag{1.28}$$

where, $r = |\boldsymbol{x}| = \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}$, is the Euclidean distance from the mass to the field point $\boldsymbol{x} = (x^i)$. Elementary gravitational force, f^i, can be represented as a gradient of gravitational potential

$$\phi = \frac{Gm}{r} \,, \tag{1.29}$$

so that

$$f^i = \frac{\partial\phi}{\partial x^i} = \left(\frac{\partial\phi}{\partial x^1}, \frac{\partial\phi}{\partial x^2}, \frac{\partial\phi}{\partial x^3}\right) . \tag{1.30}$$

If the mass m is displaced to the point with coordinates $\boldsymbol{x}' = (x'^i)$, the gradient expression (1.30) for gravitational force, f^i, remains the same but the value of the potential, ϕ, at the point x^i becomes

$$\phi = \frac{Gm}{|\boldsymbol{x} - \boldsymbol{x}'|} \,. \tag{1.31}$$

Let us now consider an extended massive body made up of a continuous distribution of matter having a compact support (enclosed in a finite volume) with a mass density $\rho(t, \boldsymbol{x})$. One assumes that the body's matter can move that explains the time dependence of the mass density which obeys the equation of continuity

$$\frac{\partial\rho}{\partial t} + \frac{\partial(\rho v^i)}{\partial x^i} = 0 \,, \tag{1.32}$$

where $v^i = v^i(t, \boldsymbol{x})$ is velocity of an infinitesimally-small element of body's matter, and the repeated Roman indices mean the Einstein rule of summation from 1 to 3. Einstein's summation rule was invented to avoid the explicit (but in many cases unnecessary) appearance of the sign of summation, $\sum$, in tensor equations. It tacitly assumes that a pair of repeated (dummy) indices assume summation over corresponding values of the indices. In what follows, one will uses the Einstein rule for summation of coordinate indices. For example, a scalar product of two vectors $\boldsymbol{a} = (a^i)$ and $\boldsymbol{b} = (b^i)$ will be written as $\boldsymbol{a}\cdot\boldsymbol{b} = a^i b^i \equiv \sum_{i=1}^{3} a^i b^i = a^1 b^1 + a^2 b^2 + a^3 b^3$.

Integration of equation (1.32) over the finite volume $\mathcal{V}$ of the body, tells us that the overall mass of the body

$$M = \int_{\mathcal{V}} \rho(t, \boldsymbol{x}) d^3 x \,, \tag{1.33}$$

is constant. Indeed, taking the time derivative from both sides of equation (1.33) yields

$$\frac{dM}{dt} = \frac{d}{dt}\int_{\mathcal{V}} \rho(t, \boldsymbol{x}) d^3 x = \int_{\mathcal{V}} \frac{\partial \rho}{\partial t} d^3 x + \oint_{\partial\mathcal{V}} \rho v^i dS_i \,, \tag{1.34}$$

where the surface integral is taken over the body's surface, $\partial\mathcal{V}$, and accounts for the presumable time-dependence of the boundary of integration due to the motion of matter. Applying now the Gauss theorem to the surface integral recasts equation (1.34) to

$$\frac{dM}{dt} = \int_{\mathcal{V}} \left[\frac{\partial \rho}{\partial t} + \frac{\partial(\rho v^i)}{\partial x^i}\right] d^3 x = 0 \,, \tag{1.35}$$

due to the equation of continuity. It proves that the overall mass of the extended body remains constant that is mass of the body is the integral of motion of matter.

Gravitational potential U of the extended body is found as an integral taken over the body's volume comprised of the contributions of the "elementary" potentials (1.31) of the "point-like" elements of the body with mass m replaced with $m \to \rho d^3 x$. It yields

$$U(t, \boldsymbol{x}) = \int_{\mathcal{V}} \frac{\rho(t, \boldsymbol{x}') d^3 x'}{|\boldsymbol{x} - \boldsymbol{x}'|} \,. \tag{1.36}$$

Equation for the gravitational force, F^i, exerted by the extended body on a probe unit mass at point x^i is defined as a gradient of the gravitational potential (1.36), that is

$$F^i = \frac{\partial U}{\partial x^i} = \left(\frac{\partial U}{\partial x^1}, \frac{\partial U}{\partial x^2}, \frac{\partial U}{\partial x^3}\right) \,. \tag{1.37}$$

By taking second-order partial derivatives from the potential U, one can prove that the gravitational potential, U, obeys the second-order partial differential equation

$$\delta^{ij} \frac{\partial^2 U}{\partial x^i \partial x^j} = -4\pi G \rho \,, \tag{1.38}$$

where $\delta^{ij} = \mathrm{diag}(1,1,1)$ is the unit matrix, and one has used the Einstein summation rule to avoid the appearance of the double sum of summation, $\sum_{i=1}^{3}\sum_{j=1}^{3}$, in the right side of this equation. Differential operator acting on gravitational potential U in the left side of this equation, is called the Laplace operator

$$\triangle \equiv \delta^{ij}\frac{\partial^2}{\partial x^i \partial x^j}\ , \tag{1.39}$$

and the equation (1.38) is known as the Poisson equation, conventionally written as

$$\triangle U = -4\pi G\rho\ . \tag{1.40}$$

Gravitational potential (1.36) is solution of the (inhomogeneous) Poisson equation that is valid both inside and outside of the body's volume. However, if one is interested in the gravitational potential and force only outside of the body, a homogeneous Laplace equation

$$\triangle U = 0\ , \tag{1.41}$$

is sufficient.

1.2.2
Gravitational Multipoles

In many practical tasks of celestial mechanics and geodesy one does not need the integral form of the Newtonian gravitational potential but its multipolar decomposition describing gravitational field in terms of multi-index objects called multipoles. It can be obtained by expanding gravitational potential (1.36) outside of the body into infinite Taylor series by making use of decomposition of the reciprocal distance $|\boldsymbol{x}-\boldsymbol{x}'|^{-1}$ around the point $x'^i = 0$ with respect to the, so-called, harmonic polynomials. One has [7],

$$\begin{aligned}\frac{1}{|\boldsymbol{x}-\boldsymbol{x}'|} &= \sum_{l=0}^{\infty}\frac{(-1)^l}{l!}x'^{<i_1}x'^{i_2}...x'^{i_l>}\frac{\partial^l}{\partial x^{i_1}\partial x^{i_2}...\partial x^{i_l}}\left(\frac{1}{r}\right) \\ &= \frac{1}{r}+\frac{x^i x'^i}{r^3}+\frac{1}{2}\left[\frac{3(x^i x'^i)(x^j x'^j)}{r^5}-\frac{r'^2}{r^3}\right]+\ldots\ ,\end{aligned} \tag{1.42}$$

where $r = |\boldsymbol{x}|$, $r' = |\boldsymbol{x}'|$, the multi-index notation of spatial indices has been used, each index $i_1, i_2, ..., i_l$ runs from 1 to 3, and the angular brackets around the indices denote an algebraic operation making an object having such indices, symmetric and traceless (STF) tensor with respect to spatial rotations (see Appendix A). This expansion is more commonly written in terms of the Legendre polynomials $P_l(\cos\theta)$. Formula (1.42) can be easily converted to these polynomials after re-writing the scalar product of two vectors $x^i x'^i$ in trigonometric form, $x^i x'^i = rr'\cos\theta$, and substituting it to

7) Notice the usage of Einstein's summation rule for indices $i_1, i_2, ..., i_l$ numerating spatial coordinates.

expression (1.42). It yields,

$$\frac{1}{|\boldsymbol{x}-\boldsymbol{x}'|} = \frac{1}{r}\sum_{l=0}^{\infty}\frac{(-1)^l}{l!}\left(\frac{r'}{r}\right)^l P_l(\cos\theta)\ . \tag{1.43}$$

Though this form of expansion of the reciprocal distance looks more simple, it requires further elaboration of $P_l(\cos\theta)$ in terms of the associated Legendre functions $P_l^m(\cos\theta)$ which is effectively equivalent to the expansion in terms of the harmonic polynomials. The harmonic polynomials has many mathematical advantages in theoretical studies [Hartmann *et al.*, 1994], and will be preferred almost everywhere in this book to describe the multipolar decompositions of gravitational potentials in classical and relativistic gravity theories.

After substituting the Taylor expansion (1.42) in the definition (1.36) of the Newtonian gravitational potential, one obtains its multipolar expansion

$$U(t,\boldsymbol{x}) = G\sum_{l=0}^{\infty}\frac{(-1)^l}{l!} I^{<i_1 i_2 \dots i_l>}\partial_{i_1 i_2 \dots i_l}\left(\frac{1}{r}\right)\ , \tag{1.44}$$

where one has used a shorthand notation for partial derivatives $\partial_i = \partial/\partial x^i$, $\partial_{i_1 i_2 \dots i_l} = \partial_{i_1}\partial_{i_2}\dots\partial_{i_l}$, and $I^{<i_1 i_2 \dots i_l>}$ are the mass multipole moments of gravitational field of the body which are integrals from the density, $\rho(t,\boldsymbol{x})$, taken over the body's volume

$$I^{<i_1 i_2 \dots i_l>} = \int_{\mathcal{V}}\rho(t,\boldsymbol{x})x^{<i_1}x^{i_2}\dots x^{i_l>}d^3x\ , \tag{1.45}$$

with $x^{<i_1}x^{i_2}\dots x^{i_l>}$ representing the harmonic polynomial of the l-th degree. The angular brackets around the indices of the polynomial denote a special kind of symmetry which is imposed on the harmonic polynomial by the condition that it must be a solution of a homogeneous Laplace equation, that is

$$\triangle\left(x^{<i_1}x^{i_2}\dots x^{i_l>}\right) = 0\ . \tag{1.46}$$

It is this condition which demands for the l-th order harmonic polynomial, $x^{<i_1}x^{i_2}\dots x^{i_l>}$, to be a fully-symmetric and trace-free (STF) tensor with respect to rotations in three-dimensional Euclidean space [Blanchet and Damour, 1989; Pirani, 1965; Thorne, 1980]. The word "trace-free" means that contraction of any pair of indices nullifies the STF tensor,

$$x^{<i_1}x^{i_1}x^{i_3}\dots x^{i_l>} = \delta_{i_1 i_2}x^{<i_1}x^{i_2}\dots x^{i_l>} \equiv 0\ , \tag{1.47}$$

which is a mathematical property of any polynomial solution of the homogeneous Laplace equation (1.46). We provide more details on the structure of the harmonic polynomials in appendix A.

Let us define a symmetric moment of inertia of the l-th order,

$$I^{i_1 i_2 \dots i_l} = \int_{\mathcal{V}}\rho(t,\boldsymbol{x})x^{i_1}x^{i_2}\dots x^{i_l}d^3x\ . \tag{1.48}$$

Then, the STF multipole moment (1.45) is expressed in terms of the moments of inertia as follows [Pirani, 1965]

$$I^{<i_1 i_2 \ldots i_l>} = \sum_{k=0}^{[l/2]} a_k^l \delta^{(i_1 i_2} \delta^{i_3 i_4} \cdots \delta^{i_{2k-1} i_{2k}} I^{i_{2k+1} \ldots i_l) j_1 j_1 \ldots j_k j_k} \;, \tag{1.49}$$

where the round brackets around a group of indices denote full symmetrization with respect to permutation of the indices, $[l/2]$ denotes the integer part of $l/2$, the repeated indices denote Einstein's summation, and the numerical coefficient

$$a_k^l = (-1)^k \frac{l!}{(l-2k)!(2k)!!} \frac{(2l-2k-1)!!}{(2l-1)!!} \;. \tag{1.50}$$

The STF multipole moments $I^{<i_1 i_2 \ldots i_l>}$ are well-known in celestial mechanics, and other ares of theoretical physics. For example, the zero-order ($l = 0$) multipole moment I is simply a constant mass M of the body having been introduced in equation (1.33). Dipole moment ($l = 1$)

$$I^i = \int_V \rho(t, \boldsymbol{x}) x^i d^3 x \;, \tag{1.51}$$

defines position of the center of mass of the body with respect to the origin of the coordinates. The quadrupole moment ($l = 2$)

$$I^{<ij>} = \int_V \rho(t, \boldsymbol{x}) \left(x^i x^j - \frac{r^2}{3} \delta^{ij} \right) d^3 x \;, \tag{1.52}$$

and the multipole moments of higher-order provide an integral characteristics of various asymmetries in the distribution of matter inside the body with respect to its equatorial and meridional planes [8]. If the origin of coordinates is placed to the center of mass of the body, the dipole moment vanishes, $I^i = 0$, and the multipolar expansion (1.44) can be written as follows

$$U(t, \boldsymbol{x}) = \frac{GM}{r} + G \sum_{l=2}^{\infty} \frac{(-1)^l}{l!} I^{<i_1 i_2 \ldots i_l>} \partial_{i_1 i_2 \ldots i_l} \left(\frac{1}{r} \right) \;, \tag{1.53}$$

where mass M is constant, but the multipole moments $I^{<i_1 i_2 \ldots i_l>}$ ($l \geq 2$) can depend on time. In many cases, contribution of higher-order multipoles to the overall gravitational field is fairly small and can be neglected, thus, leaving only the first term in the right side of equation (1.53). Extended body with spherically-symmetric distribution of mass has no multipole moments at all, so that its gravitational potential

$$U = \frac{GM}{r} \;, \tag{1.54}$$

is the same as that of the point-like mass M placed at the coordinate origin.

8) Notice that contraction of two indices of the quadrupole moment gives $I^{<ii>} \equiv 0$, that is its trace is indeed zero.

1.2.3
Equations of Motion

Let us derive the Newtonian equations of motion of extended bodies comprising the N-body system under consideration. We introduce a global inertial reference frame with time t and spatial coordinates $\boldsymbol{x} = (x^i) = (x^1, x^2, x^3)$ and assume that each body A occupies a finite volume $\mathcal{V}_A$ of space. The interior distribution of matter is characterized by mass density $\rho = \rho(t, \boldsymbol{x})$ and by the symmetric tensor of stresses $\pi^{ij}(t, \boldsymbol{x}) = \pi^{ji}(t, \boldsymbol{x})$, which is reduced in case of a perfect fluid to an isotropic pressure $p = p(t, \boldsymbol{x})$ such that the trace of this tensor, $\pi^{ij} = p\delta^{ij}$. Macroscopic equations of motion of matter are [9)]

$$\frac{\partial(\rho v^i)}{\partial t} + \frac{\partial(\rho v^i v^j)}{\partial x^j} = -\frac{\partial \pi^{ij}}{\partial x^j} + \rho\frac{\partial U}{\partial x^i} , \qquad (1.55)$$

where $v^i = dx^i/dt$ is velocity of matter and, $U = U(t, \boldsymbol{x})$, is gravitational potential that is a linear superposition of potentials of all bodies of the system

$$U(t, \boldsymbol{x}) = \sum_{B=0}^{N} U_B(t, \boldsymbol{x}) , \qquad (1.56)$$

$$U_B(t, \boldsymbol{x}) = G \int_{\mathcal{V}_B} \frac{\rho(t, \boldsymbol{x}')}{|\boldsymbol{x} - \boldsymbol{x}'|} d^3x' . \qquad (1.57)$$

Making use of equation of continuity (1.32), the equations of motion (1.55) can be recast to the form

$$\rho\frac{dv^i}{dt} = -\frac{\partial \pi^{ij}}{\partial x^j} + \rho\frac{\partial U}{\partial x^i} , \qquad (1.58)$$

where the operator of the total derivative

$$\frac{d}{dt} \equiv \frac{\partial}{\partial t} + v^j \frac{\partial}{\partial x^j} , \qquad (1.59)$$

describes differentiation along the worldline of the element of body's matter. We define the dipole moment of body A with respect to a point with coordinates $\boldsymbol{x}_A = (x^i_A)$ by expression

$$I^i_A = \int_{\mathcal{V}_A} \rho(t, \boldsymbol{x}) \left(x^i - x^i_A\right) d^3x . \qquad (1.60)$$

The point x^i_A coincides with the center of mass of the body in case when $I^i_A = 0$, and one imposes and keep this condition for any instant of time. Hence, the time-dependent position $\boldsymbol{x}_A = \boldsymbol{x}_A(t)$ of the center of mass of body A is defined in the inertial coordinates by equation

$$m_A x^i_A = \int_{\mathcal{V}_A} \rho(t, \boldsymbol{x}) x^i d^3x , \qquad (1.61)$$

9) Observe the use of Einstein's summation rule in application to the repeated indices.

where

$$m_A = \int_{\mathcal{V}_A} \rho(t, \boldsymbol{x}) d^3x \,, \tag{1.62}$$

is a constant mass of the body A. Equations of orbital motion of body A can be obtained after double differentiation of both sides of equation (1.61) with respect to time and application of the macroscopic equation of motion of matter (1.58). For doing this calculation, an important formula giving a value of the time derivative of integral quantities, will be required. More specifically, for any smooth function, $f = f(t, \boldsymbol{x})$, multiplied with density $\rho = \rho(t, \boldsymbol{x})$, the following differentiation rule is valid

$$\frac{d}{dt}\int_{\mathcal{V}_A} \rho f d^3x = \int_{\mathcal{V}_A} \rho \frac{df}{dt} d^3x \,, \tag{1.63}$$

where the total time derivative in the right side must be understood in the sense of equation (1.59). The proof of this rule is rather straightforward. Time derivative of the integral is

$$\frac{d}{dt}\int_{\mathcal{V}_A} \rho f d^3x = \int_{\mathcal{V}_A} \left[\frac{\partial \rho}{\partial t} f + \rho \frac{\partial f}{\partial t}\right] d^3x + \oint_{\partial \mathcal{V}_A} \rho f v^i dS_i \,, \tag{1.64}$$

where the surface integral in the right side of this equation takes into account that the volume of the body changes as time passes on. Applying the equation of continuity (1.32) and the Gauss theorem one can bring equation (1.64) to the following form,

$$\begin{aligned} &\int_{\mathcal{V}_A} \left[\frac{\partial \rho}{\partial t} f + \rho \frac{\partial f}{\partial t}\right] d^3x + \oint_{\partial \mathcal{V}_A} \rho f v^i dS_i = \\ &\int_{\mathcal{V}_A} \left[-\frac{\partial(\rho v^i)}{\partial x^i} f + \rho \frac{\partial f}{\partial t} + \frac{\partial(\rho f v^i)}{\partial x^i}\right] d^3x \,, \end{aligned} \tag{1.65}$$

which is immediately reduced to the right side of equation (1.63) after applying the Leibnitz rule to the partial derivative

$$\frac{\partial(\rho f v^i)}{\partial x^i} = \frac{\partial(\rho v^i)}{\partial x^i} f + \rho v^i \frac{\partial f}{\partial x^i} \,. \tag{1.66}$$

Applying equation (1.63) two times to the center-of-mass definition (1.61), one obtains

$$m_A v_A^i = \int_{\mathcal{V}_A} \rho(t, \boldsymbol{x}) v^i d^3x \,, \tag{1.67}$$

$$m_A a_A^i = \int_{\mathcal{V}_A} \rho(t, \boldsymbol{x}) \frac{dv^i}{dt} d^3x \,, \tag{1.68}$$

where $v_A^i = dx_A^i/dt$ is velocity, and $a_A^i = dv_A^i/dt$ is acceleration of the body's center of mass respectively. Now, one replaces the time derivative, dv^i/dt, in the right side of integral in equation (1.68) with the macroscopic equations of motion (1.58), and split the gravitational potential, U, in two parts - internal and external,

$$U = U_A + \bar{U} , \tag{1.69}$$

where

$$\bar{U} = \sum_{\substack{B=0 \\ B\neq A}}^{N} U_B(t, \boldsymbol{x}) . \tag{1.70}$$

It yields

$$m_A a_A^i = -\int_{\mathcal{V}_A} \frac{\partial \pi^{ij}}{\partial x^j} d^3x + \int_{\mathcal{V}_A} \rho \frac{\partial U_A}{\partial x^i} d^3x + \int_{\mathcal{V}_A} \rho \frac{\partial \bar{U}}{\partial x^i} d^3x . \tag{1.71}$$

First term in the right side of this equation vanishes,

$$\int_{\mathcal{V}_A} \frac{\partial \pi^{ij}}{\partial x^j} d^3x = \oint_{\partial\mathcal{V}_A} \pi^{ij} dS_j = 0 , \tag{1.72}$$

because stresses disappear on the surface of each gravitating body [Landau and Lifshit's, 1959]. The integral of the derivative of the internal gravitational potential also vanishes,

$$\int_{\mathcal{V}_A} \rho \frac{\partial U_A}{\partial x^i} d^3x = -\int_{\mathcal{V}_A}\int_{\mathcal{V}_A} \rho(t, \boldsymbol{x})\rho(t, \boldsymbol{x}') \frac{x^i - x'^i}{|\boldsymbol{x} - \boldsymbol{x}'|^3} d^3x d^3x' = 0 , \tag{1.73}$$

due to anti-symmetry of the integrand with respect to exchange of coordinates, $\boldsymbol{x} \leftrightarrow \boldsymbol{x}'$. Thus, all internal forces exerted on the body, cancel out exactly, and equations of motion of the center of mass of body A are reduced to

$$m_A a_A^i = \int_{\mathcal{V}_A} \rho \frac{\partial \bar{U}}{\partial x^i} d^3x . \tag{1.74}$$

External potential, $\bar{U}$, can be expanded in Taylor series around the point $\boldsymbol{x}_A$ with respect to the harmonic (STF) polynomials,

$$\bar{U} = \sum_{l=0}^{\infty} \frac{1}{l!} r_A^{<i_1} r_A^{i_2} ... r_A^{i_l>} \partial_{i_1 i_2 ... i_l} \bar{U}(t, \boldsymbol{x}_A) , \tag{1.75}$$

where $r_A^i \equiv x^i - x_A^i$, the angular brackets around indices denote STF symmetrization defined in equation (1.49), the partial derivative $\partial_i = \partial/\partial x^i$, and the l-th partial derivative $\partial_{i_1 i_2 ... i_l} \bar{U}(t, \boldsymbol{x}_A) \equiv \left[\partial_{i_1 i_2 ... i_l} \bar{U}(t, \boldsymbol{x})\right]_{\boldsymbol{x}=\boldsymbol{x}_A}$. Appearance of the harmonic polynomials in this expansion is justified because the external potential $\bar{U}$ satisfies the Laplace equation: $\triangle\bar{U} = 0$. Hence, the symmetric polynomial $r_A^{<i_1} r_A^{i_2} ... r_A^{i_l>}$ must be apparently traceless. Substituting expansion (1.75) in equation (1.74), yields

$$m_A a_A^i = \sum_{l=0}^{\infty} \frac{1}{l!} I_A^{<i_1 i_2 ... i_l>} \partial_{i i_1 ... i_l} \bar{U}(t, \boldsymbol{x}_A) , \tag{1.76}$$

which is the equation of motion of body A given in terms of its (time-dependent) STF multipole moments

$$I_A^{<i_1 i_2 \dots i_l>} = \int_{\mathcal{V}_A} \rho(t, \boldsymbol{x}) r_A^{<i_1} r_A^{i_2} \dots r_A^{i_l>} d^3 x \,, \tag{1.77}$$

coupled with the partial derivatives of the external potential $\bar{U}$ taken at the center of mass of the body. Development of this theory is getting progressively complicated if one continue to keep all multipole moments of the bodies in equations of motion [Kopejkin, 1988a]. We shall show how to deal with these complications in Chapter 6.1. The present section is restricted with the case of spherically-symmetric bodies neglecting their tidal and rotational deformations. In such case, the external potential $\bar{U}$ is simplified to a linear superposition of potentials of point-like masses

$$\bar{U}(t, \boldsymbol{x}) = \sum_{\substack{B=0 \\ B \neq A}}^{N} \frac{G m_B}{|\boldsymbol{x} - \boldsymbol{x}_B|} \,, \tag{1.78}$$

where $\boldsymbol{x}_B = \boldsymbol{x}_B(t)$ are time-dependent positions of the external bodies $B \neq A$ defined by equation being similar to equation (1.60) where index A must be replaced with index B. Substituting potential (1.78) in equation (1.76) and assuming that body A is also spherically-symmetric (so that only $l = 0$ monopole term, $I_A \equiv m_A$, remains) one arrives to the final form of dynamical equations of motion of $N + 1$ point-like masses m_A located at coordinate positions $\boldsymbol{x}_A$,

$$m_A \ddot{\boldsymbol{x}}_A = - \sum_{\substack{B=0 \\ B \neq A}}^{N} \frac{G m_A m_B}{r_{AB}^3} \boldsymbol{r}_{AB} \,, \tag{1.79}$$

with vector $\boldsymbol{r}_{AB} = \boldsymbol{x}_A - \boldsymbol{x}_B$ being directed from body B to A, $r_{AB} = |\boldsymbol{r}_{AB}|$.

1.2.4
The Integrals of Motion

The system of equations (1.79) admits 10 integrals of motion: three integrals of the linear momentum $\boldsymbol{P}$; three integrals of the initial position of the center of mass, $\boldsymbol{X}_0$; three integrals of the angular momentum, $\boldsymbol{J}$, and one integral of the energy, E. The integrals of the linear momentum and the the center of mass are obtained by summing up equations (1.79) over all the bodies of the system, followed by integration with respect to time. The net gravitational force in the sum is reduced to zero due to the third Newton's law, so one obtains the following two vectorial integrals of motion:

$$\sum_{A=0}^{N} m_A \dot{\boldsymbol{x}}_A = \boldsymbol{P} \,, \tag{1.80}$$

$$\sum_{A=0}^{N} m_A \boldsymbol{x}_A = \boldsymbol{P}(t - t_0) + \boldsymbol{D} \,, \tag{1.81}$$

with t_0 being the epoch, and the constant vector

$$\boldsymbol{D} = M\boldsymbol{X}_0 \,, \tag{1.82}$$

where the total mass of the system

$$M = \sum_{A=0}^{N} m_A \,, \tag{1.83}$$

is constant. To obtain one more integral of motion, take the dot-product of equation (1.79) with the velocity $\dot{\boldsymbol{x}}_A$, with the subsequent summation over all the bodies of the system,

$$\sum_{A=0}^{N} m_A \ddot{\boldsymbol{x}}_A \cdot \dot{\boldsymbol{x}}_A = -\sum_{A=0}^{N} \sum_{\substack{B=0\\B\neq A}}^{N} \frac{G m_A m_B}{r_{AB}^3} \boldsymbol{r}_{AB} \cdot \dot{\boldsymbol{x}}_A \,. \tag{1.84}$$

With aid of the equalities

$$\sum_{A=0}^{N} \sum_{\substack{B=0\\B\neq A}}^{N} \frac{G m_A m_B}{r_{AB}^3} \boldsymbol{r}_{AB} \cdot \dot{\boldsymbol{x}}_A = \frac{1}{2} \sum_{A=0}^{N} \sum_{\substack{B=0\\B\neq A}}^{N} \frac{G m_A m_B}{r_{AB}^3} \boldsymbol{r}_{AB} \cdot \dot{\boldsymbol{r}}_{AB} \,, \tag{1.85}$$

and

$$\frac{G m_A m_B}{r_{AB}^3} \boldsymbol{r}_{AB} \cdot \dot{\boldsymbol{r}}_{AB} = -\frac{d}{dt} \frac{G m_A m_B}{r_{AB}} \,, \tag{1.86}$$

equation (1.84) becomes

$$\frac{dE}{dt} = 0 \,, \tag{1.87}$$

with E standing for a scalar integral of motion – the energy:

$$E = \frac{1}{2} \sum_{A=0}^{N} m_A \dot{\boldsymbol{x}}_A^2 - \frac{1}{2} \sum_{A=0}^{N} \sum_{\substack{B=0\\B\neq A}}^{N} \frac{G m_A m_B}{r_{AB}} \,. \tag{1.88}$$

Clearly, the first term in the right side makes up the kinetic energy of the bodies, while the second one represents the gravitational potential energy. The former being always positive, the latter is always negative[10].

The last integral of motion - the angular-momentum vector $\boldsymbol{J}$ - is derived from equation (1.79) by taking the cross product of both sides of this equation with the position vector $\boldsymbol{x}_A$, summing up over all the equations, and subsequent integration over time. This entails

$$\boldsymbol{J} = \sum_{A=0}^{N} m_A \left(\boldsymbol{x}_A \times \dot{\boldsymbol{x}}_A \right) \,, \tag{1.89}$$

10) The gravitational potential energy being negative makes the relativistic masses of self-gravitating astronomical objects, like planets or stars, smaller than the algebraic sum of the rest masses of their constituent particles - baryons. See section 6.1.3 for further details.

the sign $\times$ denoting the Euclidean cross product of two vectors. Constant vector $\boldsymbol{J}$ defines an invariant plane of the N-body problem called the *invariable plane of Laplace*.

1.2.5
The Equations of Relative Motion with Perturbing Potential

It would be instructive to re-write the equations of motion (1.79) in terms of the relative distances of the bodies from the primary body, the one denoted with the index $B = 0$. To this end, one introduces the relative-to-the-primary vectors

$$\boldsymbol{R}_A = \boldsymbol{x}_A - \boldsymbol{x}_0 \,, \qquad \boldsymbol{R}_B = \boldsymbol{x}_B - \boldsymbol{x}_0 \tag{1.90}$$

where $\boldsymbol{x}_0$ denotes the position of the primary. The equation (1.79) written for the primary is [11)]

$$\ddot{\boldsymbol{x}}_0 = \frac{Gm_A}{R_A^3}\boldsymbol{R}_A + \sum_{\substack{B=1\\B\neq A}}^{N} \frac{Gm_B}{R_B^3}\boldsymbol{R}_B \,, \tag{1.91}$$

and for the other bodies

$$\ddot{\boldsymbol{x}}_A = -\frac{Gm_0}{R_A^3}\boldsymbol{R}_A - \sum_{\substack{B=1\\B\neq A}}^{N} \frac{Gm_B}{r_{AB}^3}\boldsymbol{r}_{AB} \,, \tag{1.92}$$

where, $\boldsymbol{r}_{AB} = \boldsymbol{x}_A - \boldsymbol{x}_B = \boldsymbol{R}_A - \boldsymbol{R}_B$, is a vector of relative distance directed from body B to A. The difference between equations (1.92) and (1.91) amounts to

$$\ddot{\boldsymbol{R}}_A = -\frac{G(m_0 + m_A)}{R_A^3}\boldsymbol{R}_A - \sum_{\substack{B=1\\B\neq A}}^{N} Gm_B\left(\frac{\boldsymbol{r}_{AB}}{r_{AB}^3} + \frac{\boldsymbol{R}_B}{R_B^3}\right) \,, \tag{1.93}$$

whose right side can be recast to a gradient form

$$\ddot{\boldsymbol{R}}_A = \frac{\partial U}{\partial \boldsymbol{R}_A} \,, \tag{1.94}$$

where

$$U = \frac{G(m_0 + m_A)}{R_A} + \mathfrak{W} \,, \tag{1.95}$$

consists of an algebraic sum of a potential of a point-like mass, $m_0 + m_A$, and the *perturbing potential*

$$\mathfrak{W} = \sum_{\substack{B=1\\B\neq A}}^{N} Gm_B\left(\frac{1}{r_{AB}} - \frac{1}{R_B} - \frac{\boldsymbol{R}_A \cdot \boldsymbol{R}_B}{R_B^3}\right) \,. \tag{1.96}$$

11) Notice that the mass, m_0, of the primary cancels out.

The perturbing potential $\mathfrak{W}$ acting on a mass m_A is generated by all external masses m_B other than m_A or the primary with the mass m_0. It depends on the total gravitational potential of the external bodies with masses m_B taken at the position of body A, from which one subtracts a monopole term, $\sim R_B^{-1}$, and a dipolar component, $\sim (\boldsymbol{R}_A \cdot \boldsymbol{R}_B)R_B^{-3}$, which can be interpreted as a force of inertia $\boldsymbol{F}_{\rm in}$ emerging in the non-inertial frame associated with the primary. It is interesting to notice that in case of two-body problem the perturbing potential vanishes identically, $\mathfrak{W} \equiv 0$.

1.2.6 The Tidal Potential and Force

When the distance R_A happens to be much smaller than any of the distances R_B, the problem of relative motion of the body A around the primary becomes a two-body problem with the perturbation caused by the tidal forces from the external bodies. An example of such a motion is rendered by the Earth-Moon system that moves in the external gravitational field of the Sun and the major planets. In this case, the Earth assumes the role of the primary, the Moon plays the role of the secondary body A, while the external bodies $B \neq A$ are the Sun and the major planets of the solar system. An expansion of the perturbing potential $\mathfrak{W}$ in the Taylor series with respect to a small parameter R_A/R_B is obtained by expanding the function r_{AB}^{-1} about the point $R_A = 0$ in terms of the harmonic polynomials. This gives us

$$\begin{aligned} \frac{1}{r_{AB}} &= \sum_{l=0}^{\infty} \frac{1}{l!} R_A^{<i_1} R_A^{i_2} \ldots R_A^{i_l>} \left[\frac{\partial^l}{\partial R_A^{i_1} \partial R_A^{i_2} \ldots \partial R_A^{i_l}} \left(\frac{1}{r_{AB}} \right) \right]_{\boldsymbol{R}_A=0} \\ &= \frac{1}{R_B} + \frac{\boldsymbol{R}_A \cdot \boldsymbol{R}_B}{R_B^3} + \frac{1}{2} \left[\frac{3(\boldsymbol{R}_A \cdot \boldsymbol{R}_B)^2}{R_B^5} - \frac{R_A^2}{R_B^3} \right] + \ldots , \end{aligned} \tag{1.97}$$

where the angular brackets around spatial indices denote STF (symmetric and trace-less) tensor, and one has used vector notation $R_A^{i_l} = \boldsymbol{R}_A$ $(l = 1, 2, \ldots)$, with each index i_l taking the values $(1, 2, 3)$ corresponding to the three Cartesian coordinates (x^1, x^2, x^3). Substituting this expansion into equation (1.96), one sees that both the monopole $(l = 0)$ and dipole $(l = 1)$ terms canceled out, so the tidal expansion of the disturbing potential acquires the following form

$$\mathfrak{W} = \sum_{l=2}^{\infty} \frac{(-1)^l}{l!} R_A^{<i_1} R_A^{i_2} \ldots R_A^{i_l>} \frac{\partial^l \bar{U}}{\partial R_B^{i_1} \partial R_B^{i_2} \ldots \partial R_B^{i_l}} , \tag{1.98}$$

where

$$\bar{U} = \sum_{\substack{B=1 \\ B\neq A}}^{N} \frac{Gm_B}{R_B} \tag{1.99}$$

is the gravitational potential created by the external bodies at the position of the primary. The lowest-order term of the tidal potential $\mathfrak{W}$ is $l = 2$, which corresponds to the quadrupole moment in the expansion of the external gravitational potential

$\bar{U}$ in the immediate neighborhood of the primary. It is worthwhile to point out that the partial derivatives in the expansion for the tidal potential possess the following property: contraction with respect to any couple of indices gives identically zero, because the external potential $\bar{U}$ satisfies the homogeneous Laplace equation

$$\triangle\bar{U} \equiv \delta^{ij}\frac{\partial^2\bar{U}}{\partial R_B^i \partial R_B^j} = 0\,, \tag{1.100}$$

where the repeated indices assume summation from 1 to 3, and $\delta^{ij} = \mathrm{diag}(1,1,1)$ is the unit matrix (the Kronecker symbol).

The tidal force $\boldsymbol{F}_{\rm tide} = \left(F^i_{\rm tide}\right)$ exerted on the body A orbiting the primary is calculated as a partial derivative of the tidal potential

$$F^i_{\rm tide} = \frac{\partial\mathfrak{W}}{\partial R_A^i} = \sum_{l=2}^{\infty}\frac{(-1)^l}{(l-1)!}R_A^{<i_1}R_A^{i_2}...R_A^{i_{l-1}>}\frac{\partial^l\bar{U}_A}{\partial R_B^{i_1}...\partial R_B^{i_{l-1}}\partial R_B^i}\,, \tag{1.101}$$

so the equation of the relative motion of the body A around the primary is

$$\begin{aligned}\ddot{\boldsymbol{R}}_A &= -\frac{G(m_0+m_A)}{R_A^3}\boldsymbol{R}_A + \boldsymbol{F}_{\rm tide} \\ &= -\frac{G(m_0+m_A)}{R_A^3}\boldsymbol{R}_A + \sum_{\substack{B=1\\B\neq A}}^{N}\frac{Gm_B}{R_B^3}\left[\frac{3(\boldsymbol{R}_A\cdot\boldsymbol{R}_B)\boldsymbol{R}_B}{R_B^2} - \boldsymbol{R}_A\right] + \ldots\,,\end{aligned} \tag{1.102}$$

where one has shown only the leading (quadrupole) term in the tidal force perturbing the motion of the body A. If the orbit of the body A around its primary is circular and there is only one external body B lying on the x axis, the quadrupole tidal force is maximal at the point of intersection of the orbit with x axis, and is minimal at the points of intersection of the orbit with the y axis. The ratio of the maximal-to-minimal values of the tidal force amounts to 2. A plot of the quadrupole tidal force at different points in space around the primary is demonstrated in Figure 1.2.

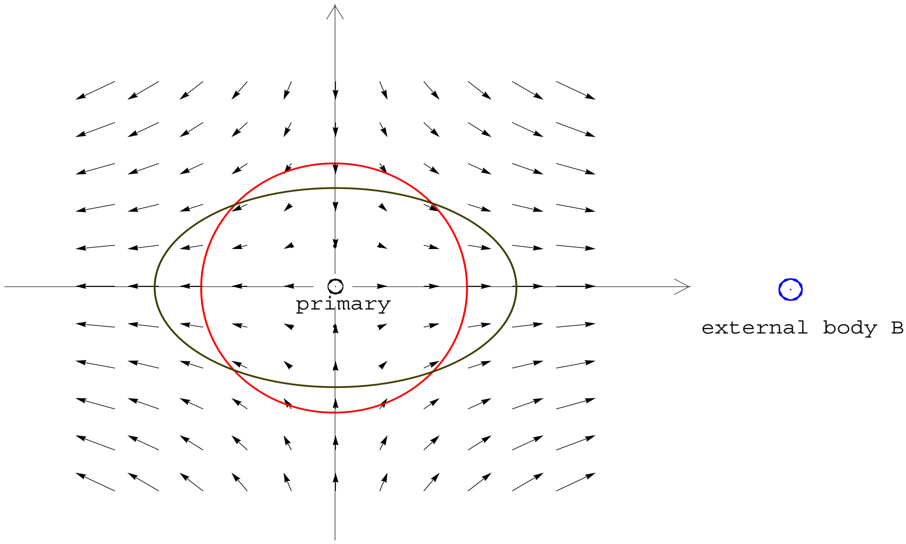

Figure 1.2 The vector field of the quadrupole tidal force is shown at different points in space around the primary kept fixed at the coordinate origin. The tidal force is caused by the external body B laid out on the x axis far away from the primary. The circle depicts a circular orbit in the plane $x - y$ that would be described by the body A about the primary in the absence of the perturbing tidal force. The quadrupole tidal force squeezes the circular orbit in the plane $x - y$ along the direction toward the body B so that the orbit becomes an ellipse with the ratio of its axes equals to 2.

1.3
The Reduced Two-Body Problem

In the simplest case of two bodies, only a primary of mass m_0 located at $\boldsymbol{x}_0$, and a secondary of mass m_1 located at $\boldsymbol{x}_1$ are present. The disturbance $\mathfrak{W}$ vanishes, because the subscript B in equation (1.96) runs through one value solely, $B = 1$, and there are no other values to be taken by the index $B \neq A$. The motion becomes mathematically equivalent to the Newtonian one-body problem, i.e., to movement about a *fixed* center of mass, $M = m_0 + m_1$, given by equation

$$\ddot{\boldsymbol{r}} = -\frac{GM}{r^3}\boldsymbol{r} , \tag{1.103}$$

where $\boldsymbol{r} \equiv \boldsymbol{r}_1$. Equation (1.103) has been derived independently in section 1.1.3. A fortunate aspect of the two-body problem is that it is integrable in terms of elementary functions. The outcome is Newton's celebrated result: the generic solution is a conic with the gravitating center in one of its foci. This then grants one a possibility to thoroughly discuss multiple aspects of the orbital motion of the bodies making use of various parameterizations of the conics.

1.3.1
Integrals of Motion and Kepler's Second Law

Let the center of attraction be located in the origin of an inertial reference frame parameterized with axes $x^i = (x^1, x^2, x^3) = (x, y, z)$, as shown in Figure 1.3. The directions of the axes are defined via three unit vectors $\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z$ with the following components

$$\boldsymbol{e}_x = (1, 0, 0) , \qquad \boldsymbol{e}_y = (0, 1, 0) , \qquad \boldsymbol{e}_z = (0, 0, 1) . \tag{1.104}$$

The position of a moving body is given by the radius-vector $\boldsymbol{r}$. The projection of the velocity $\boldsymbol{v} = \dot{\boldsymbol{r}}$ of the body [12] onto the direction of the radius-vector $\boldsymbol{r}$ gives the rate $\dot{r}$ at which the radial distance evolves. In other words, the Euclidean dot product $\dot{\boldsymbol{r}} \cdot \boldsymbol{r} = \dot{r}r$. Keeping this in mind, one sees that the dot-product of equation (1.103) by $\dot{\boldsymbol{r}}$ is

$$\dot{\boldsymbol{r}} \cdot \ddot{\boldsymbol{r}} + GM\frac{\dot{r}}{r^2} = 0 . \tag{1.105}$$

Integration of the latter results in a conservation law of the orbital energy,

$$\mathcal{E} = \frac{1}{2}\dot{\boldsymbol{r}}^2 - \frac{GM}{r} , \tag{1.106}$$

with the constant $\mathcal{E}$ being the energy per unit mass.

12) One reminds that a dot over any function of time denotes an ordinary derivative with respect to time, for example, $\dot{\boldsymbol{r}} = d\boldsymbol{r}/dt$, etc.

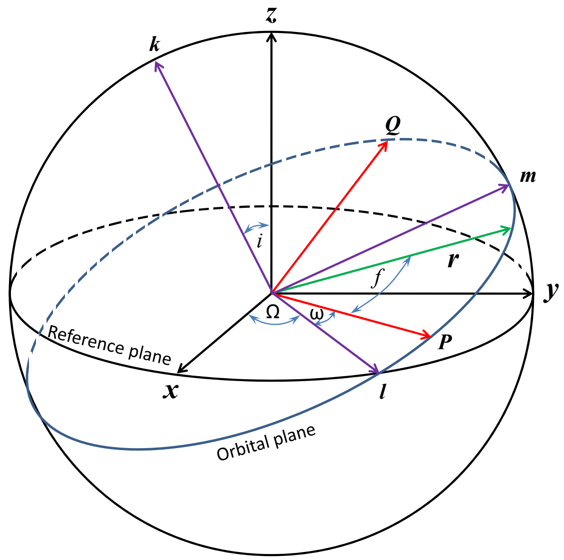

Figure 1.3 Inertial reference frame $(x^1, x^2, x^3) = (x, y, z)$ has its origin at the fixed center of gravity with mass M. Orbital plane is orthogonal to the unit vector $\boldsymbol{k}$ and intersects with the reference plane (x, y) along the apsidal line defined by the unit vector $\boldsymbol{l}$. Position of the moving body is characterized by the radius-vector $\boldsymbol{r}$.

Taking the cross-product of both sides of equation (1.103) with vector $\boldsymbol{r}$, one trivially ends up with

$$\boldsymbol{r} \times \ddot{\boldsymbol{r}} = 0 , \tag{1.107}$$

integration whereof gives us another conservation law of the orbital angular momentum,

$$\boldsymbol{J} = \boldsymbol{r} \times \dot{\boldsymbol{r}} . \tag{1.108}$$

Here the constant vector, $\boldsymbol{J}$, is orthogonal to both $\boldsymbol{r}$ and $\dot{\boldsymbol{r}}$ and is easily identifiable with the orbital angular momentum per unit mass. Conservation of this vector tells us that the plane defined by $\boldsymbol{r}$ and $\dot{\boldsymbol{r}}$ stays unchanged. This circumstance gives birth to the term *orbital plane*. Notice that both magnitude and direction of vector $\boldsymbol{J}$ are conserved, so that each component J_x, J_y, and J_z of this vector is an independent integral of motion.

Orientation of the orbital plane is uniquely defined by the orientation of the angular momentum vector $\boldsymbol{J}$ that is parallel to the unit vector $\boldsymbol{k}$, that is $\boldsymbol{J} = J\boldsymbol{k}$ as shown in Figure 1.3. The orbital plane intersects with the reference plane (x, y) along the apsidal line defined by the unit vector $\boldsymbol{l}$ that is directed towards the ascending node of the orbit which assumes that the body in Figure 1.3 moves counter-clockwise if one watches the motion from the tip of vector $\boldsymbol{k}$. The unit vector $\boldsymbol{m} = \boldsymbol{k} \times \boldsymbol{l}$, and lies in the orbital plane. The triad of unit vectors $\boldsymbol{l}, \boldsymbol{m}, \boldsymbol{k}$ is related to the triad of unit vectors $\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z$ defining orientation of three axes of the inertial frame, as follows

$$\boldsymbol{l} = \boldsymbol{e}_x \cos\Omega + \boldsymbol{e}_y \sin\Omega , \tag{1.109}$$
$$\boldsymbol{m} = -\boldsymbol{e}_x \cos i \sin\Omega + \boldsymbol{e}_y \cos i \cos\Omega + \boldsymbol{e}_z \sin i , \tag{1.110}$$
$$\boldsymbol{k} = \boldsymbol{e}_x \sin i \sin\Omega - \boldsymbol{e}_y \sin i \cos\Omega + \boldsymbol{e}_z \cos i . \tag{1.111}$$

Here, the angle Ω is the longitude of the ascending node of the orbit, and the angle i is the inclination of the orbit with respect to the reference plane.

Let us now introduce within the orbital plane the polar coordinates of the moving body which are the radial distance, r, and the argument of latitude, θ, that is the angle between vectors $\boldsymbol{r}$ and $\boldsymbol{l}$ measured counter-clockwise in the orbital plane from the direction $\boldsymbol{l}$. In terms of r and θ one has,

$$\boldsymbol{r} = r(\boldsymbol{l}\cos\theta + \boldsymbol{m}\sin\theta) , \tag{1.112}$$
$$\dot{\boldsymbol{r}} = \boldsymbol{l}\left(\dot{r}\cos\theta - r\dot{\theta}\sin\theta\right) + \boldsymbol{m}\left(\dot{r}\sin\theta + r\dot{\theta}\cos\theta\right) . \tag{1.113}$$

The angular momentum $\boldsymbol{J}$ being expressed in the polar coordinates becomes,

$$\boldsymbol{J} = \boldsymbol{k} r^2 \dot{\theta} , \tag{1.114}$$

with its absolute value

$$J = |\boldsymbol{J}| = r^2 \dot{\theta} = \text{const.} \tag{1.115}$$

In time δt, the radius-vector $\boldsymbol{r}$ sweeps out the angle $\delta\theta = \dot{\theta}\delta t$ and the area

$$\delta A = \frac{1}{2} r(r + \delta r)\sin(\delta\theta) = \frac{1}{2} r^2 \delta\theta \,. \tag{1.116}$$

After dividing each side of equation (1.116) by δt and taking the limit $\delta t \to 0$ one gets a differential equation for the area's temporal change,

$$\dot{A} = \frac{1}{2} r^2 \dot{\theta} = \frac{J}{2} \,. \tag{1.117}$$

Integration of this equation ensues Kepler's second law of planetary motion: *equal areas are swept out in equal times*, as can be envisaged from the right side of equation (1.117) telling us that the time derivative of the area equals to constant $J/2$. Thus, Kepler's second law has been shown to follow from Newton's theory. Our next step will be to demonstrate that Kepler's first and third laws do so as well.

1.3.2
The Equations of Motion and Kepler's First Law

Let us write the equation of motion (1.103) with respect to inertial Cartesian axes defined by the unit vectors $\boldsymbol{l}, \boldsymbol{m}, \boldsymbol{k}$. In fact, due to the law of conservation of the angular momentum only two components of this equation in the orbital plane will be present. Differentiating the law of conservation of the angular momentum (1.114) one obtains,

$$r\ddot{\theta} + 2\dot{r}\dot{\theta} = 0 \,. \tag{1.118}$$

Differentiating equation (1.113) with respect to time and making use of equation (1.118), allows us to express the acceleration of the body in the following form,

$$\ddot{\boldsymbol{r}} = \boldsymbol{r}\left(\frac{\ddot{r}}{r} - \dot{\theta}^2\right) . \tag{1.119}$$

Equation (1.115) also tells us that the time derivative of the angle θ is

$$\dot{\theta} = \frac{J}{r^2} \,. \tag{1.120}$$

Substituting this expression in equation (1.119), and the result of the substitution into vectorial equation of motion (1.103), brings about a differential equation for the radial motion of the body

$$\ddot{r} + \frac{GM}{r^2} - \frac{J^2}{r^3} = 0 \,, \tag{1.121}$$

where J is the constant angular momentum of the orbital motion. Solution of this differential equation was not known at the time of Newton. For this reason, in his *Principia* Newton proved only that orbits in the form of conics necessitate an inverse-square gravity law. To prove the inverse statement, i.e., that the gravity law entails this type of orbits, Newton would have to cope with equation (1.121) which solution had to wait until 1710. On the 13-th of December 1710, two Swiss mathematicians, Johann Bernoulli and Jakob Hermann,[13)] presented their solutions of equation 1.121 to a meeting of the Paris Academy of Sciences. Both speakers provided valid proofs (to which they had come independently) that Newton's gravity law yields conical orbits. For historical account of those events see [Weinstock, 1982].

In the middle of the XVIII-th century, solution of this problem was greatly simplified due to an elegant mathematical trick pioneered independently by d'Alembert and Clairaut. The first step of this method is to employ equation (1.120) as a mean of switching from differentiation with respect to time t to differentiation with respect to the angle θ. The rationale beneath this replacement of variable is to find the shape of an orbital curve, i.e., the dependence of r upon θ. Thus, one replaces the time derivatives of r with those with respect to θ

$$\dot{r} = \dot{\theta} r' , \tag{1.122}$$

$$\ddot{r} = \ddot{\theta} r' + \dot{\theta}^2 r'' , \tag{1.123}$$

where a dot signifies the time derivative, and the prime denotes the derivative with respect to θ. Making use of equations (1.118), (1.120), and (1.122) one can recast equation (1.123) to the following form:

$$\ddot{r} = \frac{J^2}{r^2}\left(\frac{r''}{r^2} - 2\frac{r'^2}{r^3}\right) . \tag{1.124}$$

The second crucial step is to replace r with a reciprocal radial variable $u = 1/r$, thus obtaining:

$$u' = -\frac{r'}{r^2} , \tag{1.125}$$

$$u'' = -\frac{r''}{r^2} + 2\frac{r'^2}{r^3} . \tag{1.126}$$

Comparing this equation with (1.124), and using the substitution $u = 1/r$, one obtains

$$\ddot{r} = -J^2 u^2 u'' . \tag{1.127}$$

This expression, along with $\dot{\theta} = J r^{-2} = J u^2$, helps us to transform (1.121) to the Binet equation

$$u'' + u = \frac{GM}{J^2} , \tag{1.128}$$

which is the equation of harmonic oscillator subject to a constant perturbation GM/J^2. Solution of this equation is a linear superposition of general solution of

13) The two were former disciples of Johann's older brother, Jakob Bernoulli; and Hermann was a distant relative of Euler.

a homogeneous equation $u'' + u = 0$, and a particular solution of the inhomogeneous equation (1.128)

$$u = B\cos(\theta - \omega) + \frac{GM}{J^2}\,, \tag{1.129}$$

where B and ω are constants of integration depending on the initial conditions. This solution looks similar to a well-known in analytic geometry expression for the reciprocal distance from a point on a conic to one of its foci,

$$\frac{p}{r} = 1 + e\cos f\,, \tag{1.130}$$

where f is the *true anomaly*, i.e., the angular separation of the point from the periapse (subtended at the said focus), while $p = a(1 - e^2)$ is a constant parameter being expressed in terms of the semimajor axis a and eccentricity e, and called *semilatus rectum* [14].

To convert the resemblance to equivalence, one must choose the constants B and J in equation (1.129) in the following form

$$B = \frac{e}{a(1-e^2)} = \frac{e}{p}\,, \tag{1.131}$$

$$J = \sqrt{GMp}\,, \tag{1.132}$$

and equate the angular variables

$$f = \theta - \omega\,. \tag{1.133}$$

After these identifications, the orbital elements e and p turn out to be interconnected with the integrals of motion via formulae

$$p = \frac{J^2}{GM}\,, \tag{1.134}$$

$$e = \sqrt{1 + \frac{2\mathcal{E}J^2}{G^2M^2}}\,, \tag{1.135}$$

so the constant B from equation (1.129) becomes

$$B = \frac{GM}{J^2}\sqrt{1 + \frac{2\mathcal{E}J^2}{G^2M^2}}\,, \tag{1.136}$$

and the integral of the reduced total energy

$$\mathcal{E} = \frac{GM}{2p}(e^2 - 1)\,. \tag{1.137}$$

14) To derive equation (1.130) for an ellipse, a circle, or a hyperbola, start with equation (1.147) written in a Cartesian coordinate system (ξ, η), whose origin is located in one of the foci, the axis ξ going through the foci, as shown in Figure 1.4. For a point on the conic, $\xi = r\cos f$ and $\eta = r\sin f$, plugging of which into equation (1.147) entails equation (1.130).

For different conics, the parameters of the orbit are defined as

$$\text{circle} \quad : \quad p = a\,, \qquad e = 0\,, \tag{1.138a}$$

$$\text{ellipse} \quad : \quad p = a(1-e^2)\,, \qquad 0 < e < 1\,, \tag{1.138b}$$

$$\text{parabola} \quad : \quad p = 2q\,, \qquad e = 1\,, \tag{1.138c}$$

$$\text{hyperbola} \quad : \quad p = a(e^2-1)\,, \qquad e > 1\,. \tag{1.138d}$$

Equation (1.132) demonstrates that the angular momentum, J, of the orbit depends only on the focal parameter, p, and is always positive for any type of the orbit

$$J^2 = GMp > 0\,. \tag{1.139}$$

On the other hand, equation (1.137) reveals that the reduced total energy $\mathcal{E}$ of the two-body system depends only on the semi-major axis, a, and has either positive or negative, or zero value for different types of conics

$$\text{ellipse} \quad : \quad \mathcal{E} = -\frac{GM}{2a} < 0\,, \tag{1.140a}$$

$$\text{parabola} \quad : \quad \mathcal{E} = 0\,, \tag{1.140b}$$

$$\text{hyperbola} \quad : \quad \mathcal{E} = \frac{GM}{2a} > 0\,. \tag{1.140c}$$

The case of parabola is exceptional in that its eccentricity $e = 1$, and the semi-latus rectum is defined as $p = 2q$, where q is the minimal distance of the orbit to the gravitating center at body's closest approach. As parabolic (or near-parabolic) orbits are considered in extremely rare situations, they will be omitted below [15].

The point of the closest approach of the orbit to the attracting center is called *pericenter*, and the opposite point on the orbit is called *apocenter* [16]. Let us define a unit vector $\boldsymbol{P}$ directed from the attracting center towards pericenter and another unit vector $\boldsymbol{Q}$ lying in the orbital plane so that the triad $\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{k}$ make a right-handed system of three vectors (see Figure 1.3). Vectors $\boldsymbol{P}$ and $\boldsymbol{Q}$ are related to vectors $\boldsymbol{l}$ and $\boldsymbol{m}$ by rotation at the angle ω

$$\boldsymbol{P} = \boldsymbol{l}\cos\omega + \boldsymbol{m}\sin\omega\,, \tag{1.141}$$

$$\boldsymbol{Q} = -\boldsymbol{l}\sin\omega + \boldsymbol{m}\cos\omega\,. \tag{1.142}$$

15) One possibility of integrating such orbits is through switching to the Kustaanheimo-Stiefel variables [Stiefel, 1976]. Treatment by more conventional means is offered, for example, by Osman and Ammar [2006].

16) In case of a planetary orbit around the Sun these orbital points are called respectively *perihelion* and *aphelion*. Corresponding points on the orbit of the Moon and on the orbit of artificial satellites of the Earth are called *perigee* and *apogee*.

The angle ω is the same as in equation (1.133). It measures in the orbital plane the angular distance of the pericenter from the ascending node. In terms of the constant unit vectors $\boldsymbol{P}$ and $\boldsymbol{Q}$ the radius-vector of the body is expressed as follows

$$\boldsymbol{r} = r(\boldsymbol{P}\cos f + \boldsymbol{Q}\sin f) \ . \tag{1.143}$$

Derivation of velocity, $\boldsymbol{v} = \dot{\boldsymbol{r}}$, is achieved by direct differentiation of equation (1.143),

$$\dot{\boldsymbol{r}} = \dot{r}(\boldsymbol{P}\cos f + \boldsymbol{Q}\sin f) + r\dot{f}(-\boldsymbol{P}\sin f + \boldsymbol{Q}\cos f) \ . \tag{1.144}$$

Making use of equation (1.130) describing the first Kepler's law along with equation $\dot{f} = \dot{\theta}$, the integral of the angular momentum taken in the form of equations (1.120) and (1.132) allows us to get the time derivative of the radial distance,

$$\dot{r} = \frac{e}{p}\sin f(r^2\dot{f}) = \frac{Je}{p}\sin f = \sqrt{\frac{GM}{p}}e\sin f \ . \tag{1.145}$$

Substituting this expression in equation (1.144) and making use of the law of conservation of the angular momentum, result in

$$\dot{\boldsymbol{r}} = \sqrt{\frac{GM}{p}}\left[-\boldsymbol{P}\sin f + \boldsymbol{Q}(\cos f + e)\right] \ . \tag{1.146}$$

1.3.3
The mean and eccentric anomalies. Kepler's third law.

To start with, introduce a coordinate system (ξ, η) with an origin fixed at the attraction center F, and with the unit vectors $\boldsymbol{P}$ and $\boldsymbol{Q}$ directed along the axes ξ and η, respectively. Draw an auxiliary circle of radius a equal to the semi-major axis, and centered at the midpoint O between the foci. Let the body be located on the conic at point B at time t, as shown in Figure 1.4. One draws a straight line parallel to the η axis and passing through the orbiter. The line is orthogonal to the ξ axis, and intersects the auxiliary circle at point C. The *eccentric anomaly* E is defined as the angle subtended at the circle's center, between the axis ξ and a straight line pointing at the point C on the circle.

As the distance between the origin and a focus is ea, one writes down the equation for the conic

$$\frac{(\xi + ae)^2}{a^2} + \frac{\eta^2}{a^2(1-e^2)} = 1 \ , \tag{1.147}$$

valid for a circle, an ellipse, or a hyperbola.

The abscissa of point the B on the conic is equal to $\xi = r\cos f$. At the same time, it is equal to

$$\xi = a\cos E - ae = a(\cos E - e) \ , \tag{1.148}$$

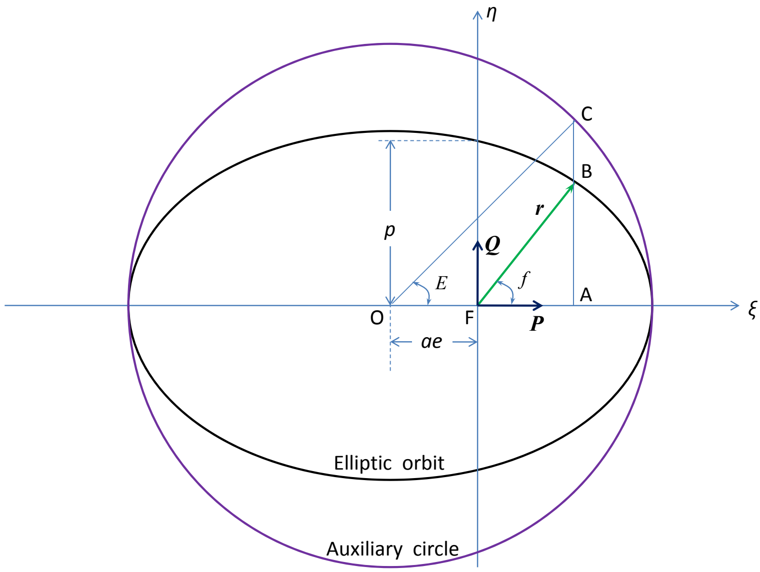

Figure 1.4 The orbiting body is at the point B. The eccentric anomaly is the angle E between directions OA and OC. The true anomaly f is the angle between the ξ axis and the radius-vector $\boldsymbol{r}$. The distance between the center of the auxiliary circle and the conic is ae, where a is the semi-major axis and e is the eccentricity.

because the distance OF between the attraction center at point F and the center of the auxiliary circle at point O is equal to ae, in accordance with definitions of eccentricity e and the semi-major axis a. Equating the two expressions for the abscissa of the point B, and making use of equation (1.130) of the conic, one gets a formula interconnecting the two anomalies:

$$\cos E = \frac{e + \cos f}{1 + e\cos f} \ . \tag{1.149}$$

The ordinate of point B on the conic is $\eta = r\sin f$. Accounting for equation (1.147) of the conic and equation (1.148), one obtains

$$\eta = a\sqrt{1 - e^2}\sin E \ . \tag{1.150}$$

Again, equating the two expressions for η gives us another interconnection,

$$\sin E = \frac{\sqrt{1 - e^2}\sin f}{1 + e\cos f} \ , \tag{1.151}$$

which leads us to the expression for the distance from the focus, as a function of the eccentric anomaly:

$$r = a(1 - e\cos E) \ . \tag{1.152}$$

In combination with (1.149), the latter renders:

$$r = \frac{a(1 - e^2)}{1 + e\cos f} \ . \tag{1.153}$$

The equation (1.149) can be rewritten in the following, equivalent forms:

$$1 - \cos f = (1 + e)\frac{1 - \cos E}{1 - e\cos E} \ , \tag{1.154a}$$

$$1 + \cos f = (1 - e)\frac{1 + \cos E}{1 - e\cos E} \ . \tag{1.154b}$$

With aid of the standard double-angle formulae, these relations can be reshaped correspondingly into

$$\sin^2\frac{f}{2} = \frac{1 + e}{1 - e\cos E}\sin^2\frac{E}{2} \ , \tag{1.155a}$$

$$\cos^2\frac{f}{2} = \frac{1 - e}{1 - e\cos E}\cos^2\frac{E}{2} \ . \tag{1.155b}$$

The ratio of these two formulae furnishes yet another elegant interconnection between the true and eccentric anomalies,

$$\tan\frac{f}{2} = \sqrt{\frac{1 + e}{1 - e}}\tan\frac{E}{2} \ . \tag{1.156}$$

Our next project is to link the anomalies with the time. To this end, one employs equations (1.120) and (1.133) to write

$$df = d\theta = \frac{J}{r^2}dt \ . \tag{1.157}$$

As the formulae (1.132) and (1.153) enable us to express J and r via the elements, one can rewrite (1.157) as

$$df = \sqrt{\frac{GM}{a^3}}\frac{(1+e\cos f)^2}{(1-e^2)^{3/2}}dt \ . \tag{1.158}$$

Introducing a widely used quantity

$$n = \sqrt{\frac{GM}{a^3}} \ , \tag{1.159}$$

called *mean motion* or *mean angular frequency*, one can write down (1.158) as

$$ndt = \frac{\left(1-e^2\right)^{3/2}}{(1+e\cos f)^2}df \ . \tag{1.160}$$

In the special case of a bound orbit, i.e., when the conic is a circle or an ellipse, the above formula shows an important property. As integration of its right side over a period, i.e., from $f = 0$ through $f = 2\pi$, gives exactly 2π, one has,

$$nT = 2\pi \ , \tag{1.161}$$

where T being the time of the orbital period. Obviously, the mean motion n is the angular velocity, in the case of a circular orbit, and an *average* angular velocity (as seen from the focus), in the case of ellipse. Combining expression (1.161) with the definition (1.159), one arrives at Kepler's third law in two-body problem,

$$T^2 = \frac{4\pi^2 a^3}{GM} \ , \tag{1.162}$$

which reads: *the square of the orbital period of a body is directly proportional to the cube of the semi-major axis of its orbit.*

Our next step is to interconnect the time with the eccentric anomaly. From the afore-proven equality (1.156) one deduces

$$dE = \frac{\sqrt{1-e^2}}{1+e\cos f}df \ . \tag{1.163}$$

Moreover, equations (1.152) and (1.153) yields

$$1 - e\cos E = \frac{1-e^2}{1+e\cos f} \ . \tag{1.164}$$

Formulae (1.163), (1.164) and (1.160) being put together, entail

$$ndt = (1 - e\cos E)dE = d(E - e\sin E)\ , \tag{1.165}$$

wherefrom a simple integration yields

$$E - e\sin E = n(t - t_0) + \mathcal{M}_0\ , \tag{1.166}$$

where t_0 is the fiducial time, called the *epoch*, and $\mathcal{M}_0$ is the integration constant. The latter compels us to define, following Kepler, a convenient quantity

$$\mathcal{M} = \mathcal{M}_0 + n(t - t_0)\ , \tag{1.167}$$

called *mean anomaly*, with $\mathcal{M}_0$ now termed as the *mean anomaly at epoch*. According to (1.166), $\mathcal{M}$ obeys the *Kepler equation*

$$\mathcal{M} = E - e\sin E\ , \tag{1.168}$$

whence it is clear that, for elliptic and circular orbits, $\mathcal{M}$ changes by 2π over a period, because the eccentric anomaly changes by 2π, and $\sin E$ is a periodic function with period 2π.

Finally, let us notice that (1.163) and (1.168) enable us to interconnect the mean and eccentric anomalies:

$$d\mathcal{M} = (1 - e\cos E)dE = \frac{\left(1 - e^2\right)^{3/2}}{(1 + e\cos f)^2}df\ . \tag{1.169}$$

For bound orbits, this yields for one period of orbital revolution

$$\oint\limits_{\substack{\text{orbital}\\ \text{period}}} d\mathcal{M} = \int_0^{2\pi} \frac{\left(1 - e^2\right)^{3/2}}{(1 + e\cos f)^2}df = 2\pi\ , \tag{1.170}$$

whence we, once again see that $\mathcal{M}$ changes by 2π over a period.

Be mindful that one did *not* necessarily imply t_0 to be the instant of the first periapse passage τ. The time τ_N of the N-th periapse passage is defined from the condition that the eccentric anomaly

$$E = 2\pi(N - 1)\ , \tag{1.171}$$

N being an integer chosen so that $N = 1$ corresponds to $E = 0$ that is the first periapse passage. As evident from equation (1.168), condition 1.171) can be rewritten also as

$$\mathcal{M} = 2\pi(N - 1)\ . \tag{1.172}$$

From here and equation (1.167), it is easy to demonstrate that the N-th periapse passage takes place at the time

$$\tau_N = t_0 + \frac{2\pi(N - 1) - \mathcal{M}_0}{n}\ . \tag{1.173}$$

In celestial mechanics, the time of the first periapse passage, τ_1, is denoted simply as τ, which is

$$\tau = t_0 - \frac{\mathcal{M}_0}{n} \; . \tag{1.174}$$

Hence, the mean anomaly expressed in terms of the first periapse passage will look

$$\mathcal{M} = n(t - \tau) \; . \tag{1.175}$$

Over one orbital revolution, the mean anomaly changes by 2π, while the time changes by period T. So equation (1.175) naturally renders (1.161).

1.3.4
The Laplace-Runge-Lenz Vector

One has already learned that the reduced two-body problem obeying the Newton gravity law (1.103) permits four integrals of motion - the energy, $\mathcal{E}$, and the three components of the angular-momentum vector, $\boldsymbol{J}$. It is remarkable that the reduced two-body problem admits one more integral of motion. To demonstrate this fact, consider the so-called *Laplace-Runge-Lenz vector*

$$\boldsymbol{A}_{\rm L} = \dot{\boldsymbol{r}} \times \boldsymbol{J} - GM\frac{\boldsymbol{r}}{r} \; , \tag{1.176}$$

where $\boldsymbol{J}$ is the conserved angular-momentum vector. Despite its name, the Laplace-Runge-Lenz vector was discovered by neither of these three scholars. The honor of its discovery belongs to the afore mentioned Jakob Hermann. It was also he who demonstrated that $\boldsymbol{A}_{\rm L}$ is conserved in the two-body problem governed by Newton's gravity law of inverse squares [Hermann, 1710].

One way to explore this vector is to make use of equations (1.143), (1.146) defining the position $\boldsymbol{r}$ and velocity $\dot{\boldsymbol{r}}$ of the orbiting body in terms of the orthogonal unit vectors $\boldsymbol{P}$ and $\boldsymbol{Q}$ shown in Figures 1.3 and 1.4. Together, these equations yield the following expression for the angular-momentum vector

$$\boldsymbol{J} = \boldsymbol{k}\sqrt{GMp} \; , \tag{1.177}$$

where one has used equation (1.132) and employed the relationship $\boldsymbol{k} = \boldsymbol{P} \times \boldsymbol{Q}$. Insertion of equations (1.146) and (1.177) in the right side of equation (1.176) entails,

$$\boldsymbol{A}_{\rm L} = GMe\boldsymbol{P} \; , \tag{1.178}$$

which tells us that the Laplace-Runge-Lenz vector is a constant vector directed towards the pericenter, the closest point of the orbit.

It would also be instructive to express the magnitude of the Laplace-Runge-Lenz vector through the other conserved quantities. Squaring both sides of equation (1.176) gives

$$A_{\rm L}^2 = GM^2 - \frac{2GM}{r}\boldsymbol{r} \cdot (\dot{\boldsymbol{r}} \times \boldsymbol{J}) + J^2\dot{r}^2 \; . \tag{1.179}$$

At the same time, permuting of the scalar triple product as

$$\boldsymbol{r}\cdot(\dot{\boldsymbol{r}}\times\boldsymbol{J}) = \boldsymbol{J}\cdot(\boldsymbol{r}\times\dot{\boldsymbol{r}}) = J^2 \,, \tag{1.180}$$

enables us to rewrite the latter equation as

$$A_{\rm L}^2 = GM^2 + 2J^2\left(\frac{1}{2}\dot{r}^2 - \frac{GM}{r}\right) = GM^2 + 2J^2\mathcal{E} \,, \tag{1.181}$$

with $\mathcal{E}$ being the constant reduced energy per unit mass.

Conservation of both the absolute value and the direction of $\boldsymbol{A}_{\rm L}$ tells us that its three components $(A_{{\rm L}x}, A_{{\rm L}y}, A_{{\rm L}z})$ are integrals of motion. However, only one of these three components can be regarded as an *independent* integral of motion. Indeed, as can be seen from equation (1.181), the magnitude of the Laplace-Runge-Lenz vector, $\boldsymbol{A}_{\rm L}^2 = A_{{\rm L}x}^2 + A_{{\rm L}y}^2 + A_{{\rm L}z}^2$, can be expressed through two known integrals - the energy, $\mathcal{E}$, and the magnitude of the angular-momentum vector, J. Another constraint follows from the fact that the vector $\boldsymbol{A}_{\rm L}$ belongs to the orbital plane and, thus, is always orthogonal to the vector of the angular momentum:

$$\boldsymbol{A}_{\rm L}\cdot\boldsymbol{J} = A_{{\rm L}x}J_x + A_{{\rm L}y}J_y + A_{{\rm L}z}J_z = 0 \,. \tag{1.182}$$

This explains why the conservation of the Laplace-Runge-Lenz vector increases the number of the *independent* integrals of motion in the reduced two-body problem not by three but only by one - from four to five.[17] Together with the initial condition (expressed, for example, by fixing $\mathcal{M} = \mathcal{M}_0$ at the epoch t_0), this gives us totally six constants that should single out a particular trajectory of the body out of the entire multitude of parameterized conics. This observation agrees with the afore mentioned fact that an arbitrary solution to the reduced two-body Kepler problem should depend upon six constants like the initial position and velocity of the body.

It can also be demonstrated that the existence of five independent integrals of motion makes it possible to integrate the equations of motion of the reduced two-body problem in quadratures. This fact turns out to be intimately related to the highly symmetrical nature of the reduced two-body problem. A central-force problem is trivially invariant under the spatial rotations making up the group SO(3), hence, the conservation of the angular momentum. However, the inverse-square force proportional to $1/r^2$, and the space harmonic oscillator with the force of elasticity proportional to r^2, possess a symmetry under a bigger group [Landau and Lifshitz, 1975]. In the case of the inverse-square gravity law, this is: SO(4) for a negative energy $\mathcal{E} < 0$; or SO(1,3), for a positive energy $\mathcal{E} > 0$ [Dubrovin *et al.*, 1984, §34]. Both SO(4) and SO(1,3) are rotational groups of symmetry in four-dimensional Euclidean and pseudo-Euclidean space respectively, and the Lie algebras of their generators have the dimension of 6. This circumstance gives birth to 6 conserved quantities - the three components of $\boldsymbol{J}$ and the three components of $\boldsymbol{A}_{\rm L}$. Mind, though, that these components and the energy are interconnected by the two constraints, (1.181) and

17) Generally, a system that has k degrees of freedom and has, at the same time, more than k integrals of motion is called *superintegrable*, while a system with $2k - 1$ integrals is called *maximally superintegrable*, the reduced two-body problem being the case with $k = 3$.

(1.182). Further details on this interesting topic can be found in excellent textbooks by Vozmischeva [2003] and Mathúna [2008].

1.3.5 Parameterizations of the Reduced Two-body Problem

1.3.5.1 *A Keplerian Orbit in the Euclidean Space*

As has been shown in section 1.2, the two-body problem is equivalent in the barycentric frame of reference to its reduced version (1.103) which, mathematically, looks as a motion of a particle of reduced mass μ about a fixed gravitating center of mass M. The generic solution to equation (1.103) is a Keplerian conic characterized by six constant parameters:

a – the semi-major axis,
e – the eccentricity,
$\mathcal{M}_0$ – the *mean anomaly at epoch*,
Ω – the angle of the ascending node,
i – the orbital inclination,
ω – the longitude of pericenter.

Three of these parameters, Ω, i, and ω, define the orientation of the orbit in space. Two parameters, a and e, fix the shape of the orbit. The remaining one, $\mathcal{M}_0$, determines the position of the body on the orbit at the initial epoch t_0. In arbitrary inertial reference frame, there exist six additional constant parameters - the linear momentum (1.80) and the position of the center of mass (1.81). As a rule, a barycentric reference frame is chosen, so these integrals of motion get nullified and do not appear explicitly in any equation. Still, it is useful to keep in mind that these integrals actually exist.

Substituting equations (1.109)-(1.111), (1.141), (1.142) and (1.153) to expressions (1.143), (1.146) and applying simple trigonometric identities, the explicit form of the position and velocity of the body can be written down in the barycentric inertial coordinates as

$$\boldsymbol{r} = x\boldsymbol{e}_x + y\boldsymbol{e}_y + z\boldsymbol{e}_z \,, \tag{1.183}$$

$$\dot{\boldsymbol{r}} = \dot{x}\boldsymbol{e}_x + \dot{y}\boldsymbol{e}_y + \dot{z}\boldsymbol{e}_z \,, \tag{1.184}$$

where

$$x = \frac{a(1-e^2)}{1+e\cos f}\left[\cos\Omega\cos(\omega+f) - \sin\Omega\sin(\omega+f)\cos i\right] \,, \tag{1.185a}$$

$$y = \frac{a(1-e^2)}{1+e\cos f}\left[\sin\Omega\cos(\omega+f) + \cos\Omega\sin(\omega+f)\cos i\right] \,, \tag{1.185b}$$

$$z = \frac{a(1-e^2)}{1+e\cos f}\sin(\omega+f)\sin i \,, \tag{1.185c}$$

and

$$\dot{x} = -\frac{na}{\sqrt{1-e^2}}\Big[\cos\Omega\sin(\omega+f)+\sin\Omega\cos(\omega+f)\cos i \tag{1.186a}$$
$$+e(\cos\Omega\sin\omega+\sin\Omega\cos\omega\cos i)\Big]\ ,$$
$$\dot{y} = -\frac{na}{\sqrt{1-e^2}}\Big[\sin\Omega\sin(\omega+f)-\cos\Omega\cos(\omega+f)\cos i \tag{1.186b}$$
$$+e(\sin\Omega\sin\omega-\cos\Omega\cos\omega\cos i)\Big]\ ,$$
$$\dot{z} = \frac{na}{\sqrt{1-e^2}}\left[\cos(\omega+f)+e\cos\omega\right]\sin i\ , \tag{1.186c}$$

with n being the mean motion (1.159).

Expressions (1.185) - (1.186) give us one possible form of the generic solution of equation of motion (1.103) - a form corresponding to parametrization of a conic by a set of six *Keplerian constants* $(a, e, \Omega, i, \omega, \mathcal{M}_0)$ and the variable true anomaly f. Since the true anomaly is a function of time, through the relation (1.160), then equations (1.185 - 1.186) do define a dependence of $\boldsymbol{r}$ and $\dot{\boldsymbol{r}}$ upon the constants and the *time*, as required. This dependence, however, is *implicit*, and requires solution of the transcendental Kepler equation (1.168) along with a non-linear trigonometric equation (1.156).

The same solution can be parameterized via some other constants, for example those of Delaunay: $\mathcal{M}_0, \omega, \Omega, \sqrt{GMa}, \sqrt{GMb}, \sqrt{GMb}\cos i$, where $b = a\sqrt{1-e^2}$ is the semi-minor axis of the conic. Another possibility is to consider the solution as a function of the initial conditions: then the constants $(x_0, y_0, z_0, \dot{x}_0, \dot{y}_0, \dot{z}_0)$ are the six parameters defining a particular orbit. The latter option is natural when the integration is carried out numerically in the Cartesian coordinates, but is impractical for analytic treatments. Numerous other parameterizations have been introduced for various purposes. Whatever the set of the constants of integration is chosen, their number should, on general grounds, be six. A switch from the Keplerian constants to the Delaunay ones, or to any other parametrization will still give the same geometric image of the curve in the coordinate (x, y, z) space. The velocity vector of the body being a tangent vector to that curve will not depend on curve's parametrization either. However, the specific mathematical presentation of $\boldsymbol{r}$ and $\dot{\boldsymbol{r}}$ as functions of the new parameters will, of course, be different from those given by equations (1.185) - (1.186).

To avoid confusion one would like to point out that the set of Delaunay constants differs from the set of Delaunay elements, as the latter set includes $\mathcal{M}$ instead of $\mathcal{M}_0$. In the same way, the set of Keplerian constants differs from the set of Keplerian elements: the former set contains $\mathcal{M}_0$, the latter $\mathcal{M}$. Employing the mean anomaly $\mathcal{M} = \mathcal{M}_0 + n(t - t_0)$ is convenient, because this enables one to keep track, via one variable, both of the explicit time dependence and of the dependence upon the integration constant $\mathcal{M}_0$ - see the comprehensive treatise by Plummer [1918].

1.3.5.2 *A Keplerian Orbit in the Projective Space*

A generic solution to the reduced two-body problem is a section of cone i.e., a plane conic curve described by a quadratic polynomial. It would be instrumental to study a conic from the viewpoint of projective geometry, as this approach will allow us to treat any Keplerian orbit - circular, elliptical, parabolic, hyperbolic, and two-body collisional linear orbit - in a unified way. The approach adopted in this section has been proposed by Satō [1998].

A Keplerian orbit in the plane (ξ, η) is given by equations (1.143), (1.130)

$$\xi = \frac{q(1+e)\cos f}{1+e\cos f}\,, \qquad \eta = \frac{q(1+e)\sin f}{1+e\cos f}\,, \tag{1.187}$$

where $q \equiv r_{\min} = a(1-e)$ is the distance to pericenter. Eliminating the true anomaly f, one obtains an ordinary quadratic form for the orbit

$$[(1-e)\xi + eq]^2 + \frac{1-e}{1+e}\eta^2 = q^2\,. \tag{1.188}$$

We now assume that coordinates ξ and η are dimensionless[18)] and replace $\xi = \xi_1/\xi_0$, $\eta = \xi_2/\xi_0$ using homogeneous coordinates (ξ_0, ξ_1, ξ_2) that map equation (1.188) to the projective plane $\mathbb{P}^2$ [Casse, 2006]

$$-(1+e)q\xi_0^2 + 2e\xi_0\xi_1 + \frac{1-e^2}{(1+e)q}\xi_1^2 + \frac{1}{(1+e)q}\xi_2^2 = 0\,. \tag{1.189}$$

Equation (1.189) can be written in a matrix form as

$$[\xi_0, \xi_1, \xi_2]\begin{bmatrix} -(1+e)q & e & 0 \\ e & (1+e)k & 0 \\ 0 & 0 & \dfrac{1}{(1+e)q} \end{bmatrix}\begin{bmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \end{bmatrix} = 0\,, \tag{1.190}$$

where $k \equiv 1/r_{\max} = (1-e)/[(1+e)q]$ is the reciprocal of the apocenter distance $r_{\max}$. We emphasize that there are only two free parameters (q, e) with k being expressed in terms of these two. Furthermore, parameters q and k are dimensionless in accordance with definition of the homogeneous coordinates to which they are related through equation (1.190).

Equation (1.189) can be reduced to a canonical quadratic form by rotation in the projective space which diagonalizes the matrix in equation (1.190). The characteristic equation for eigenvalues λ of the matrix, is

$$[(1+e)q\lambda - 1][\lambda^2 - (1+e)(k-q)\lambda - 1] = 0\,, \tag{1.191}$$

and it has three solutions

$$\lambda_0 = \frac{1}{2}\left[-(1+e)(q-k) - \sqrt{(1+e)^2(q+k)^2 + 4e^2}\right]\,, \tag{1.192}$$

$$\lambda_1 = \frac{1}{2}\left[-(1+e)(q-k) + \sqrt{(1+e)^2(q+k)^2 + 4e^2}\right]\,, \tag{1.193}$$

$$\lambda_2 = \frac{1}{(1+e)q}\,, \tag{1.194}$$

18) Dimensionless of coordinates is an integral part of the definition of the projective plane $\mathbb{P}^2$ [Casse, 2006, section 4].

where one has used identity $(1+e)^2kq = 1-e^2$ for transforming the root square terms in λ_0 and λ_1.

Next step is to replace parameters (q,e) with another set (ϕ,ψ) by making use of identifications

$$\cosh\phi\cos 2\psi = \frac{1}{2}(1+e)(q+k)\,, \tag{1.195a}$$

$$\cosh\phi\sin 2\psi = e\,, \tag{1.195b}$$

$$\sinh\phi = \frac{1}{2}(1+e)(q-k)\,, \tag{1.195c}$$

which makes the eigenvalues $\lambda_0 = -\exp\phi$, $\lambda_1 = \exp(-\phi)$. It is now rather straightforward to find the eigenvectors of the matrix in equation (1.190) that are

$$\boldsymbol{E}_0 = [\cos\psi, -\sin\psi, 0]\;, \quad \boldsymbol{E}_1 = [\sin\psi, \cos\psi, 0]\;, \quad \boldsymbol{E}_2 = [0,0,1]\;. \tag{1.196}$$

Canonical homogeneous coordinates of conic in the projective space $\mathbb{P}^2$ corresponding to these eigenvectors are

$$\{\zeta_0 : \zeta_1 : \zeta_2\} = \left\{\frac{1}{\sqrt{-\lambda_0}} : \frac{\cos\Theta}{\sqrt{\lambda_1}} : \frac{\sin\Theta}{\sqrt{\lambda_2}}\right\}\,, \tag{1.197}$$

where Θ is called the *projective anomaly* [Satō, 1998]. Matrix of rotation from the canonical homogeneous coordinates $(\zeta_0,\zeta_1,\zeta_2)$ to the original ones is made of the components of the eigenvectors. The transformation has the following form of the rotation about ξ_2 axis

$$\begin{bmatrix}\xi_0\\ \xi_1\\ \xi_2\end{bmatrix} = \begin{bmatrix}\cos\psi & \sin\psi & 0\\ -\sin\psi & \cos\psi & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}\zeta_0\\ \zeta_1\\ \zeta_2\end{bmatrix} \tag{1.198}$$

$$= \begin{bmatrix}1+\exp\phi\tan\psi\cos\Theta\\ -\tan\psi+\exp\phi\cos\Theta\\ \sqrt{\exp 2\phi-\tan^2\psi}\sin\Theta\end{bmatrix}\exp\left(-\frac{\phi}{2}\right)\cos\psi\,.$$

Therefore, substituting

$$\alpha \equiv \exp\phi = \frac{1}{2}\left[(1+e)(q-k)+\sqrt{(1+e)^2(q+k)^2+4e^2}\right]\,, \tag{1.199}$$

$$\beta \equiv \tan\psi = \frac{2e}{(1+e)(q+k)+\sqrt{(1+e)^2(q+k)^2+4e^2}}\,, \tag{1.200}$$

one obtains parametrization of the Keplerian conic in the projective space

$$\{\xi_0 : \xi_1 : \xi_2\} = \left\{1+\alpha\beta\cos\Theta : -\beta+\alpha\cos\Theta : \sqrt{\alpha^2-\beta^2}\sin\Theta\right\}\,, \tag{1.201}$$

where α is the semi-major axis and β is a coordinate of the center of the conic in $\mathbb{P}^2$.

Semi-major axis a and eccentricity e of the Keplerian orbit in the Euclidean space are related with the parameters α and β of the projective space. The relationship is established after matching the Euclidean coordinates of pericenter and apocenter in equations (1.187) with similar points in the homogeneous coordinates

$$\xi = \frac{\xi_1}{\xi_0} = \frac{\alpha\cos\Theta - \beta}{1+\alpha\beta\cos\Theta} \tag{1.202}$$

$$\eta = \frac{\xi_2}{\xi_0} = \frac{\sqrt{\alpha^2-\beta^2}\sin\Theta}{1+\alpha\beta\cos\Theta} \tag{1.203}$$

$$r = \sqrt{\xi^2+\eta^2} = \frac{\alpha - \beta\cos\Theta}{1+\alpha\beta\cos\Theta}\ . \tag{1.204}$$

It yields for semi-major axis, a, and eccentricity, e, the following relationships

$$a = \frac{\alpha(1+\beta^2)}{1-\alpha^2\beta^2}\ , \qquad e = \frac{\beta(1+\alpha^2)}{\alpha(1+\beta^2)}\ . \tag{1.205}$$

Moreover, parameters

$$q = \frac{\alpha-\beta}{1+\alpha\beta}\ , \qquad k = \frac{1-\alpha\beta}{\alpha+\beta}\ , \tag{1.206}$$

and

$$1-e = \frac{(\alpha-\beta)(1-\alpha\beta)}{\alpha(1+\beta^2)}\ , \quad 1+e = \frac{(\alpha+\beta)(1+\alpha\beta)}{\alpha(1+\beta^2)}\ . \tag{1.207}$$

The orbit is circular for $\beta = 0$, elliptical for $\alpha\beta < 1$, parabolic for $\alpha\beta = 1$, and hyperbolic for $\alpha\beta > 1$. The orbit is degenerated to a straight line (collisional trajectory) if $\alpha = \beta$.

Kepler's equation for the projective anomaly Θ is obtained by using the law of conservation of angular momentum,

$$\xi\dot{\eta} - \dot{\xi}\eta = \sqrt{GMq(1+e)}\ . \tag{1.208}$$

Substituting the formulas for the corresponding quantities into this law, one obtains

$$\frac{\alpha-\beta\cos\Theta}{(1+\alpha\beta\cos\Theta)^2}\frac{d\Theta}{dt} = \sqrt{\frac{GM}{\alpha(1+\beta^2)}}\ . \tag{1.209}$$

In the case of elliptic orbit ($\alpha\beta < 1$), the integral with respect to Θ is reduced by the substitution

$$\tan\frac{\Theta}{2} = \sqrt{\frac{1+\alpha\beta}{1-\alpha\beta}}\ \tan\frac{E}{2}\ , \tag{1.210}$$

where E is the eccentric anomaly, into an ordinary form of Kepler's equation

$$E - e\sin E = n(t-t_0) + \mathcal{M}_0\ , \tag{1.211}$$

with

$$n = \sqrt{\frac{GM}{\alpha^3}\left(\frac{1-\alpha^2\beta^2}{1+\beta^2}\right)^3} = \sqrt{\frac{GM}{a^3}}\ , \tag{1.212}$$

being the mean orbital motion defined earlier in equation (1.159). If $\alpha = \beta$ (a parabolic orbit), the integral is reduced by the substitution

$$s = \tan\frac{\Theta}{2}\,, \tag{1.213}$$

into a cubic equation

$$s^3 - 3\frac{1-\alpha^2}{1+\alpha^2}\,s = 6\sqrt{\frac{GM}{\alpha(1+\alpha^2)^3}}(t-t_0) + \mathcal{M}_0\,. \tag{1.214}$$

A hyperbolic case with $\alpha\beta > 1$ is treated similarly to the elliptic case after replacements: $1 - \alpha\beta \to \alpha\beta - 1$ and $\tan(\Theta/2) \to \tanh(\Theta/2)$.

1.3.6
The Freedom of Choice of the Anomaly

One has already encountered several options for defining the instantaneous position of a body on its orbit as a function of time. One option was to keep $\mathcal{M}_0$ among the integration constants, and to use the time t as a variable defining the position of the body at each moment. This method is seldom employed in practice, because it is impossible to analytically express the components of $\boldsymbol{r}$ and $\dot{\boldsymbol{r}}$ as explicit functions of t. A better option would be to use the true anomaly f instead of the time. This method is more practical, as it is directly implemented by the explicit expressions (1.185), (1.186). Similarly, one can keep $\mathcal{M}_0$ among the constants and use the eccentric anomaly E instead of the time. It is also possible to unite the integration constant $\mathcal{M}_0$ and the epoch t_0 into one constant parameter τ defined in equation (1.174) as the instant of the body's first passage through the pericenter. This is possible, because in the unperturbed two-body problem these two quantities show up in the linear combination $\mathcal{M} = \mathcal{M}_0 + n\,(t - t_0)$, when the transition from the eccentric anomaly to the time is performed - see equation (1.166). Sometimes it is also convenient to employ the angular anomaly subtended the empty focus. A transition to this, so-called *anti-focal anomaly* was offered by Callandreau [1902a,b], and has proven to be very useful for numerical integration of weakly-perturbed elliptic trajectories with low eccentricities [Fukushima, 2004].

Thus, one sees that the freedom of parametrization is not exhausted by one's preferences in choosing the constants of integration. Another freedom lies in one's choice of the “fast" variable - the anomaly. One can exploit this freedom in analytical calculations by parameterizing the relationship between the time and the anomaly and keeping it arbitrary through the calculation. The arbitrariness is eliminated by breaking the freedom at the end to simplify the resulting expressions or to serve an other particular goal. The generalized anomaly W is *defined* as a solution to a simple differential equation

$$\frac{dE}{\sin E} = \frac{dW}{\sin W}\,, \tag{1.215}$$

that establishes the following trigonometric mapping between the eccentric anomaly E and the generalized anomaly W,

$$\tan\frac{W}{2} = \kappa\tan\frac{E}{2}\,, \tag{1.216}$$

with κ emerging as an integration constant and, thus, playing the role of a free constant parameter. It parameterizes a particular choice of the general anomaly W among a whole family of such anomalies. Equation (1.216) shows that the new anomaly W generalizes relationship (1.156) between the true and eccentric anomalies, and relationship (1.210) between the projective and eccentric anomalies.

Making use of equation (1.216), in order to express the eccentric anomaly E in terms of the generalized anomaly W and the parameter κ, and substituting the so-obtained expression to equations (1.149), (1.150), and (1.152), one can derive the following formulae for the perifocal coordinates defined in Figure 1.4,

$$\xi = \frac{\kappa^2-(1+e)/(1-e)+\left[\kappa^2+(1+e)/(1-e)\right]\cos W}{\kappa^2+1+(\kappa^2-1)\cos W}\,q\,, \tag{1.217}$$

$$\eta = \frac{2\kappa\sqrt{(1+e)/(1-e)}\sin W}{\kappa^2+1+(\kappa^2-1)\cos W}\,q\,, \tag{1.218}$$

and for the distance from the gravitating center to the orbiter,

$$r = \frac{\kappa^2+(1+e)/(1-e)+\left[\kappa^2-(1+e)/(1-e)\right]\cos W}{\kappa^2+1+(\kappa^2-1)\cos W}\,q\,, \tag{1.219}$$

where $q = a(1-e)$ is the distance from the pericenter to the focus wherein the attracting mass is located.

The generalized anomaly W was originally introduced into celestial mechanics by Subbotin [1936a,b] who used a different parameter σ instead of κ [19]. Introducing an auxiliary quantity

$$\chi = \sqrt{1-e^2+\sigma^2}\,, \tag{1.220}$$

Subbotin [1936a,b] was able to derive the following expressions for the perifocal coordinates:

$$\xi = a\left(\frac{\chi\cos W+\sigma}{\chi+\sigma\cos W}-e\right), \tag{1.221}$$

$$\eta = a\frac{\left(1-e^2\right)\sin W}{\chi+\sigma\cos W}\,. \tag{1.222}$$

To establish a relationship between the parameter κ standing in equations (1.216) - (1.219) and the parameter σ introduced by Subbotin, equate the expression (1.217)

19) In fact, Subbotin denoted his parameter α. Here, the notation σ is used to avoid confusion with the projective parameter α from previous section.

with (1.221), and that (1.218) with (1.222) for the perifocal coordinates. This will render

$$\chi = \frac{\kappa^2+1}{2\kappa}\sqrt{1-e^2}\,, \qquad \sigma = \frac{\kappa^2-1}{2\kappa}\sqrt{1-e^2}\,. \tag{1.223}$$

From here one sees that the parameter κ admits two values corresponding to one value of Subbotin's parameter σ,

$$\kappa_+ = \frac{\sigma+\chi}{\sqrt{1-e^2}}\,, \tag{1.224}$$

$$\kappa_- = \frac{\sigma-\chi}{\sqrt{1-e^2}}\,. \tag{1.225}$$

The value κ_+ corresponds to the generalized *focal* anomaly, W_+, and the value κ_- corresponds to the generalized *anti-focal anomaly*, W_-, that is

$$\tan\frac{W_+}{2} = \kappa_+\tan\frac{E}{2}\,, \qquad \tan\frac{W_-}{2} = \kappa_-\tan\frac{E}{2}\,. \tag{1.226}$$

One can see that

$$\tan\frac{W_-}{2}\tan\frac{W_+}{2} = -\tan^2\frac{E}{2}\,. \tag{1.227}$$

Now one can see that equations (1.221) and (1.222) are actually valid for the focal anomaly $W = W_+$. For the anti-focal anomaly W_-, the corresponding equations should read

$$\xi = a\left(\frac{\chi\sigma\cos W_-}{\chi-\sigma\cos W_-} - e\right)\,, \tag{1.228}$$

$$\eta = a\frac{\left(1-e^2\right)\sin W_-}{\chi-\sigma\cos W_-}\,. \tag{1.229}$$

For more details on Subbotin's anomalies see the paper by Sokolov [2009] who also corrected some misprints in Subbotin's works.

The advantage of parametrization (1.217) - (1.219) stems from the fact that it enables one to establish a correspondence between the generalized anomaly of Subbotin [1936a,b] and the projective anomaly Θ of Satō [1998]. Indeed, equations (1.217) - (1.219) tell us that the (focal) generalized anomaly W coincides with the eccentric anomaly E, for $\kappa = 1$ ($\sigma = 0$); with the true anomaly f, for $\kappa = \sqrt{(1+e)/(1-e)}$ ($\sigma = e$); and with the projective anomaly Θ, for $\kappa = \sqrt{(1+\alpha\beta)/(1-\alpha\beta)}$ corresponding to Subbotin's parameter $\sigma = \beta\sqrt{\alpha^2-\beta^2}/(1+\beta^2)$.

The freedom of the anomaly choice in the Newtonian celestial mechanics remains greatly under-exploited, the work by Fukushima [2004] being a rare exception. We believe, though, that the future use of this freedom will yield fruits. We also think that equally productive may be the use of this freedom in the relativistic two-body problem, including this problem's practical application in data processing of binary-pulsar-timings. We shall return to this topic in section 6.4.

1.4
A Perturbed Two-Body Problem

1.4.1
Prefatory Notes

Celestial mechanics of two-body problem is insufficient in many practical applications where one has to take into account gravitational perturbations exerted on the Keplerian motion by external agents. Hence, one needs a mathematical extension of the unperturbed two-body formalism to perturbed settings. One such setting is that of a binary system embedded to the gravitational field of N external bodies which affect the orbital motion of the binary through tidal forces. Other types of perturbations include triaxiality of the interacting bodies, as well as atmospheric drag, magnetic fields, tides, relativistic corrections, or forces of inertia if a non-inertial reference frame is used for calculations. Since the disturbing forces normally are small compared to the gravitational interaction between the two bodies, one may presume that each body moves along a conic that is *osculating* (tangent) at each instant of time to the actual physical trajectory and slowly evolving (see Figure 1.5). This approach was offered circa 1687 by Newton in his unpublished *Portsmouth Papers*. Very succinctly, in purely geometric terms, Newton mentioned it also in Corollaries 3 and 4 of Proposition 17 in the first book of his *Principia*. Implementation of this idea in the language of calculus was initiated by Euler [1748, 1753] and got its final shape in the works of Lagrange [1778, 1783, 1788a,b, 1808a,b, 1809].

Before explaining their developments, let us point out that the smallness of perturbations is, by itself, a rather shaky foundation for the varying-conic method. Indeed, ones are immediately faced by the following questions:

1) To what degree of rigor can a perturbed orbit be modeled with a family of instantaneously osculating conics having the primary body in one of their foci?
2) Does this modeling admit an exact mathematical formulation?
3) Is this representation of the perturbed orbit by a family of the osculating conics unique?

These questions will not seem trivial, if one recalls that the concept of evolving instantaneous conics had been introduced into practice (and that major developments of the disturbing-function theory had been accomplished) long before Frenet and Serret developed the theory of curves with their concept of the moving Frenet-Serret frame being closely associated with the *curvature* and *torsion* of the curve [Dubrovin *et al.*, 1984]. This order of historical events explains the reason why terms curvature and torsion of the orbit are rarely used in the dynamic astronomy books. Fortunately, Lagrange fortified his developments with the tools of calculus, which were powerful enough to completely surpass the theory of curves. Moreover, these tools in no way relied on the smallness of the disturbing forces. Hence, Lagrange's treatment of the problem already contains an affirmative answer to the first two questions. The answer to the third question, surprisingly, turns out to be negative. Below this point is explained in more detail and one demonstrates that celestial mechanics permits

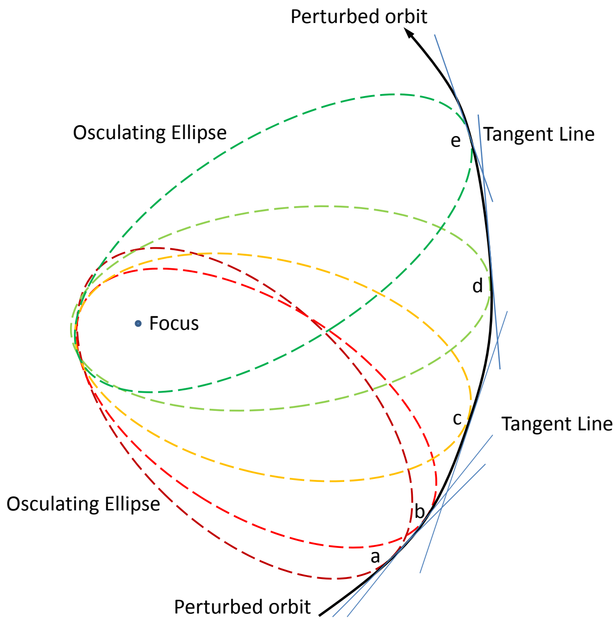

Figure 1.5 This picture illustrates the method of variation of parameters in application to the reduced two-body problem. A perturbed orbit can be presented as an envelope of a family of instantaneous conics sharing the common focus F. Each instantaneous conic is osculating - it touches the physical orbit, sharing with it the tangent line at the point of contact. The Keplerian parameters of the instantaneous conic evolve in time, as the body moves through the positions a, b, c, d, e. Be mindful that the method of variation of parameters, in application to this problem, implies variation of the elements of the instantaneous conic, but not of the position of the focus. Hence the instantaneous conics are always confocal.

internal freedom in description of perturbed orbits.

1.4.2
Variation of Constants. Osculating Conics

We shall start in the spirit of Lagrange [1808a,b, 1809], but shall soon deviate on two points. First, in distinction from Lagrange, one will not assume that the disturbing force is conservative and depends upon the positions solely, but shall permit it to depend also upon velocities. Second, one's intention is to relax eventually the Lagrange constraint i.e., the assumption that the instantaneous conics should be tangent to the resulting perturbed curve. This will bring up orbital variables, which will not be osculating yet mathematically useful.

In the modern, vectorial notations, Lagrange's line of reasoning looks as follows. A generic solution to the reduced two-body problem described by equation

$$\ddot{\boldsymbol{r}} + \frac{GM}{r^3}\boldsymbol{r} = 0 \;, \tag{1.230}$$

is a Keplerian conic that is defined by the set of six orbital elements $\{C_i\} = C_1, \ldots, C_6$ implementing the chosen orbital parametrization. In some fixed inertial Cartesian coordinate system, this conic reads

$$\boldsymbol{r} = \boldsymbol{r}\left(C_1, \ldots, C_6, t\right) \;, \qquad \dot{\boldsymbol{r}} = \boldsymbol{v}\left(C_1, \ldots, C_6, t\right) \;. \tag{1.231}$$

The expressions for orbital radius-vector, $\boldsymbol{r}$, and velocity, $\boldsymbol{v}$, were written down in the previous section. By definition, function $\boldsymbol{v}$ is the partial derivative of $\boldsymbol{r}$ with respect to time,

$$\boldsymbol{v} \equiv \left(\frac{\partial \boldsymbol{r}}{\partial t}\right)_{C_i=\text{const.}} \;. \tag{1.232}$$

Of course, since the orbital elements are constants of motion in the unperturbed two-body problem, the partial and ordinary time derivatives of vector $\boldsymbol{r}$, coincide.

The functions entering expression (1.231) can be used as an *Ansatz* for solving the perturbed two-body problem

$$\ddot{\boldsymbol{r}} + \frac{GM}{r^3}\boldsymbol{r} = \boldsymbol{F} \;, \tag{1.233}$$

with vector $\boldsymbol{F}$ being a known disturbing force of whatever nature (including inertial forces). To solve (1.233), one assumes that the perturbed orbit coincides at each instant of time with an *instantaneous* Keplerian conic. This way, by going smoothly from one instantaneous conic to another, one endows the orbital parameters C_i with a time-dependence of their own,

$$\boldsymbol{r} = \boldsymbol{r}\left[C_1(t), \ldots, C_6(t), t\right] \;, \tag{1.234}$$

keeping the functional form of $\boldsymbol{r}$ the same as in equation (1.231). As the parameters C_i are now time-dependent, the velocity of the body,

$$\dot{\boldsymbol{r}} = \frac{d\boldsymbol{r}}{dt} = \frac{\partial \boldsymbol{r}}{\partial t} + \sum_{i=1}^{6} \frac{\partial \boldsymbol{r}}{\partial C_i}\frac{dC_i}{dt} = \boldsymbol{v} + \sum_{i=1}^{6} \frac{\partial \boldsymbol{r}}{\partial C_i}\frac{dC_i}{dt} \;, \tag{1.235}$$

acquires an additional input besides $\boldsymbol{v}$, while the term $\boldsymbol{v}$ retains the same functional form as it used to have in the unperturbed setting (1.231).

Substitution of expression (1.235) into the perturbed equation of motion (1.233) gives birth to three independent scalar differential equations of the second order. These three equations contain one independent variable – time, and six time-dependent parameters, $C_i(t)$, whose evolution is to be determined. Evidently, this cannot be done in a unique way because the number of the parameters exceeds, by three, the number of equations. This means that though the *perturbed* orbit given by the locus of points in space and by the values of velocity at each of these points, is unique, its parametrization in terms of the orbital elements admits a certain freedom. The fact, that the system of differential equations for the parameters $C_i(t)$ is underdetermined was noticed by Lagrange in his treatment. To make it solvable, he decided to amend it with three supplementary conditions imposed on functions C_i and their *first* time derivatives. His choice was

$$\sum_{i=1}^{6} \frac{\partial \boldsymbol{r}}{\partial C_i} \frac{dC_i}{dt} = 0 \,, \tag{1.236}$$

a so-called *Lagrange constraint* that is often imposed in the theory of ordinary differential equations. Imposition of this supplementary constraint was motivated by both physical considerations and by Lagrange's desire to simplify calculations. Since, physically, the perturbed orbit $\boldsymbol{r}$ with a time-dependent set of orbital elements $\{C_i(t)\}$ can, at each fixed time t, be interpreted as an instantaneous conic, Lagrange decided to make the $C_i(t)$ osculating, that is to keep the instantaneous conics tangential to the perturbed trajectory, as displayed in Figure 1.5. This means that the physical trajectory of a body defined by $C_1(t), ..., C_6(t)$ must, at each instant of time, coincide *locally* with the unperturbed orbit that the moving body would follow if perturbations were to cease instantaneously. This can be achieved only when the dependence of the velocities upon the elements, in the perturbed setting, is the same as that in the original unperturbed case, $\dot{\boldsymbol{r}} = \boldsymbol{v}$. This, in turn, can be true only if the second term on the right side of (1.235) vanishes, i.e., if one sets the extra condition (1.236). This vector condition, the *Lagrange constraint*, consists of three scalar equations which, together with the three equations of motion (1.233), constitute a well-defined system of six equations for six variables $C_1(t), ..., C_6(t)$.

As it was recently pointed out in [Efroimsky, 2002a,b], the choice of the supplementary condition in the form of (1.236) is not always optimal. Moreover, as explained by Efroimsky and Goldreich [2003, 2004], in some important situations this choice is simply unavailable. We shall address this topic below in Section 1.5. For now, though, one sticks to the supplementary condition in the form of Lagrange's constraint (1.236).

1.4.3
The Lagrange and Poisson Brackets

As a prerequisite to the subsequent calculations, it would be of use to introduce the, so-called, Lagrange and Poisson brackets of the orbital elements. The *Lagrange bracket* of two elements, C_k and C_i, is denoted by $[C_k C_i]$, and the entire set of the Lagrange brackets form a 6×6 matrix. Each element of the matrix is defined as a certain linear combination of scalar products of partial derivatives of the components of vectors $\boldsymbol{r}(C_1, ..., C_6, t)$ and $\boldsymbol{v}(C_1, ..., C_6, t)$ with respect to the orbital elements C_i. Namely,

$$[C_k C_i] \equiv \frac{\partial \boldsymbol{r}}{\partial C_k} \cdot \frac{\partial \boldsymbol{v}}{\partial C_i} - \frac{\partial \boldsymbol{r}}{\partial C_i} \cdot \frac{\partial \boldsymbol{v}}{\partial C_k} \,, \tag{1.237}$$

where the dot between two vectors denote the Euclidean dot product.

The *Poisson bracket* of two elements, C_k and C_i, is denoted by $\{C_k C_i\}$. It is defined as a scalar product between two vectors that are partial derivatives of the orbital elements C_i with respect to coordinates and velocity of the body [20)]

$$\{C_k C_i\} \equiv \frac{\partial C_k}{\partial \boldsymbol{r}} \cdot \frac{\partial C_i}{\partial \boldsymbol{v}} - \frac{\partial C_k}{\partial \boldsymbol{v}} \cdot \frac{\partial C_i}{\partial \boldsymbol{r}} \,. \tag{1.238}$$

The 6×6 matrix of the Poisson brackets is the *negative inverse* to the matrix of the Lagrange brackets,

$$\sum_{i=1}^{6} [C_j C_i] \, \{C_i C_k\} = -\delta_{jk} \,, \tag{1.239}$$

where $\delta_{jk} = \mathrm{diag}(1, 1, 1, 1, 1, 1)$ $(j, k = 1, 2, ..., 6))$ is the Kronecker symbol (the unit matrix) in six-dimensional Euclidean space of the orbital parameters, C_i. This relation can be easily derived if one express the orbital parameters in terms of radius-vector and velocity, $C_i = C_i(\boldsymbol{r}, \boldsymbol{v})$, and apply the chain rule of differentiation,

$$\frac{\partial C_i}{\partial \boldsymbol{r}} \cdot \frac{\partial \boldsymbol{r}}{\partial C_j} + \frac{\partial C_i}{\partial \boldsymbol{v}} \cdot \frac{\partial \boldsymbol{v}}{\partial C_j} = \frac{\partial C_i}{\partial C_j} = \delta_{ij} \,, \tag{1.240}$$

in calculation of the product of two brackets in equation (1.239). Both the Lagrange and Poisson brackets are antisymmetric,

$$[C_k C_j] = -[C_j C_k] \,, \qquad \left\{C_k C_j\right\} = -\left\{C_j C_k\right\} \,. \tag{1.241}$$

The antisymmetry of the brackets evidently yields vanishing diagonal elements:

$$[C_k C_k] = 0 \,, \qquad \{C_k C_k\} = 0 \,. \tag{1.242}$$

20) In differential geometry, the partial derivatives with respect to the coordinates and velocity are associated with covectors. This detail is ignored here, because one works in a Euclidean space where vectors and covectors are formally equivalent to each other.

A remarkable property of both the Lagrangian and Poisson brackets, which greatly facilitates their evaluation, is that they do not depend on time explicitly [Brouwer and Clemence, 1961; Schaub and Junkins, 2003], that is their partial time derivatives vanish:

$$\frac{\partial}{\partial t}\left[C_k C_i\right] = 0\,, \qquad \frac{\partial}{\partial t}\left\{C_k C_i\right\} = 0\,. \tag{1.243}$$

To verify this, let us take the partial time derivative of the Lagrange brackets (1.237). One obtains

$$\frac{\partial}{\partial t}\left[C_k C_i\right] = \frac{\partial \boldsymbol{r}}{\partial C_k}\cdot\frac{\partial \boldsymbol{a}}{\partial C_i} - \frac{\partial \boldsymbol{r}}{\partial C_i}\cdot\frac{\partial \boldsymbol{a}}{\partial C_k}\,, \tag{1.244}$$

where $\boldsymbol{a} = \partial\boldsymbol{v}/\partial t$ is the orbital acceleration on an instantaneous conic. For Keplerian conics,

$$\boldsymbol{a} = -\frac{GM}{r^3}\boldsymbol{r} = \frac{\partial}{\partial \boldsymbol{r}}\left(\frac{GM}{r}\right)\,, \tag{1.245}$$

with $r = |\boldsymbol{r}|$. Hence,

$$\begin{aligned}\frac{\partial}{\partial t}\left[C_k C_i\right] &= \frac{\partial \boldsymbol{r}}{\partial C_k}\cdot\frac{\partial}{\partial \boldsymbol{r}}\frac{\partial}{\partial C_i}\left(\frac{GM}{r}\right) - \frac{\partial \boldsymbol{r}}{\partial C_i}\cdot\frac{\partial}{\partial \boldsymbol{r}}\frac{\partial}{\partial C_k}\left(\frac{GM}{r}\right) \\ &= \frac{\partial^2}{\partial C_k \partial C_i}\left(\frac{GM}{r}\right) - \frac{\partial^2}{\partial C_i \partial C_k}\left(\frac{GM}{r}\right) = 0\,,\end{aligned} \tag{1.246}$$

since the second partial derivatives commute. It proves that the Lagrange brackets bear no explicit dependence on the time variable. The proof that the Poisson brackets do not depend explicitly on time follows immediately after taking partial time derivative from both sides of equation (1.239).

Because of this remarkable property, it does not matter at which point of the instantaneous orbit one evaluates the brackets. Thus, one can chose the most convenient point of the orbit in order to reduce the amount of algebra involved. After the Lagrange brackets are found, the elements of the Poisson brackets (1.238) can be obtained by matrix inversion from equation (1.239). The results are presented in Table 1.1 and Table 1.2 below.

1.4.4
Equations of Perturbed Motion for Osculating Elements

One is now prepared to derive the equations describing evolution of the osculating elements C_i, with the Lagrange constrain (1.236) imposed. As the second term on the right side of equation (1.235) vanishes, one can write the acceleration as

$$\frac{d^2\boldsymbol{r}}{dt^2} = \frac{\partial \boldsymbol{v}}{\partial t} + \sum_{i=1}^{6}\frac{\partial \boldsymbol{v}}{\partial C_i}\frac{dC_i}{dt} = \frac{\partial^2 \boldsymbol{r}}{\partial t^2} + \sum_{i=1}^{6}\frac{\partial \boldsymbol{v}}{\partial C_i}\frac{dC_i}{dt}\,, \tag{1.247}$$

insertion whereof in the perturbed equation of motion (1.233) entails:

$$\frac{\partial^2 \boldsymbol{r}}{\partial t^2} + \frac{GM}{r^3}\boldsymbol{r} + \sum_{i=1}^{6}\frac{\partial \boldsymbol{v}}{\partial C_i}\frac{dC_i}{dt} = \boldsymbol{F}\,, \tag{1.248}$$

where $r \equiv |\boldsymbol{r}|$. The function $\boldsymbol{r}$ is, by definition, a Keplerian solution to the unperturbed two-body problem, with constant orbital elements. So this function obeys the unperturbed equation (1.230), which means that the sum of the first two terms on the left side of equation (1.248) must be equated to zero. This simplifies equation (1.248) to

$$\sum_{i=1}^{6} \frac{\partial \boldsymbol{v}}{\partial C_i} \frac{dC_i}{dt} = \boldsymbol{F} \,. \tag{1.249}$$

This is the equations of disturbed motion, written in terms of the osculating orbital elements $C_i = C_i(t)$. Together with Lagrange's constraint (1.236), they constitute a well-defined system of six equations that can be solved with respect to C_i for all $i = 1, 2, ..., 6$. However, the mathematical form of equation (1.249) is not optimal for finding its solution because the time derivatives dC_i/dt in the left side of this equation are algebraically coupled with vectors $\partial \boldsymbol{v}/\partial C_i$. Therefore, the next step is to decouple derivatives dC_i/dt from $\partial \boldsymbol{v}/\partial C_i$. It can be achieved with the formalism of the Lagrange brackets.

Let us make a Euclidean dot-product of both sides of equation (1.249) with vector, $\partial \boldsymbol{r}/\partial C_k$. It yields,

$$\sum_{i=1}^{6} \left(\frac{\partial \boldsymbol{r}}{\partial C_k} \cdot \frac{\partial \boldsymbol{v}}{\partial C_i} \right) \frac{dC_i}{dt} = \frac{\partial \boldsymbol{r}}{\partial C_k} \cdot \boldsymbol{F} \,, \tag{1.250}$$

where $k = 1, 2, ..., 6$, while the dot between the two vectors denotes their Euclidean dot product. Making a dot-product of the Lagrange constraint (1.236) with vector $\partial \boldsymbol{v}/\partial C_k$ brings about

$$\sum_{i=1}^{6} \left(\frac{\partial \boldsymbol{v}}{\partial C_k} \cdot \frac{\partial \boldsymbol{r}}{\partial C_i} \right) \frac{dC_i}{dt} = 0 \,, \tag{1.251}$$

Subtraction of equation (1.251) from (1.250) results in

$$\sum_{i=1}^{6} [C_i C_k] \frac{dC_k}{dt} = \frac{\partial \boldsymbol{r}}{\partial C_i} \cdot \boldsymbol{F} \,. \tag{1.252}$$

With aid of equation (1.239), the derivative dC_i/dt can be decoupled so that one obtains,

$$\frac{dC_i}{dt} = -\sum_{j=1}^{6} \left\{ C_i C_j \right\} \frac{\partial \boldsymbol{r}}{\partial C_j} \cdot \boldsymbol{F} \,, \tag{1.253}$$

which is a system of six ordinary differential equations of the first order for the osculating elements of the perturbed orbit. The equations are valid in an arbitrary Cartesian coordinates and for arbitrary parametrization of the Keplerian conic. For this reason, equations (1.253) have a wide range of applications in celestial mechanics. The method of variation of parameters is also used in the theory of ordinary differential equations for finding their general solutions.

Table 1.1 The Lagrange brackets for the Keplerian osculating elements $(C_1, C_2, C_3, C_4, C_5, C_6)$=$(a, e, i, \Omega, \omega, \mathcal{M}_0)$. The notation $b = a\sqrt{1-e^2}$ stands for the semi-minor axis.

$[ea] = 0,$	$[\mathcal{M}_0\ \Omega] = 0,$	$[\mathcal{M}_0\ \omega] = 0,$
$[i\ a] = 0,$	$[i\ e] = 0,$	$[\omega\ \Omega] = 0,$
$[\Omega\ a] = \frac{n\,b}{2}\ \cos i,$	$[\Omega\ e] = -\frac{n\,a^3\,e}{b}\cos i,$	$[\Omega\ i] = -n\,a\,b\ \sin i,$
$[\omega\ a] = \frac{nb}{2},$	$[\omega\ e] = -\frac{na^3 e}{b},$	$[\omega\ i] = 0,$
$[\mathcal{M}_0\ a] = \frac{na}{2},$	$[\mathcal{M}_0\ e] = 0,$	$[\mathcal{M}_0\ i] = 0.$

1.4.5
Equations for Osculating Elements in the Euler-Gauss form

Though equations (1.253) are invariant with respect to the change of coordinates and/or the orbital parametrization of the unperturbed orbit of two-body system, choosing an appropriate parametrization can simplify their right side, thus, facilitating solution. One of the most convenient parameterizations is given in terms of the Keplerian elements $(a, e, i, \Omega, \omega, \mathcal{M}_0)$ by formulae (1.185) - (1.186). From these formulae, one can calculate the matrix of the Lagrange brackets in a fairly straightforward way. The calculation can be greatly simplified if one recalls that the Lagrange brackets do not explicitly depend on time and, therefore, also upon the true anomaly, f, that may be set nil, $f = 0$. The outcome of the calculation is displayed in Table 1.1 where one has shown only 15 elements of the matrix since the anti-symmetry of the Lagrange brackets gives immediately the other 15 off-diagonal elements. Remaining 6 diagonal elements are, of course, identically zero. As many elements of the matrix of the Lagrange brackets are zero, it is relatively easy to invert the matrix in order to calculate the Poisson brackets. The result is presented in Table 1.2. Calculating the partial derivatives $\partial\boldsymbol{r}/\partial a$, $\partial\boldsymbol{r}/\partial e$, etc., from equations (1.185) - (1.186), forming their scalar dot-products with the perturbing force $\boldsymbol{F}$, and making transvection of these products with the Poisson brackets from Table 1.2 yield a system of ordinary

Table 1.2 The Poisson brackets for the Keplerian osculating elements $(C_1, C_2, C_3, C_4, C_5, C_6)=(a, e, i, \Omega, \omega, \mathcal{M}_0)$. The notation $b = a\sqrt{1-e^2}$ stands for the semi-minor axis.

$\{ea\} = 0,$	$\{\mathcal{M}_0\ \Omega\} = 0,$	$\{\mathcal{M}_0\ \omega\} = 0,$
$\{i\ a\} = 0,$	$\{i\ e\} = 0,$	$\{\omega\ \Omega\} = 0,$
$\{\Omega\ a\} = 0,$	$\{\Omega e\} = 0,$	$\{\Omega\ i\} = -\frac{1}{n\,a\,b\ \sin i},$
$\{\omega\ a\} = 0,$	$\{\omega e\} = -\frac{b}{n\,a^3\,e},$	$\{\omega\ i\} = \frac{\cos i}{n\,a\,b\ \sin i},$
$\{\mathcal{M}_0\ a\} = \frac{2}{n\,a},$	$\{\mathcal{M}_0\ e\} = \frac{b^2}{n\,a^4\,e},$	$\{\mathcal{M}_0\ i\} = 0.$

differential equations for the Keplerian elements [21] in the *Euler-Gauss form*:

$$\frac{da}{dt} = \frac{2}{n\sqrt{1-e^2}}\left(eF_R \sin f + F_T\frac{p}{r}\right) , \tag{1.254a}$$

$$\frac{de}{dt} = \frac{\sqrt{1-e^2}}{na}\left[F_R \sin f + F_T\left(\cos f + \cos E\right)\right] , \tag{1.254b}$$

$$\frac{di}{dt} = \frac{r\cos(f+\omega)}{na^2\sqrt{1-e^2}}F_N , \tag{1.254c}$$

$$\frac{d\Omega}{dt} = \frac{r\sin(f+\omega)}{na^2\sqrt{1-e^2}\sin i}F_N , \tag{1.254d}$$

$$\frac{d\omega}{dt} = -\cos i\frac{d\Omega}{dt} + \frac{\sqrt{1-e^2}}{nae}\left[-F_R\cos f + F_T\left(1+\frac{r}{p}\right)\sin f\right] , \tag{1.254e}$$

$$\frac{d\mathcal{M}_0}{dt} = -\sqrt{1-e^2}\left(\frac{d\omega}{dt} + \cos i\frac{d\Omega}{dt}\right) - \frac{2r}{na^2}F_R , \tag{1.254f}$$

where the radial distance r has to be expressed in terms of the orbital elements of the two-body problem,

$$r = \frac{a(1-e^2)}{1+e\cos f} , \tag{1.255}$$

while F_R, F_T, and F_N are the radial, transversal, and normal to the orbit components of the perturbing force $\boldsymbol{F}$,

$$F_R = \boldsymbol{n}\cdot\boldsymbol{F} , \qquad F_T = (\boldsymbol{k}\times\boldsymbol{n})\cdot\boldsymbol{F} , \qquad F_N = \boldsymbol{k}\cdot\boldsymbol{F} . \tag{1.256}$$

with the unit vector $\boldsymbol{n} = \boldsymbol{r}/r$ and the unit vector $\boldsymbol{k}$ is set to be orthogonal to the instantaneous orbital plane. The mean anomaly of the perturbed motion obeys the

21) These equations are called in celestial mechanics the *planetary equations*.

following equation,

$$\mathcal{M} = \mathcal{M}_0 + \int_{t_0}^{t} n(t')dt' \ , \tag{1.257}$$

where $\mathcal{M}_0$ is solution of equation (1.254f), and $n(t)$ is the mean orbital frequency given by equation (1.159) where the semi-major axis $a = a(t)$ is solution of the corresponding equation (1.254a).

Equations (1.254a)-(1.254f) can be also independently derived by differentiating the two vectorial integrals of motion, the angular-momentum vector $\boldsymbol{J}$ and the Laplace-Runge-Lenz vector $\boldsymbol{A}_{\rm L}$, which are not conserved in the perturbed motion. Differentiating their vectorial definitions, (1.108) and (1.176), and making use of the equations of perturbed motion (1.233), result in:

$$\frac{d\boldsymbol{J}}{dt} = \boldsymbol{r}\times\boldsymbol{F} \ , \tag{1.258}$$

$$\frac{d\boldsymbol{A}_{\rm L}}{dt} = 2(\dot{\boldsymbol{r}}\cdot\boldsymbol{F})\boldsymbol{r} - (\boldsymbol{r}\cdot\boldsymbol{F})\dot{\boldsymbol{r}} - (\boldsymbol{r}\cdot\dot{\boldsymbol{r}})\boldsymbol{F} \ . \tag{1.259}$$

As the Lagrange's principle of the variation of orbital elements demands that all relationships of the unperturbed Keplerian orbit remained valid in the perturbed motion, one can also use the expressions (1.177) and (1.178) of these vector integrals for calculating the time derivatives. Differentiating equations (1.177) and (1.178) and keeping in mind that the unit vectors $\boldsymbol{k}$ and $\boldsymbol{P}$ are also functions of time via equations (1.111) and (1.141), lead us to the following results,

$$\frac{d\boldsymbol{J}}{dt} = \frac{na}{2\sqrt{1-e^2}}\frac{dp}{dt}\boldsymbol{k} + na^2\sqrt{1-e^2}\left(\boldsymbol{l}\sin i\frac{d\Omega}{dt} - \boldsymbol{m}\frac{di}{dt}\right) \ , \tag{1.260}$$

$$\begin{aligned}\frac{d\boldsymbol{A}_{\rm L}}{dt} &= GMe\left[\left(\frac{d\omega}{dt} + \cos i\frac{d\Omega}{dt}\right)\boldsymbol{Q} + \left(\sin\omega\frac{di}{dt} - \cos\omega\sin i\frac{d\Omega}{dt}\right)\boldsymbol{k}\right] \\ &+ GM\frac{de}{dt}\boldsymbol{P} \ ,\end{aligned} \tag{1.261}$$

where the unit vectors $\boldsymbol{P}$ and $\boldsymbol{Q}$ have been defined in equations (1.141), (1.142) and are also shown in Figure 1.3. After decomposing the perturbing force in the three components (1.256), and equating right sides of the corresponding equations, (1.258) and (1.260) as well as (1.259) and (1.261), it can then be demonstrated that the Euler-Gauss equations (1.254) are again obtained.

As a historical aside, one would mention that in his work on the lunar motion, Euler [1753] derived the planetary equations for the longitude of the node, Ω, the inclination, i, and the semilatus rectum, p, with the time derivatives of these three elements expressed through the three components of the disturbing force $\boldsymbol{F}$. Sixty years later, the method was amended by Gauss, who wrote down similar equations for the other three Keplerian elements, thus obtaining a full system of the planetary equations. Although many books refer to the system of equations (1.254) as the *Gauss* equations, it looks more fair to pay the tribute evenly, calling them as the *Euler-Gauss* equations.

1.4.6
The Planetary Equations in the form of Lagrange

So far, one did not impose any limitation on the functional form of the perturbing force. Now, let us assume that the force is conservative and depends only on coordinates of perturbing bodies. This assumption was made by Lagrange [1778, 1783, 1788a,b, 1808a,b, 1809] in his treatment of planetary motions in the solar system. Being dependent solely on the positions of the moving bodies, this force can be expressed as a gradient of the *disturbing function*,

$$\boldsymbol{F} = \frac{\partial R}{\partial \boldsymbol{r}} = -\frac{\partial \mathfrak{W}}{\partial \boldsymbol{r}} , \tag{1.262}$$

with the *disturbing function*, R, being negative to the disturbing potential $\mathfrak{W}$ introduced above in equation (1.96). The chain rule for partial derivatives provides us with

$$\frac{\partial \boldsymbol{r}}{\partial C_i} \cdot \boldsymbol{F} = \frac{\partial \boldsymbol{r}}{\partial C_i} \cdot \frac{\partial R}{\partial \boldsymbol{r}} = \frac{\partial R}{\partial C_i} . \tag{1.263}$$

Hence, substituting the force (1.262) into equation (1.253) for osculating elements results in

$$\frac{dC_i}{dt} = -\sum_{k=1}^{6} \{C_i C_k\} \frac{\partial R}{\partial C_k} . \tag{1.264}$$

Finally, insertion of the Poisson brackets from Table 1.2 to equation (1.264) takes us to the celebrated system of planetary equations in the form of Lagrange:

$$\frac{da}{dt} = \frac{2}{na}\frac{\partial R}{\partial \mathcal{M}_0} , \tag{1.265a}$$

$$\frac{de}{dt} = \frac{1-e^2}{na^2 e}\frac{\partial R}{\partial \mathcal{M}_0} - \frac{\sqrt{1-e^2}}{na^2 e}\frac{\partial R}{\partial \omega} , \tag{1.265b}$$

$$\frac{di}{dt} = \frac{\cos i}{na^2\sqrt{1-e^2}\sin i}\frac{\partial R}{\partial \omega} - \frac{1}{na^2\sqrt{1-e^2}\sin i}\frac{\partial R}{\partial \Omega} , \tag{1.265c}$$

$$\frac{d\Omega}{dt} = \frac{1}{na^2\sqrt{1-e^2}\sin i}\frac{\partial R}{\partial i} , \tag{1.265d}$$

$$\frac{d\omega}{dt} = -\frac{\cos i}{na^2\sqrt{1-e^2}\sin i}\frac{\partial R}{\partial i} + \frac{\sqrt{1-e^2}}{na^2 e}\frac{\partial R}{\partial e} , \tag{1.265e}$$

$$\frac{d\mathcal{M}_0}{dt} = -\frac{1-e^2}{na^2 e}\frac{\partial R}{\partial e} - \frac{2}{na}\frac{\partial R}{\partial a} . \tag{1.265f}$$

The advantage of the Lagrange planetary equations is that the right sides of (1.265) depend on a single function R, in contrast to the Euler-Gauss equations (1.254) whose right sides depend on three components of the disturbing force $\boldsymbol{F}$.

Disadvantages of employing the Lagrange (and the Euler-Gauss) equations are that the right ascension of the ascending node becomes indeterminate as the inclination tends to zero, and the argument of perigee becomes indeterminate as the eccentricity tends to zero. The difficulty is, of course, of a purely mathematical nature, and has nothing to do with the actual motion. It can be sidestepped by switching from the Keplerian elements to the, so-called, equinoctial orbital elements $(\mathfrak{p}, \mathfrak{f}, \mathfrak{g}, \mathfrak{h}, \mathfrak{k}, \mathfrak{L})$, that are related to the Keplerian orbital parameters as follows [Broucke and Cefola, 1972; Walker *et al.*, 1985];

$$\begin{aligned}
\mathfrak{p} &= a(1-e^2)\,,\\
\mathfrak{f} &= e\cos(\Omega+\omega)\\
\mathfrak{g} &= e\sin(\Omega+\omega)\\
\mathfrak{h} &= \tan\tfrac{i}{2}\cos\Omega\,,\\
\mathfrak{k} &= \tan\tfrac{i}{2}\sin\Omega\,,\\
\mathfrak{L} &= \Omega+\omega+f\,.
\end{aligned}$$

The equinoctial orbital elements are useful for trajectory analysis and optimization of space flights. They are valid for circular, elliptic, and hyperbolic orbits. The Lagrange equations for equinoctial elements exhibit no singularity for zero eccentricity and orbital inclinations equal to 0 and 90 degrees. However, two of the components are singular for an orbital inclination of 180 degrees but this can be handled by an appropriate re-definition.

1.4.7
The Planetary Equations in the form of Delaunay

Another advantageous set of orbital elements is the Delaunay variables l, g, h, L, G, H. In terms of the Keplerian orbital elements, they are defined as

$l \equiv \mathcal{M}_0$ – the mean anomaly at epoch,
$g \equiv \omega$ – the argument of the pericenter,
$h \equiv \Omega$ – the longitude of the ascending node,
$L \equiv \sqrt{GMa}$,
$G \equiv L\sqrt{1-e^2}$ – the magnitude of the angular-momentum vector $\boldsymbol{J}$,
$H \equiv G\cos i$ – the angular-momentum component normal to equatorial plane.

By using the chain rule of differentiation, one can easily re-write the planetary equations (1.265) in terms of these variables:

$$\frac{dL}{dt} = \frac{\partial R}{\partial l}\,, \qquad \frac{dl}{dt} = -\frac{\partial R}{\partial L} \tag{1.266a}$$

$$\frac{dG}{dt} = \frac{\partial R}{\partial g}\,, \qquad \frac{dg}{dt} = -\frac{\partial R}{\partial G}\,, \tag{1.266b}$$

$$\frac{dH}{dt} = \frac{\partial R}{\partial h}\,, \qquad \frac{dh}{dt} = -\frac{\partial R}{\partial H}\,, \tag{1.266c}$$

where one deliberately wrote the equations in pairs, to emphasize their symplectic structure with the Hamiltonian being equal to the disturbing function R. In both sets of equations, (1.265) and (1.266), the element, $\mathcal{M}_0 = l$, can be substituted by the mean anomaly, $\mathcal{M}$, provided R is simultaneously substituted with $R' = R + GM/2a$. The advantage of the Delaunay equations is that they can be analyzed with the powerful mathematical technique of a symplectic geometry [Arnold, 1995]. Just as the Euler-Gauss and Lagrange planetary equations, the equations in the form of Delaunay become singular both in the limits of circular and/or equatorial orbits. Accordingly, a transition to the, so-called, canonical elements of Poincar'e can be an option [Brouwer and Clemence, 1961, p. 540].

1.4.8 Marking a Minefield

The most logical way of introducing the Delaunay variables would be to start out with the polar coordinates and their conjugate momenta, and to carry out the Hamilton-Jacobi procedure to find a canonical transformation to variables, which will remain mutually conjugated with respect to a vanishing Hamiltonian, i.e., they will be canonical integrals of motion [Arnold, 1995; Landau and Lifshitz, 1975]. These canonical variables are found through solution of the Hamilton-Jacobi equation, as demonstrated in numerous books - see, for example[22], [Plummer, 1918, sections 135 - 136], [Kovalevsky, 1967, sections 25 - 32], or [Vinti, 1998, chapter 6]. Within this approach, the Lagrange-type planetary equations are derived from those of Delaunay.

Unfortunately, neither of these books address the following important question: will the Hamilton-Jacobi procedure always result in *osculating* canonical elements? As one will see in the subsequent sections, the answer to this question is affirmative if the disturbance depends solely on positions of bodies, and is negative for velocity-dependent perturbations. This happens, because the Hamilton-Jacobi procedure implies that the resulting Delaunay elements are canonical, while the condition of canonicity becomes incompatible with the condition of osculation when the disturbance depends not only on the coordinates but also upon velocities.

As a result, the customary Delaunay and Lagrange planetary equations, when employed for such velocity-dependent perturbations, furnish orbital elements which are *not* osculating. In other words, the instantaneous conics parameterized by the so-obtained elements will not be tangent to the orbit. This means that these elements will render the correct position of the body, $\boldsymbol{r}$, but the partial derivative of $\boldsymbol{r}$ with respect to time will not provide its instantaneous velocity as in case of the osculating elements. While the use of such non-osculating elements is sometimes beneficial mathematically, their physical interpretation is not always apparent. Interestingly, the Andoyer elements which are the analogues of the Delaunay elements employed in the canonical description of a rigid-body rotation, are subject to the same reservation.

22) Plummer used notations β and β_2 for the negative Delaunay elements, $-l$ and $-g$, correspondingly.

1.5
Re-examining the Obvious

> Don't ever take a fence down until you know the reason it was put up.
> *G. K. Chesterton*

1.5.1
Why did Lagrange Impose his Constraint? Can It Be Relaxed?

When deriving the planetary equations, Lagrange amended the equations of motion (1.249) with constraint (1.236), in order to make the overall system of the equations well-defined mathematically. In the case of the conservative perturbing force, which are represented as a gradient of the perturbing potential, it provided a maximal simplification of the resulting equation (1.264) for osculating elements. Besides, the physical interpretation of the elements, C_i, obeying constraint (1.236) was geometrically straightforward. Indeed, by assuming that at each instant of time the perturbed velocity $\dot{\boldsymbol{r}}$ is equal to the unperturbed Keplerian velocity $\boldsymbol{v}$, Lagrange set the instantaneous conics tangent to the perturbed orbital curve and, thus, made the appropriate orbital elements osculating – see Figure 1.5.

It would then be natural to enquire if the Lagrange constraint should *always* be imposed on the orbital elements. Specifically: are there situations in which this constraint is not the best choice or is incompatible with a particular form of the equations for orbital elements? Indeed, a careful examination by Efroimsky and Goldreich [2003, 2004] reveals that the attempt of modeling of an orbit by tangential (osculating) confocal conics may be in conflict with the canonical equations. More specifically, if the perturbing function $R = R(\boldsymbol{r}, \boldsymbol{v})$, depends not only on positions of the bodies but also on their velocities[23] the demand of osculation comes into a contradiction with one's desire to keep the Delaunay variables canonical. Analytical solution of the Delaunay equations will then furnish an answer that will be mathematically consistent but it will come out in terms of *non-osculating* orbital elements violating the Lagrange constraint (1.236). The Delaunay elements in this case will render a correct instantaneous position of the perturbed body (satellite, planet) but a wrong value of its instantaneous velocity. In celestial calculations, this "booby trap" is often encountered but noticed fairly rare [Brumberg, 1972, 1991].

To explore the mathematical consequences of choosing a particular constraint imposed on the perturbed orbital parameters one will deliberately permit a measured degree of non-osculation in the differential equations describing their evolution. This gives us some, so-called, *gauge freedom* in choosing the constraint so that by imposing a constraint different from that of Lagrange one can sometimes greatly simplify

23) Examples of such forces are dissipative forces, the Coriolis forces in a precessing frame, or relativistic perturbations (see section 6.4).

the resulting equations for the orbital, but no longer osculating, elements. The situation resembles that emerging in the Maxwell electrodynamics where a suitable choice of the gauge imposed on electromagnetic vector potential can considerably simplify calculations without changing the physical observables. Planetary equations, obeying a constraint more general than that of Lagrange, were derived in Efroimsky [2002a,b]. Before addressing that formalism, one would show an elementary example of the gauge freedom of differential equations due to Newman and Efroimsky [2003], to illustrate the idea underlying the method.

1.5.2
Example: the Gauge Freedom of a Harmonic Oscillator

A one-dimensional harmonic oscillator with coordinate, $x = x(t)$, and disturbed by a force $F(t)$ obeys the second-order differential equation,

$$\ddot{x} + \omega_0^2 x = F(t) \, , \tag{1.267}$$

where overdot denotes a time derivative and ω_0 is the oscillation frequency of unperturbed motion. We also impose some initial conditions, $x(0)$, and, $\dot{x}(0)$, at the initial instant of time $t_0 = 0$.

The method of variation of parameters suggests that solution of equation (1.267) be sought for using a solution of homogeneous equation with the integration constants replaced with yet unknown functions of time,

$$x = C_1(t) \sin \omega_0 t + C_2(t) \cos \omega_0 t \, , \tag{1.268}$$

where $C_1(t)$ and $C_2(t)$ are to be determined. Differentiation of $x(t)$ will lead us to

$$\dot{x} = \dot{C}_1(t) \sin \omega_0 t + \dot{C}_2(t) \cos \omega_0 t + \omega_0 \left[C_1(t) \cos \omega_0 t - C_2(t) \sin \omega_0 t \right] \, . \tag{1.269}$$

It is common, at this point, to set the sum, $\dot{C}_1(t) \sin \omega_0 t + \dot{C}_2(t) \cos \omega_0 t$, equal to zero, in order to remove the indeterminacy which stems from having only one equation for the two variables, $C_1(t)$ and $C_2(t)$. This is equivalent to imposing the Lagrange constraint which is convenient though not obligatory. Let us see what will happen if one does not impose this particular constraint by assuming that

$$\dot{C}_1(t) \sin \omega_0 t + \dot{C}_2(t) \cos \omega_0 t = \Phi(t) \, , \tag{1.270}$$

with $\Phi(t)$ being an arbitrary smooth function of time. We call equation (1.270) the *gauge condition* and $\Phi(t)$ as the *gauge function* because picking up various $\Phi(t)$ leads to different solutions for $C_1(t)$ and $C_2(t)$ without changing the solution of the original equation (1.267) for function $x(t)$, as demonstrated below.

Substituting the gauge condition (1.270) in equation (1.269) and differentiating one more time, entails

$$\ddot{x} = \dot{\Phi} + \omega_0 \left[\dot{C}_1(t) \cos \omega_0 t - \dot{C}_2(t) \sin \omega_0 t \right] - \omega_0^2 \left[C_1(t) \sin \omega_0 t + C_2(t) \cos \omega_0 t \right] \, .$$

(1.271)

Replacing this result along with equation (1.268) in the original equation of motion (1.267), yields the dynamical equation rewritten in terms of the new variables, C_1 and C_2. Together with the *gauge condition* (1.270), it constitutes the following system of two differential equations,

$$\dot{\Phi} + \omega_0 \left[\dot{C}_1(t)\cos\omega_0 t - \dot{C}_2(t)\sin\omega_0 t\right] = F(t) \;, \tag{1.272a}$$

$$\dot{C}_1(t)\sin\omega_0 t + \dot{C}_2(t)\cos\omega_0 t = \Phi(t) \;. \tag{1.272b}$$

This system can be algebraically solved with respect to time derivatives of functions C_1 and C_2:

$$\frac{dC_1}{dt} = \omega_0^{-1}\left[F\cos\omega_0 t - \frac{d}{dt}\left(\Phi\cos\omega_0 t\right)\right] \;, \tag{1.273a}$$

$$\frac{dC_2}{dt} = \omega_0^{-1}\left[-F\sin\omega_0 t + \frac{d}{dt}\left(\Phi\sin\omega_0 t\right)\right] \;, \tag{1.273b}$$

with the initial conditions

$$C_1(0) = \frac{\dot{x}(0) - \Phi(0)}{\omega_0} \;, C_2(0) = x(0) \qquad , \tag{1.274}$$

imposed on the variables C_1 and C_2 in terms of the known values of $x(0)$ and $\dot{x}(0)$ taken at the time $t_0 = 0$. Notice that the initial value, $\Phi(0)$, of the gauge function remains arbitrary. Clearly, equations (1.273) are a simple analogue to the Lagrange system of planetary equations, whence one expects that the concept of gauge freedom may be equally applicable to the planetary equations of celestial mechanics.

In the example under consideration, the unperturbed problem was deliberately chosen extremely simple - a harmonic oscillator. So, one ended up with a very simple system of equations for variables, C_1 and C_2, with the gauge-dependent terms being the total time-derivatives. While, in general, one would have arrived to a system of differential equations that can be integrated only numerically, in this simple case the analytical integration is possible,

$$C_1(t) = C_1(0) + \omega_0^{-1}\left[\int_0^t F(t')\cos\omega_0 t' dt' - \Phi(t)\cos\omega_0 t\right] \;, \tag{1.275a}$$

$$C_2(t) = C_2(0) - \omega_0^{-1}\left[\int_0^t F(t')\sin\omega_0 t' dt' - \Phi(t)\sin\omega_0 t\right] \;. \tag{1.275b}$$

One sees that the solution for functions, $C_1(t)$ and $C_2(t)$, explicitly depends on the gauge function, Φ, which vanishes if the Lagrange constraint, $\Phi = 0$, is imposed. On the other hand, substitution of equations (1.275) in equation (1.268) leads to a complete cancelation of the Φ-dependent terms:

$$x = \omega_0^{-1}\int_0^t F(t')\sin\omega_0(t-t')dt' + C_1(0)\sin\omega_0 t + C_2(0)\cos\omega_0 t \;, \tag{1.276}$$

where the first term is a particular solution of the inhomogeneous equation, and the last two terms is a general solution of a homogeneous equation (1.267). This simple exercise proves that the physical trajectory, $x = x(t)$, of the perturbed oscillator remains invariant irrespectively of the choice of the gauge function, $\Phi(t)$, though the mathematical description (1.275) of its motion in terms of the variables $C_1(t)$ and $C_2(t)$ *is* gauge-dependent and rather arbitrary up to the following transformation of the variables

$$C_1 \longrightarrow \tilde{C}_1 = C_1 + \Phi(t)\cos\omega_0 t\ , \qquad C_2 \longrightarrow \tilde{C}_2 = C_2 - \Phi(t)\sin\omega_0 t\ . \tag{1.277}$$

This gauge-dependence of the variables C_1 and C_2, if not taken into account properly, may greatly influence the numerical error in finding solution for $x(t)$. Indeed, in settings more complicated than the perturbed harmonic pendulum, a choice of gauge may change the numerical error of integration by several orders of magnitude [Gurfil and Klein, 2006]. Specifically, choosing the Lagrange constraint $\Phi(t) = 0$ is not necessarily optimal.

An equally important feature illustrated by this example may also concern timescales. Suppose the unperturbed oscillator's frequency, ω_0, is much higher than an upper cut-off frequency, γ_0, of the spectrum of the perturbing force $F(t) \sim A\cos\gamma_0 t$. Naively, one may expect that a "slow" disturbance would cause an appropriately slow modulation of $C_1(t)$ and $C_2(t)$ in the perturbed problem. That this is not necessarily so, can be easily seen after integration of equations (1.275), where the slow and fast frequencies mix under the integral. The perturbing force brings about the following "fast" components to the solution

$$C_1(t) \sim \frac{A\cos\gamma_0 t}{\omega_0^2 - \gamma_0^2}\sin\omega_0 t\ , \qquad C_2(t) \sim -\frac{A\cos\gamma_0 t}{\omega_0^2 - \gamma_0^2}\cos\omega_0 t\ , \tag{1.278}$$

This tells us that in principle, $C_1(t)$ and $C_2(t)$, can undergo fast changes even under a slowly-evolving, $\gamma_0 \ll \omega_0$, disturbance. To a numerist, this indicates that the integration step used in solving equations (1.275) should not be much larger than the one employed in the unperturbed setting. Again, in some situations a clever choice of gauge function Φ may relax this restriction. As demonstrated by Gurfil and Klein [2006], a special choice of gauge turns the integration problem, in the linear case, into a simple quadrature. This means that for a complicated systems with significant timescale differences, for which integration of the variational equations with the Lagrange constraint using a fixed time step is impossible, the variational form of the equations with the specially-adapted gauge, $\Phi(t)$, can be integrated using a fixed time step.

1.5.3
Relaxing the Lagrange Constraint in Celestial Mechanics

1.5.3.1 *The Gauge Freedom*

Recall that a solution to the unperturbed equation of motion (1.230) of a restricted two-body problem is a conic whose functional form can be denoted with

$$\boldsymbol{r} = \boldsymbol{r}\,(C_1, ..., C_6,\, t) \,, \tag{1.279a}$$

$$\dot{\boldsymbol{r}} = \frac{\partial \boldsymbol{r}\,(C_1, ..., C_6,\, t)}{\partial t} \,, \tag{1.279b}$$

where, here and everywhere else, overdot is standing for a total time derivative d/dt, C_i $(i = 1, 2, ..., 6)$ are the constants of integration that do not depend on time. In the presence of a perturbing force, $\boldsymbol{F}$, the two-body system obeys equation (1.233). Solving the perturbed equation (1.233) by the method of variation of parameters implies that the functional form of the solution for the perturbed radius-vector, $\boldsymbol{r}$, remains the same as in the unperturbed problem

$$\boldsymbol{r} = \boldsymbol{r}\,[C_1(t), ..., C_6(t),\, t] \,, \tag{1.280}$$

while the constants of integration become functions of time. The perturbed velocity of the body is given by the total time derivative

$$\dot{\boldsymbol{r}} = \boldsymbol{v} + \boldsymbol{\Phi} \,, \tag{1.281}$$

where the vector function,

$$\boldsymbol{v} = \frac{\partial \boldsymbol{r}\,[C_1(t), ..., C_6(t),\, t]}{\partial t} \,, \tag{1.282}$$

has the same functional form as the unperturbed two-body velocity (1.279b), and

$$\boldsymbol{\Phi} = \sum_{i=1}^{6} \frac{\partial \boldsymbol{r}}{\partial C_i} \dot{C}_i \,, \tag{1.283}$$

denotes a, so-called, *gauge function* taking into account that the orbital elements of the perturbed motion are not kept constant any longer but become functions of time.

Standard procedure in application of the method of variation of variables is to use the Lagrange constraint, $\boldsymbol{\Phi} = 0$. Let us step away from the standard procedure and explore what will happen if one does not set the function $\boldsymbol{\Phi}$ nil but keep it unconstrained. Then, one can proceed further to calculate the acceleration of the body,

$$\ddot{\boldsymbol{r}} = \frac{\partial \boldsymbol{v}}{\partial t} + \sum_{i=1}^{6} \frac{\partial \boldsymbol{v}}{\partial C_i} \dot{C}_i + \dot{\boldsymbol{\Phi}} \,. \tag{1.284}$$

After substituting this result in the perturbed equation of motion (1.233) and recalling that the method of variation of parameters implies, $\partial \boldsymbol{v}/\partial t = -GM\boldsymbol{r}/r^3$, one will obtain three equations of motion for the variables $C_i(t)$,

$$\sum_{i=1}^{6} \frac{\partial \boldsymbol{v}}{\partial C_i} \frac{dC_i}{dt} + \frac{d\boldsymbol{\Phi}}{dt} = \boldsymbol{F} . \tag{1.285}$$

This equation should be compared with equation (1.249) that was derived on the basis of the Lagrange constraint. Equation (1.285) clearly demonstrate that the system of ordinary differential equations of the second order for the orbital elements, $C_i = C_i(t)$, admits rather large freedom of transformations associated with the *gauge function* $\boldsymbol{\Phi}$.

To find $C_i(t)$, one will have to solve a system comprised by the equations of motion (1.285) and the expression (1.283), which so far is merely a notation for the yet unspecified, vector function $\boldsymbol{\Phi}$. The identity (1.283) will become an additional differential equation for variables, $C_i(t)$, if one chooses a particular functional form for the gauge function $\boldsymbol{\Phi} = \boldsymbol{\Phi}\,[C_1(t), ..., C_6(t)\; t]$ as a function of time and the variables $C_i(t)$. The necessity to fix a functional form of $\boldsymbol{\Phi}$, i.e., to impose three additional differential conditions upon $C_i(t)$, evidently follows from the fact that one has six variables $C_i(t)$ while the number of equations of motion (1.285) is only three. What functional form to attribute eventually to $\boldsymbol{\Phi}$ will depend on the specific type of the perturbation. This *gauge freedom* of the differential equations of the perturbed motion corresponds to a specific freedom of transformations in the space of six parameters C_i. Particular example of these transformations is delivered by the canonical transformations preserving the form-invariance of the Hamiltonian equations like the Delaunay equations [Arnold, 1995, Section 44]. Our approach, however, goes beyond the canonical transformations and includes more general class of transformations of the orbital elements which is discussed in section 1.5.3.2. The gauge freedom of the solutions of differential equations of motion can be used in:

1) computer simulations of orbits where one's choice of a gauge considerably influences the error propagation process. A good or bad choice of gauge function, $\boldsymbol{\Phi}$, can optimize or destroy the numerical procedure. Specifically, the Lagrange gauge, $\boldsymbol{\Phi} = 0$, is not guaranteed to be always optimal.
2) analytical treatment, in order to simplify the integration procedure, perhaps, reducing it to quadratures.

The functional dependence of $\boldsymbol{\Phi}$ can be chosen arbitrary insofar as its substitution in equation (1.283) entails no conflict with the equations of motion in the sense that if a specific function, $\boldsymbol{\Phi}$, is chosen in equation (1.283), exactly the same function should appear in the equations of motion (1.285), and vise versa. The caveat here is that taking a particular form of the equations of motion (1.285) also fixes the gauge function $\boldsymbol{\Phi}$ which may be not nil. If this fact is overlooked and the Lagrange constraint, $\boldsymbol{\Phi} = 0$, is used, it will lead to erroneous solution for C_i. For example, taking the perturbed equations for the Delaunay canonical variables, fixes the gauge. Under position-dependent disturbances the gauge coincides with the Lagrange constraint $\boldsymbol{\Phi} = 0$, and the resulting Delaunay elements are osculating. However, in case of velocity-dependent perturbations the Delaunay gauge turns out to be different from the Lagrange constraint, so the ensuing Delaunay elements must be treated as nonosculating. If one works with the gauge function $\boldsymbol{\Phi}$ properly, its particular choice will never influence the eventual solution for the physical variable $\boldsymbol{r}$, similar to that having been discussed in previous section 1.5.2 for one-dimensional case of harmonic oscillator.

1.5.3.2 *The Gauge Transformations*

As emphasized above the split of the orbital velocity, $\dot{\boldsymbol{r}}$, given by equation (1.281), is not unique since there is no any limitation on the freedom of choice of the gauge function $\boldsymbol{\Phi}$. It means that the solution of the perturbed problem of motion given in terms of the orbital elements C_i admits a large freedom of the infinitesimal gauge transformations of the variables generated by various choices of $\boldsymbol{\Phi}$. The gauge transformation of the variables is given by equation

$$\tilde{C}_i = C_i(t) + \alpha_i(C_k, t) \;, \qquad (i = 1, 2, ..., 6) \tag{1.286}$$

where α_i are smooth functions of the "old" variables $C_i = C_i(t)$ $(i = 1, 2, ..., 6)$, and the time t. The group of the transformations is defined by the condition that the coordinate position of the body has the same value under the change (1.286) of the variables,

$$\boldsymbol{r}\left[\tilde{C}_1(t), ..., \tilde{C}_6(t)\,, t\right] = \boldsymbol{r}\left[C_1(t), ..., C_6(t)\,, t\right] \;, \tag{1.287}$$

and the functional form of equation (1.283) remains the same

$$\boldsymbol{\Phi} = \sum_{j=1}^{6} \frac{\partial \boldsymbol{r}}{\partial C_j}\dot{C}_j \;, \qquad \tilde{\boldsymbol{\Phi}} = \sum_{i=j}^{6} \frac{\partial \boldsymbol{r}}{\partial \tilde{C}_j}\dot{\tilde{C}}_j \;, \tag{1.288}$$

but the gauge functions $\boldsymbol{\Phi} = \boldsymbol{\Phi}(C_i, t)$ and $\tilde{\boldsymbol{\Phi}} = \tilde{\boldsymbol{\Phi}}(\tilde{C}_i, t)$ are different: $\boldsymbol{\Phi} \neq \tilde{\boldsymbol{\Phi}}$.

In order to derive relationship between functions α_i and the gauge function $\boldsymbol{\Phi}$, let us expand the left side of equation (1.287) in the Taylor series with respect to α_i which is considered as a small parameter of the expansion. Canceling the radius-vector $\boldsymbol{r}(C_i, t)$ in both parts of the equation one obtains an algebraic equation

$$\sum_{j=1}^{6} \frac{\partial \boldsymbol{r}}{\partial C_j}\alpha_j = 0 \,. \tag{1.289}$$

Taylor expansion of the second equation (1.288) with making use of equation (1.286), yields

$$\tilde{\boldsymbol{\Phi}} = \boldsymbol{\Phi} + \sum_{j=1}^{6} \frac{\partial \boldsymbol{r}}{\partial C_j}\frac{\partial \alpha_j}{\partial t} \,. \tag{1.290}$$

Taking a partial time derivative from equation (1.289) provides us with a useful equality

$$\sum_{j=1}^{6} \frac{\partial \boldsymbol{r}}{\partial C_j}\frac{\partial \alpha_j}{\partial t} = -\sum_{j=1}^{6} \frac{\partial \boldsymbol{v}}{\partial C_j}\alpha_j \,. \tag{1.291}$$

Making use of this identity in equation (1.290) transforms it to

$$\sum_{j=1}^{6} \frac{\partial \boldsymbol{v}}{\partial C_j} \alpha_j = \boldsymbol{\Phi} - \tilde{\boldsymbol{\Phi}} \,. \tag{1.292}$$

Next step is to make a dot product of equation (1.289) with vector $\partial \boldsymbol{v}/\partial C_i$ and a dot product of equation (1.292) with vector $\partial \boldsymbol{r}/\partial C_i$, then, to subtract one equation from another. Accounting for definition of the Lagrange brackets (1.237), one arrives to an algebraic equation for the transformation functions α_i,

$$\sum_{j=1}^{6} \left[C_i C_j\right] \alpha_j = \frac{\partial \boldsymbol{r}}{\partial C_i} \cdot \left(\boldsymbol{\Phi} - \tilde{\boldsymbol{\Phi}}\right) \,, \tag{1.293}$$

which can be solved with the help of the matrix of the Poisson brackets (1.238). Indeed, after performing the matrix multiplication of equation (1.293) with the Poisson brackets and accounting for their property of orthogonality with the Lagrange brackets (1.239), yield

$$\alpha_i = \sum_{j=1}^{6} \left\{C_i C_j\right\} \frac{\partial \boldsymbol{r}}{\partial C_j} \cdot \left(\tilde{\boldsymbol{\Phi}} - \boldsymbol{\Phi}\right) \,. \tag{1.294}$$

This equation substituted in equation (1.286) allows us to calculate the correspondence between one set of the orbital elements, C_i, associated with the gauge function $\boldsymbol{\Phi}$, and another set of the elements, $\tilde{C}_i$, associated with the choice of another gauge function, $\tilde{\boldsymbol{\Phi}}$. For a fixed gauge function $\boldsymbol{\Phi}$, there is a residual gauge freedom of transformations of the orbital elements C_i given by the smooth functions α_i which has no explicit dependence on time, that is $\partial \alpha_i/\partial t = 0$. The residual gauge freedom is limited by the class of functions α_i that satisfy equation (1.289). Right side of equations (1.294) is functionally similar to the right side of equations (1.253) for the osculating elements. Hence, expanding the difference $\tilde{\boldsymbol{\Phi}} - \boldsymbol{\Phi}$ in the radial, transversal and normal to-the-orbit components one can write the right side of equation (1.294) in the form being similar with the right side of the Euler-Gauss equations (1.254).

1.5.4
The Gauge-Invariant Perturbation Equation in Terms of the Disturbing Force

Let us assume that one has picked up a particular function $\boldsymbol{\Phi} = \boldsymbol{\Phi}[C_1(t), ..., C_6(t), t]$. Then, the perturbed problem of motion is reduced to a system of two vector differential equations for six variables $C_i(t)$:

$$\sum_{i=j}^{6} \frac{\partial \boldsymbol{v}}{\partial C_j} \dot{C}_j = -\dot{\boldsymbol{\Phi}} + \boldsymbol{F} \,, \tag{1.295a}$$

$$\sum_{i=j}^{6} \frac{\partial \boldsymbol{r}}{\partial C_j} \dot{C}_j = \boldsymbol{\Phi} \,. \tag{1.295b}$$

Now, take the dot product of the first equation with $\partial \boldsymbol{r}/\partial C_i$, and the dot product of the second equation with $\partial \boldsymbol{v}/\partial C_i$. The difference between these two products will amount to

$$\sum_{j=1}^{6} \left[C_i C_j\right] \dot{C}_j = \left(\boldsymbol{F} - \dot{\boldsymbol{\Phi}}\right) \cdot \frac{\partial \boldsymbol{r}}{\partial C_i} - \boldsymbol{\Phi} \cdot \frac{\partial \boldsymbol{v}}{\partial C_i} \quad , \tag{1.296}$$

where the left side contains the Lagrange brackets defined in equation (1.237). It is worth emphasizing that the Lagrange brackets are defined in a gauge-invariant, i.e., $\boldsymbol{\Phi}$-independent fashion. Indeed, the dependence on $\boldsymbol{\Phi}$ could appear, if and only if, the brackets contained time derivatives from the variables $C_i(t)$. However, neither the function, $\boldsymbol{r}$, nor the function $\boldsymbol{v} = \partial \boldsymbol{r}(C_i, t)/\partial t$ include differentiation of parameters C_i with respect to time. The Poisson brackets defined in equation (1.238) are gauge-invariant for the same reason. Equation (1.296) implements the gauge-invariant generalization of the planetary equations (1.252) in the Euler-Gauss form.

Be mindful that $\boldsymbol{\Phi}$ is set to be a single-valued function $\boldsymbol{\Phi}(C_i, t)$ of the time t and parameters $C_i = C_i(t)$, but not of their time-derivatives, $\dot{C}_i$. In principle, the gauge functions with dependence upon the parameters' time derivatives of all orders are also conceivable, especially in the post-Newtonian celestial mechanics of binary pulsars [Damour, 1983; Grishchuk and Kopeikin, 1986; Lorimer and Kramer, 2004] and coalescing binary stars [Pati and Will, 2000]. Such gauge functions generate second and higher-order derivatives in the system of equations for parameters C_i [Damour and Schäfer, 1985; Grishchuk and Kopeikin, 1986] which solution is a highly non-trivial mathematical endeavor [Chicone *et al.*, 2001].

The full time derivative of the chosen $\boldsymbol{\Phi} = \boldsymbol{\Phi}(C_i, t)$ contains the time derivatives of the parameters C_i,

$$\dot{\boldsymbol{\Phi}} = \frac{\partial \boldsymbol{\Phi}}{\partial t} + \sum_{j=1}^{6} \frac{\partial \boldsymbol{\Phi}}{\partial C_j} \frac{dC_j}{dt} \ . \tag{1.297}$$

It will then be reasonable to move these derivatives to the left side of equation (1.296), thus, recasting the equation into

$$\sum_{j=1}^{6} \left([C_i C_j] + \frac{\partial \boldsymbol{r}}{\partial C_i} \cdot \frac{\partial \boldsymbol{\Phi}}{\partial C_j} \right) \frac{dC_j}{dt} = \frac{\partial \boldsymbol{r}}{\partial C_i} \cdot \boldsymbol{F} - \frac{\partial \boldsymbol{r}}{\partial C_i} \cdot \frac{\partial \boldsymbol{\Phi}}{\partial t} - \frac{\partial \boldsymbol{v}}{\partial C_i} \cdot \boldsymbol{\Phi} \ , \tag{1.298}$$

which is the general form of the gauge-invariant perturbation equation [Efroimsky and Goldreich, 2003, 2004]. If the Lagrange gauge, $\boldsymbol{\Phi} = 0$, is imposed, equation (1.298) naturally coincides with equation (1.252) which is equivalent to the Lagrange equation (1.264) when the perturbing force has a potential $R = R(\boldsymbol{r})$ depending only on positions, $\boldsymbol{r}$, but not on velocities, $\dot{\boldsymbol{r}}$, of the bodies.

Sometimes other gauges become advantageous for analytical calculations. In those gauges, the orbital elements, $C_i(t)$, are non-osculating with the instantaneous conics are not tangent to the actual orbit. A useful example of non-osculating elements is the set of *contact orbital elements* which is discussed below. Other settings wherein employment of nonosculating variables considerably simplifies calculations are the Gyldén-Meshcherskii problem, i.e., the orbital motion of a body of variable mass [Gurfil and Belyanin, 2008]; the Lense-Thirring effect, i.e, the relativistic motion of a satellite about a rotating mass [Ashby and Allison, 2007; Chashchina *et al.*, 2009; Ciufolini, 1986]; evolution of relative orbits of spacecrafts under perturbations [Gurfil, 2007]. An important example of such forces appears in the equations of motion of the post-Newtonian celestial mechanics, a topic to be discussed at length in section 6.3 below especially in conjunction with the different parameterizations of the relativistic two-body problem.

1.5.5 The Gauge-Invariant Perturbation Equation in Terms of the Disturbing Function

Let us assume that the perturbed dynamics of the reduced two-body problem can be described by the Lagrangian

$$L(\boldsymbol{r}, \dot{\boldsymbol{r}}, t) = \frac{\dot{\boldsymbol{r}}^2}{2} + \frac{GM}{r} + \Delta L(\boldsymbol{r}, \dot{\boldsymbol{r}}, t) , \tag{1.299}$$

where the first two terms in the right side defines the unperturbed Lagrangian and the perturbation, $\Delta L = \Delta L(\boldsymbol{r}, \dot{\boldsymbol{r}}, t)$, depends on both position, $\boldsymbol{r}$, and velocity, $\dot{\boldsymbol{r}}$, of the body, and on the time t. The linear momentum of the body is defined by

$$\boldsymbol{p} = \frac{\partial L}{\partial \dot{\boldsymbol{r}}} = \dot{\boldsymbol{r}} + \frac{\partial \Delta L}{\partial \dot{\boldsymbol{r}}} . \tag{1.300}$$

If the perturbation ΔL, is not singular, and one assumes that it is not, equation (1.300) can be solved, thus, obtaining velocity $\dot{\boldsymbol{r}}$ as a function of the momentum $\boldsymbol{p}$ and position $\boldsymbol{r}$,

$$\dot{\boldsymbol{r}} = \boldsymbol{p} - \frac{\partial \Delta L}{\partial \dot{\boldsymbol{r}}} , \tag{1.301}$$

where the second term in the right side is a function of $\boldsymbol{p}$ and $\boldsymbol{r}$. One can derive the Hamiltonian function, H, by making use of the Legendre transformation of the Lagrangian supplemented by equation (1.301),

$$H(\boldsymbol{r}, \boldsymbol{p}, t) = \boldsymbol{p} \cdot \dot{\boldsymbol{r}} - L . \tag{1.302}$$

Straightforward calculation reveals that the Hamiltonian corresponding to the Lagrangian (1.299) is

$$H = \frac{\boldsymbol{p}^2}{2} - \frac{GM}{r} + \Delta H , \tag{1.303}$$

where the perturbation

$$\Delta H \equiv -\Delta L - \frac{1}{2}\left(\frac{\partial \Delta L}{\partial \dot{\boldsymbol{r}}}\right)^2 . \tag{1.304}$$

The disturbing function R, being often used in celestial mechanics, is defined as the Lagrangian's perturbation,

$$R(\boldsymbol{r}, \dot{\boldsymbol{r}}, t) \equiv \Delta L(\boldsymbol{r}, \dot{\boldsymbol{r}}, t) \ . \tag{1.305}$$

while the Hamiltonian's perturbation is denoted as

$$V(\boldsymbol{r}, \boldsymbol{p}, t) \equiv \Delta H(\boldsymbol{r}, \boldsymbol{p}, t) \ . \tag{1.306}$$

In virtue of equation (1.304), the interconnection of the Hamiltonian's perturbation with the disturbing function is written as

$$V = -R - \frac{1}{2}\left(\frac{\partial R}{\partial \dot{\boldsymbol{r}}}\right)^2 \ . \tag{1.307}$$

In many situations the disturbance bears no dependence upon the velocity $\dot{\boldsymbol{r}}$, so that the disturbing function in these cases coincides with the negative Hamiltonian's perturbation as the second term in (1.307) becomes nil. One assumes more general case and intend to address disturbances with a velocity-dependence present. Hence, the necessity to use the full formula (1.307).

The Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\boldsymbol{r}}} - \frac{\partial L}{\partial \boldsymbol{r}} = 0 \ , \tag{1.308}$$

written for the perturbed Lagrangian (1.299) are,

$$\ddot{\boldsymbol{r}} + \frac{GM}{r^3}\boldsymbol{r} = \boldsymbol{F} \ , \tag{1.309}$$

with the perturbing force given by

$$\boldsymbol{F} = \frac{\partial R}{\partial \boldsymbol{r}} - \frac{d}{dt}\left(\frac{\partial R}{\partial \dot{\boldsymbol{r}}}\right) \ , \tag{1.310}$$

which should be substituted in the right side of equation (1.298). One notices that the perturbation R is an implicit function of the orbital elements,

$$R = R(\boldsymbol{r}, \dot{\boldsymbol{r}}, t) = R\left[\boldsymbol{r}(C_i, t), \dot{\boldsymbol{r}}(C_i, t), t\right] \ , \tag{1.311}$$

where $C_i = C_i(t)$. Hence, the partial derivative

$$\frac{\partial R}{\partial C_i} = \frac{\partial R}{\partial \boldsymbol{r}}\frac{\partial \boldsymbol{r}}{\partial C_i} + \frac{\partial R}{\partial \dot{\boldsymbol{r}}}\frac{\partial \boldsymbol{v}}{\partial C_i} + \frac{\partial R}{\partial \dot{\boldsymbol{r}}}\frac{\partial \boldsymbol{\Phi}}{\partial C_i} \ . \tag{1.312}$$

The time derivative entering equation (1.310), can be written as

$$\frac{d}{dt}\frac{\partial R}{\partial \dot{\boldsymbol{r}}} = \frac{\partial}{\partial t}\left(\frac{\partial R}{\partial \dot{\boldsymbol{r}}}\right) + \sum_{j=1}^{6}\frac{\partial}{\partial C_j}\left(\frac{\partial R}{\partial \dot{\boldsymbol{r}}}\right)\frac{dC_j}{dt} \ . \tag{1.313}$$

Inserting the expression (1.310) for the force $\boldsymbol{F}$ in the generic equation (1.298), and making use of the expressions (1.312) and (1.313), one arrives at the general form of the gauge-invariant equation of orbital evolution for an arbitrarily-chosen set of orbital elements C_i [Efroimsky, 2005b; Efroimsky and Goldreich, 2003, 2004]:

$$\sum_{j=1}^{6}\left\{[C_i C_j] + \frac{\partial \boldsymbol{r}}{\partial C_i}\cdot\frac{\partial \boldsymbol{\Psi}}{\partial C_j}\right\}\frac{dC_j}{dt} = -\frac{\partial V}{\partial C_i} - \left(\frac{\partial \boldsymbol{v}}{\partial C_i} + \frac{\partial R}{\partial \dot{\boldsymbol{r}}}\frac{\partial}{\partial C_i} + \frac{\partial \boldsymbol{r}}{\partial C_i}\frac{\partial}{\partial t}\right)\cdot\boldsymbol{\Psi}\,, \tag{1.314}$$

where one has used definition (1.307) of the Hamilton's perturbation V, and introduced a new notation

$$\boldsymbol{\Psi} \equiv \boldsymbol{\Phi} + \frac{\partial R}{\partial \dot{\boldsymbol{r}}}\,, \tag{1.315}$$

for the gauge function shifted from its original value, $\boldsymbol{\Phi}$, by the partial derivative of the perturbing function, R, with respect to velocity of the body. The gauge function $\boldsymbol{\Psi}$ is arbitrary with the only limitation that comes from the decision to keep $\boldsymbol{\Phi}$ as a function of the time and the orbital elements but not of their time derivatives.

1.5.6
The Delaunay Equations without the Lagrange Constraint

As an example, let us consider the generic perturbation equations (1.314) for the Delaunay elements, defined in section 1.4.7. We permit the perturbation to depend both on the position and the velocity of the orbiting body but do not impose the condition of osculation, so the gauge function $\boldsymbol{\Psi}$ remains arbitrary. The system of the generic Delaunay equations read Efroimsky and Goldreich [2003]:

$$\frac{dL}{dt} = -\frac{\partial V}{\partial l} - \frac{\partial R}{\partial \dot{\boldsymbol{r}}}\cdot\frac{\partial \boldsymbol{\Psi}}{\partial l} - \frac{\partial \boldsymbol{r}}{\partial l}\cdot\frac{d\boldsymbol{\Psi}}{dt} - \frac{\partial \boldsymbol{v}}{\partial l}\cdot\boldsymbol{\Psi}\,, \tag{1.316a}$$

$$\frac{dl}{dt} = \frac{\partial V}{\partial L} + \frac{\partial R}{\partial \dot{\boldsymbol{r}}}\cdot\frac{\partial \boldsymbol{\Psi}}{\partial L} + \frac{\partial \boldsymbol{r}}{\partial L}\cdot\frac{d\boldsymbol{\Psi}}{dt} + \frac{\partial \boldsymbol{v}}{\partial L}\cdot\boldsymbol{\Psi}\,, \tag{1.316b}$$

$$\frac{dG}{dt} = -\frac{\partial V}{\partial g} - \frac{\partial R}{\partial \dot{\boldsymbol{r}}}\cdot\frac{\partial \boldsymbol{\Psi}}{\partial g} - \frac{\partial \boldsymbol{r}}{\partial g}\cdot\frac{d\boldsymbol{\Psi}}{dt} - \frac{\partial \boldsymbol{v}}{\partial g}\cdot\boldsymbol{\Psi}\,, \tag{1.316c}$$

$$\frac{dg}{dt} = \frac{\partial V}{\partial G} + \frac{\partial R}{\partial \dot{\boldsymbol{r}}}\cdot\frac{\partial \boldsymbol{\Psi}}{\partial G} + \frac{\partial \boldsymbol{r}}{\partial G}\cdot\frac{d\boldsymbol{\Psi}}{dt} + \frac{\partial \boldsymbol{v}}{\partial G}\cdot\boldsymbol{\Psi}\,, \tag{1.316d}$$

$$\frac{dH}{dt} = -\frac{\partial V}{\partial h} - \frac{\partial R}{\partial \dot{\boldsymbol{r}}}\cdot\frac{\partial \boldsymbol{\Psi}}{\partial h} - \frac{\partial \boldsymbol{r}}{\partial h}\cdot\frac{d\boldsymbol{\Psi}}{dt} - \frac{\partial \boldsymbol{v}}{\partial h}\cdot\boldsymbol{\Psi}\,, \tag{1.316e}$$

$$\frac{dh}{dt} = \frac{\partial V}{\partial H} + \frac{\partial R}{\partial \dot{\boldsymbol{r}}}\cdot\frac{\partial \boldsymbol{\Psi}}{\partial H} + \frac{\partial \boldsymbol{r}}{\partial H}\cdot\frac{d\boldsymbol{\Psi}}{dt} + \frac{\partial \boldsymbol{v}}{\partial H}\cdot\boldsymbol{\Psi}\,. \tag{1.316f}$$

In case of a disturbance depending only on the position, one has, $V = -R$. In this situation, it would be most convenient to fix the gauge in the Lagrange-constrain form: $\boldsymbol{\Phi} = 0$. This sets the Delaunay elements osculating, while the equations (1.316) become the well-known canonical equations of Delaunay (1.266). When the disturbance depends also upon the velocity, equations (1.316) still can be reduced to the canonical form (1.266) by choosing another gauge, $\boldsymbol{\Psi} = 0$, though this can be done only at the cost of osculation loss. Indeed, while after imposing such a gauge the equations (1.316) will look similar to equations (1.266) but the instantaneous conics parameterized by the Delaunay elements will be nontangent to the perturbed orbit. This situation is depicted in Figure 1.6.

Our example reveals that a blithe use of the Delaunay elements in problems with velocity-dependent perturbations $R = R(\boldsymbol{r}, \dot{\boldsymbol{r}})$, may lead to erroneous geometric interpretation of the orbital motion as the loss of osculation may be not noticed. Another moral of the story is that often the Delaunay elements are considered in the framework of the Hamilton-Jacobi theory of canonical transformations which treats these elements as canonical variables obeying the Hamilton equations. However, equations (1.316) are not necessarily Hamiltonian depending upon the gauge chosen. Hence, the canonicity of the Delaunay elements should not be taken for granted without checking upon which gauge conditions have been imposed. For example, the nonosculating orbital inclination can differ in the first order from the osculating inclination of the orbit that can be important for correct interpretation of the theory of Iapetus's inclination evolution developed by Ward [1981].

The main conclusion is: whenever one encounters a disturbance that depends not only upon positions but also upon velocities or momenta, implementation of the canonical-perturbation method necessarily yields equations that render nonosculating canonical elements. It is possible to keep the elements osculating, but only at the cost of sacrificing canonicity. For example, under velocity-dependent orbital perturbations (like inertial forces, or atmospheric drag, or relativistic correction) the equations for osculating Delaunay elements ($\boldsymbol{\Phi} = 0$ constraint is imposed) will no longer be Hamiltonian [Efroimsky, 2002a,b].

For the first time, non-osculating orbital variables were encountered probably by Poincaré in his studies of the three-body problem, though he never explored these variables from the viewpoint of a non-Lagrange constraint choice. Having performed a transition from the barycentric to the heliocentric reference frame Poincaré [1896]; Poincare [1897] noticed a subtle difference between the instantaneous conics parameterized by the canonical Delaunay variables defined in the two frames. A conic parameterized by the Delaunay elements in the barycentric frame deviated from the perturbed trajectory at the rate of t^2 of the time t. At the same time, a conic parameterized by the Delaunay elements in the heliocentric frame deviated from the perturbed trajectory at the linear rate t. As became evident later the variables in the barycentric frame were osculating, while the heliocentric variables were not. Evolution of the heliocentric set of variables was governed by a Hamiltonian perturbation that turned out to be velocity-dependent, which is natural because the heliocentric frame is non-inertial with the Hamiltonian perturbation depending on linear momenta of the bodies. In the language of symplectic geometry, Poincaré's finding, including the issue of choosing either osculating or non-osculating elements in the three-body problem, was briefly addressed in Abdullah and Albouy [2001].

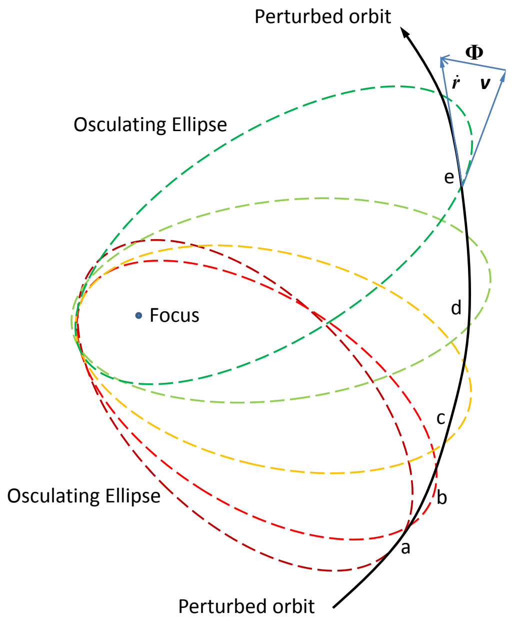

Figure 1.6 This picture illustrates the method of variation of parameters in the reduced two-body problem without imposing the condition of osculation. A perturbed orbit is a set of points, each of which is donated by a representative of a sequence of confocal instantaneous conics a, b, c, d, e, that are *not* supposed to be tangent, nor even coplanar to the orbit. As a result, the physical velocity $\dot{\boldsymbol{r}} = d\boldsymbol{r}/dt$ that is tangent to the perturbed orbit, differs from the Keplerian velocity $\boldsymbol{v}$ that is tangent to the conic. The sequence of nonosculating conics is characterized by vector $\boldsymbol{\Phi}(\tilde{C}_1, ..., \tilde{C}_6, t) = \dot{\boldsymbol{r}}(\tilde{C}_1, ..., \tilde{C}_6, t) - \boldsymbol{v}(\tilde{C}_1, ..., \tilde{C}_6, t)$ expressed as a function of time and six (nonosculating) orbital elements.

1.5.7 Contact Orbital Elements

The generic equations (1.314) evidently reveal the convenience of the constraint

$$\boldsymbol{\Psi} = 0 \quad \Longleftrightarrow \quad \boldsymbol{\Phi} = -\frac{\partial R}{\partial \dot{\boldsymbol{r}}} \,. \tag{1.317}$$

It cancels many terms in equations (1.314), reducing them to

$$\sum_{j=1}^{6} [\tilde{C}_i \tilde{C}_j] \frac{d\tilde{C}_j}{dt} = -\frac{\partial V}{\partial \tilde{C}_i} \,, \tag{1.318}$$

where one has denoted, $\tilde{C}_i$, the orbital elements corresponding to the constraint (1.317) in order to distinguish them from the osculating elements elements C_i. The so-defined orbital elements $\tilde{C}_i$ are called the *contact elements*. They are often used for analysis of orbits perturbed by velocity-dependent forces [Efroimsky, 2005a].

The term "contact elements" was offered in celestial mechanics by Brumberg *et al.* [1971]. Later, Kinoshita [1993] employed these variables. At that time, though, it was not yet clear that such variables obey conditions (1.317). It can be proven that variables obeying the same conditions show up also when one tries to preserve the interrelation $|\boldsymbol{J}| \equiv |\boldsymbol{r}\times\boldsymbol{p}|$ in the a frame precessing at a rate $\boldsymbol{\mu}$ where the momentum per unit mass, $\boldsymbol{p} = \dot{\boldsymbol{r}} + \boldsymbol{\mu}\times\boldsymbol{r}$ is not equal to $\dot{\boldsymbol{r}}$ making this situation similar to the case of the velocity-dependent perturbations of orbital motion Goldreich [1965]. Calculations carried out in terms of these variables are often greatly simplified. At the same time, one should be aware that the instantaneous conics parameterized by these variables are not tangent to the actual trajectory. Brumberg *et al.* [1971]; Kinoshita [1993], and Goldreich [1965] employed the contact variables to describe motion of a satellite orbiting a precessing oblate massive body. Although the instantaneous values of the contact variables differ from their osculating counterparts already in the first order, their averages differ only in the second order, *provided the motion is periodic*. However, in other situations, the absence of a periodic precession can invalidate geometric interpretation of the averaged values of these elements already in the first order.

Derivatives of the contact elements can be decoupled from the Lagrange brackets in equations (1.318) with the help of the matrix orthogonality condition (1.239). Thus, one obtains a system of ordinary differential equations for the contact elements,

$$\frac{d\tilde{C}_i}{dt} = \sum_{j=1}^{6} \left\{\tilde{C}_i \tilde{C}_j\right\} \frac{\partial V}{\partial \tilde{C}_j} \,, \tag{1.319}$$

which looks similar to the Lagrange planetary equations (1.264) except that now the perturbation, V, is given by more complicated expression (1.307) that involves the partial derivative of the disturbing function, R, with respect to velocity of the body, and the elements $\tilde{C}_i$ obey the constraint (1.317), instead of the Lagrange constraint,

$\boldsymbol{\Phi} = 0$. More specifically, for the set of the contact elements $(\tilde{a}, \tilde{e}, \tilde{\imath}, \tilde{\Omega}, \tilde{\omega}, \tilde{\mathcal{M}}_0)$ one gets [Brumberg, 1972, 1991]:

$$\frac{d\tilde{a}}{dt} = \frac{2}{\tilde{n}\tilde{a}}\frac{\partial V}{\partial \tilde{\mathcal{M}}_0}\,, \tag{1.320a}$$

$$\frac{d\tilde{e}}{dt} = \frac{1-\tilde{e}^2}{\tilde{n}\tilde{a}^2\tilde{e}}\frac{\partial V}{\partial \tilde{\mathcal{M}}_0} - \frac{\sqrt{1-\tilde{e}^2}}{\tilde{n}\tilde{a}^2\tilde{e}}\frac{\partial V}{\partial \tilde{\omega}}\,, \tag{1.320b}$$

$$\frac{d\tilde{\imath}}{dt} = \frac{\cos\tilde{\imath}}{\tilde{n}\tilde{a}^2\sqrt{1-\tilde{e}^2}\sin\tilde{\imath}}\frac{\partial V}{\partial \tilde{\omega}} - \frac{1}{\tilde{n}\tilde{a}^2\sqrt{1-\tilde{e}^2}\sin\tilde{\imath}}\frac{\partial V}{\partial \tilde{\Omega}}\,, \tag{1.320c}$$

$$\frac{d\tilde{\Omega}}{dt} = \frac{1}{\tilde{n}\tilde{a}^2\sqrt{1-\tilde{e}^2}\sin\tilde{\imath}}\frac{\partial V}{\partial \tilde{\imath}}\,, \tag{1.320d}$$

$$\frac{d\tilde{\omega}}{dt} = -\frac{\cos\tilde{\imath}}{\tilde{n}\tilde{a}^2\sqrt{1-\tilde{e}^2}\sin\tilde{\imath}}\frac{\partial V}{\partial \tilde{\imath}} + \frac{\sqrt{1-\tilde{e}^2}}{\tilde{n}\tilde{a}^2\tilde{e}}\frac{\partial V}{\partial \tilde{e}}\,, \tag{1.320e}$$

$$\frac{d\tilde{\mathcal{M}}_0}{dt} = -\frac{1-\tilde{e}^2}{\tilde{n}\tilde{a}^2\tilde{e}}\frac{\partial V}{\partial \tilde{e}} - \frac{2}{\tilde{n}\tilde{a}}\frac{\partial V}{\partial \tilde{a}}\,. \tag{1.320f}$$

As the right side of the resulting equation (1.319) contains only the Hamiltonian variation V, it may be logical to christen the constraint (1.317) the *Hamiltonian gauge*. Insertion of this gauge in the expression (1.281) for the perturbed velocity makes this velocity read,

$$\dot{\boldsymbol{r}} = \boldsymbol{v} - \frac{\partial R}{\partial \dot{\boldsymbol{r}}}\,. \tag{1.321}$$

Comparing this with equations (1.300) and (1.305), one sees that in the Hamiltonian gauge the partial time derivative of the perturbed coordinates, $\boldsymbol{v} = \partial \boldsymbol{r}/\partial t$, is equal to the canonical momentum,

$$\boldsymbol{v}\left[\tilde{C}_1(t), ..., \tilde{C}_6(t)\,, t\right] = \boldsymbol{p}\left[\tilde{C}_1(t), ..., \tilde{C}_6(t)\,, t\right]\,. \tag{1.322}$$

Equation (1.322) allows us to interchange the velocities and the corresponding momenta in the expressions for the Lagrange and Poisson brackets, whenever one is working in the Hamiltonian gauge. It also tells that the contact elements $\tilde{C}_i$ represent an osculating instantaneous orbit in the phase space $(\boldsymbol{r}, \boldsymbol{p})$ in contrast to the canonical osculating elements in the Lagrange gauge, which represent an osculating conic in the configuration space $(\boldsymbol{r}, \dot{\boldsymbol{r}})$. Relationship between the osculating elements C_i in the configuration space and the contact elements $\tilde{C}_i$ in the phase space can be found from equation (1.294). Indeed, substituting the Lagrange constraint, $\boldsymbol{\Phi} = 0$, and the Hamiltonian gauge, $\tilde{\boldsymbol{\Phi}} = -\partial R/\partial \dot{\boldsymbol{r}}$, one obtains

$$\tilde{C}_i - C_i = -\sum_{j=1}^{6}\left\{C_i C_j\right\}\frac{\partial \boldsymbol{r}}{\partial C_j}\cdot\frac{\partial R}{\partial \dot{\boldsymbol{r}}}\,. \tag{1.323}$$

The right side of this equation looks the same as that (1.253) with the "force" $\boldsymbol{F} = \tilde{\boldsymbol{\Phi}} = -\partial R/\partial \dot{\boldsymbol{r}}$. Therefore, the differences between the elements, $(\tilde{a}, \tilde{e}, \tilde{\imath}, \tilde{\Omega}, \tilde{\omega}, \tilde{\mathcal{M}}_0)$ and $(a, e, i, \Omega, \omega, \mathcal{M}_0)$, are given by the right side of equations (1.254).

It is possible to prove that the Hamiltonian gauge condition (1.317) is compulsory imposed by the canonical perturbation theory of the Hamiltonian equations for any dynamic system. Indeed, let us assume that the orbital elements $\tilde{C}_1, ..., \tilde{C}_6$ are associated with generalized coordinates $\boldsymbol{Q} = (Q_i) = (\tilde{C}_1, \tilde{C}_2, \tilde{C}_3)$ and generalized momentum $\boldsymbol{P} = (P_i) = (\tilde{C}_4, \tilde{C}_5, \tilde{C}_6)$, which obey the Hamiltonian equations

$$\frac{dQ_i}{dt} = \frac{\partial \tilde{H}}{\partial P_i} , \qquad \frac{dP_i}{dt} = -\frac{\partial \tilde{H}}{\partial Q_i} , \tag{1.324}$$

where the perturbed Hamiltonian

$$\tilde{H} = \tilde{H}(\boldsymbol{Q}, \boldsymbol{P}, t) = V(\boldsymbol{r}, \boldsymbol{p}, t) + \partial\chi/\partial t , \tag{1.325}$$

and $\chi = \chi(\boldsymbol{r}, \boldsymbol{p}, t)$ is a generating function defining the canonical transformation from $(\boldsymbol{r}, \boldsymbol{p})$ to $(\boldsymbol{Q}, \boldsymbol{P})$ [Arnold, 1995; Landau and Lifshitz, 1969] so that radius-vector $(r^i) = \boldsymbol{r}(\boldsymbol{Q}, \boldsymbol{P}, t)$ and momentum $(p^i) = \boldsymbol{p}(\boldsymbol{Q}, \boldsymbol{P}, t)$.

The total time derivative of $\boldsymbol{r}$ is expressed as follows

$$\dot{\boldsymbol{r}} = \frac{\partial \boldsymbol{r}}{\partial t} + \frac{\partial \boldsymbol{r}}{\partial Q_i}\frac{dQ_i}{dt} + \frac{\partial \boldsymbol{r}}{\partial P_i}\frac{dP_i}{dt} = \frac{\partial \boldsymbol{r}}{\partial t} + \left\{\boldsymbol{r}\tilde{H}\right\} , \tag{1.326}$$

where

$$\left\{\boldsymbol{r}\tilde{H}\right\} = \frac{\partial \boldsymbol{r}}{\partial Q_i}\frac{\partial \tilde{H}}{\partial P_i} - \frac{\partial \boldsymbol{r}}{\partial P_i}\frac{\partial \tilde{H}}{\partial Q_i} , \tag{1.327}$$

is the Poisson brackets of $\boldsymbol{r}$ and $\tilde{H}$ expressed in terms of the canonical variables (Q_i, P_i), and the Einstein summation rule is used for the repeated indices. However, due to the invariance of the Poisson brackets with respect to the canonical transformations [Landau and Lifshitz, 1969] one has

$$\left\{\boldsymbol{r}\tilde{H}\right\} = \{\boldsymbol{r}V\} = \frac{\partial \boldsymbol{r}}{\partial r^i}\frac{\partial V}{\partial p^i} - \frac{\partial \boldsymbol{r}}{\partial p^i}\frac{\partial V}{\partial r^i} = \frac{\partial V}{\partial \boldsymbol{p}} . \tag{1.328}$$

However, the partial derivative $\partial V/\partial \boldsymbol{p} = \partial R/\partial \dot{\boldsymbol{r}}$ [Landau and Lifshitz, 1969, §40]. Hence, going back to equation (1.326) it reveals that any system of canonical variables leads to equation (1.321) which implies the Hamiltonian constraint (1.317) telling us that the variables $(\boldsymbol{Q}, \boldsymbol{P})$ must be interpreted as the contact elements $\tilde{C}_1, ..., \tilde{C}_6$. This gauge-stiffness [24] property of the system of the Hamiltonian equations should be taken into account in the theory of the gauge transformations of the parameter space.

1.5.8
Osculation and nonosculation in rotational dynamics

Interestingly, the phenomenon of osculation vs nonosculation emerges not only in the theory of orbits but also in rotational dynamics, provided the method of variation of parameters is employed [Efroimsky and Escapa, 2007]. This should not

24) This term was suggested by Peter Goldreich.

be surprising, because the mathematics underlying rotational mechanics is virtually identical to that underlying orbital mechanics. In orbital mechanics, a perturbed trajectory of a body consists of points, each of which is donated by a representative of a sequence of instantaneous Keplerian conics. If one now disembodies this idea of its particular implementation, one should agree that:

(a) a trajectory may be assembled of points contributed by a family of algebraic curves of an essentially arbitrary type, not necessarily conics;
(b) it is not obligatory to set the family of curves tangent to the perturbed trajectory. In fact, it is often beneficial to choose them nontangent.

In its generality, the approach can be applied, for example, to describe the time evolution of Euler's angles characterizing orientation of a rotating body with respect to inertial space. A disturbed rotation can be thought of as consisting of a series of small turns along different Eulerian cones each of which is an orbit on the Euler angles' manifold corresponding to an unperturbed state of the angular momentum (spin) of the body. Just as in orbital mechanics, a transition from one instantaneous Keplerian conic to another is caused by a disturbing force, so a transition from one instantaneous Eulerian cone to another is governed by either an external torque, or the torque due to precession of the frame, or other perturbations like re-distribution of matter within the rotating body. Thus, in rotational mechanics, the Eulerian cones play the same role as the Keplerian conics do in the orbital dynamics. Most importantly, a perturbed rotation may be parameterized by the elements of the Eulerian cones in an osculating or in a nonosculating manner that is picked up by imposing a constraint on the rotational elements that is similar to choosing the gauge function, $\mathbf{\Phi}$, in the orbital dynamics. In many cases, the osculating Eulerian cones are convenient but the nonosculating parametrization may sometimes be more beneficial.

When the equations for the rotational elements are required to be canonical, the so-called Andoyer variables are typically chosen. However, like in case of the Delaunay orbital elements, the Andoyer variables may share the important peculiarity: under certain circumstances, the standard Hamiltonian equations of rotational motion render the elements nonosculating. In the theory of orbits, the standard form of the Lagrange and Delaunay planetary equations yield contact elements when perturbations depend on velocities. To keep the elements osculating in the configuration space, the equations must be amended with extra terms that are partial derivatives of the disturbing function with respect to velocities. It complicates the planetary equations and makes, for example, the Delaunay equations noncanonical. In rotational dynamics, whenever a perturbation depends upon the angular velocity, the canonical approach makes the Andoyer variables entering the Hamiltonian equations to be nonosculating to the Eulerian cones. To make them osculating, extra terms should be added to the standard Hamiltonian equations but then the equations will no longer be canonical [Efroimsky and Escapa, 2007].

1.6 Epilogue to the Chapter

As any physical theory, the Newtonian mechanics has a restricted realm of applicability. The first example of this realm's limitations, the problem of Mercury's apsidal precession, was encountered by astronomers back in the second part of the XIX-th century. By then, it had long been known that within the unperturbed Keplerian two-body problem the Laplace-Runge-Lenz vector is preserved and is always pointing toward the pericenter. Hence, if one neglects the planets' mutual disturbances, the periapses of their orbits would be staying idle. The disturbances however make the periapses move. The effect is especially pronounced in the case of Mercury, which has a small mass and therefore is most sensitive to the pull of the other planets. At the same time, since Mercury has an orbit of a high eccentricity and a small period, the advance of its pericenter is fairly easy observable, and it was accurately measured back in the XIX-th century. However, the rate of this advance turned out to differ from the predictions of the classical planetary theory by about 43 arcseconds per century. To explain the discrepancy, astronomers had to wait until the theory of general relativity was created. Explanation of Mercury's anomalous apsidal precession then became one of the first triumphs of Einstein's theory.

Over the XX-th century, astronomers came across many other examples of celestial motion, for whose accurate description in terms of the Newtonian mechanics turned out to be insufficient. An incomplete list includes the orbital motion of neutron stars and accretion-disc particles in binary systems; the motion of artificial satellites, the Moon, asteroids, and inner planets in the Solar System; and propagation of light through gravitational field. Recently, several so-far-unexplained anomalies in the orbital motion of spacecraft, planets, and the Moon have been registered [Anderson and Nieto, 2010]. They may indicate that even more subtle relativistic effects in the orbital motion of the bodies should be taken into account. Still, the problem of Mercury's pericenter advance marks the starting point whence the science of relativistic celestial mechanics reckons its history.

2
Introduction to Special Relativity

2.1
From Newtonian Mechanics to Special Relativity

2.1.1
The Newtonian Space-Time

Let us consider a physical process taking place within a sufficiently small volume of space and a sufficiently short period of time. One assumes that it is possible to reduce the actual physical process to its limiting content by shrinking the volume and the time interval to zero. This brings to life an idealized notion of event, an often used abstraction of a physical phenomenon taking place at one point in space and at one instant of time. Examples of the events are the photon's emission or absorbtion, collision of two elementary particles, meteorite's burst in planetary atmosphere, etc. . The event is mathematically identified with a point and the physical world is considered as consisting of a continuous set of points modeling its underlying mathematical structure. The set of events is further equipped with additional mathematical paraphernalia which converts it to a well-explored mathematical object called *manifold*.

Referring the reader to Chapter 3 for more extended treatment, one defines an n-dimensional manifold as a continuous set of points which can be covered by a finite or countable collection of overlapping n-dimensional coordinate charts. This means that each point of the manifold belongs to two or more charts, for each of which there exists a one-to-one continuous map to an open domain of n-dimensional coordinate space $\mathbb{R}^n$ that is represented by a set of n-tuples of real numbers. In principle, one has to define the terms *open* and *continuous* and accept that points of the manifold are separable from each other that constitutes the important Hausdorff property. Nonetheless, for now, this intuitive concept of the manifold is sufficient. A rigorous definition (in terms of topological spaces, etc.) can be found, for example, in Arnold [1995]; de Felice and Clarke [1990] and other mathematical textbooks.

In the Newtonian physics, the spacetime is modeled with a four-dimensional manifold $\mathbb{R}^4$. This space has a special structure selected by the mapping: $\mathbb{R}^4 \rightarrow \mathbb{R}^1$, with the one-dimensional subspace $\mathbb{R}^1$ called the absolute time. It is parameterized with a continuous parameter t, which measures time intervals between events in the sense of the Euclidean space. In other words, if one event occurs at time t_a and another event does at time t_b, the time interval between the two events is given by the absolute value $|t_b - t_a|$. The events are called simultaneous if they have the same numerical value of the parameter t. The set of events simultaneous with a given event form a three-dimensional subspace, $\mathbb{R}^3$, of simultaneous events called the absolute space. The absolute time flows from past to future at the same rate at each point of the absolute space. It is assumed that the absolute time can be measured with an ideal clock. The rate of the clock is independent from the motion of observer. Two travelers can meet to synchronize their clocks, then can move apart, and meet again later – just to find that the synchronization remains intact. This selected role of the absolute time tells us that the Newtonian spacetime is built as a direct product of $\mathbb{R}^3 \times \mathbb{R}^1$ that is a fiber-bundle structure with $\mathbb{R}^1$ being the base and copies of $\mathbb{R}^3$ playing the role of bundles [Schutz, 1995]. It is unimportant in classical physics whether one interprets the time as a fourth dimension of the world or as just a parameter t that labels a sequence of snapshots of the three-dimensional absolute space in the overall evolution of the world. The absolute time has a natural direction to future at any point of the absolute space. It is the existence of the special direction of time which makes the world look geometrically as a stack of three-dimensional sections impaled on the time axis, each section being marked with a value of t the same for all points of the section (see Figure 2.1). To be more specific, in the so-assembled four-dimensional space, a pair of initially synchronized clocks moving arbitrary with respect to each other, will continue to show the same time irrespectively of the speed of their relative motion. Calling the absolute time t fourth dimension brings no new physical aspect to the Newtonian spacetime.

The Newtonian spacetime is a vector space with vectors defined as differences between the coordinates of events in the absolute space. Let us consider two arbitrary events, a and b, that occur at times, t_a and t_b, respectively and have coordinates, $\boldsymbol{x}_a = (x^i_a) = (x^1_a, x^2_a, x^3_a)$ and $\boldsymbol{x}_b = (x^i_b) = (x^1_b, x^2_b, x^3_b)$. Currently, one does not assume any limitation on whether the events are simultaneous or not. The distance between the events is defined with the help of the Euclidean norm of a three dimensional vector, $\boldsymbol{x}_b - \boldsymbol{x}_a$,

$$d(\boldsymbol{x}_a, \boldsymbol{x}_b) = |\boldsymbol{x}_b - \boldsymbol{x}_a| = \sqrt{(x^1_b - x^1_a)^2 + (x^2_b - x^2_a)^2 + (x^3_b - x^3_a)^2} \ . \tag{2.1}$$

Coordinates, in which the distance is expressed by formula (2.1), are called Cartesian [1]. Other properties of the Euclidean vector space are applied as well. For example, the dot product, $\boldsymbol{a} \cdot \boldsymbol{b}$, between any two vectors, $\boldsymbol{a} = (a^i) = (a^1, a^2, a^3)$ and $\boldsymbol{b} = (b^i) = (b^1, b^2, b^3)$, is defined in the Cartesian coordinates as

$$\boldsymbol{a} \cdot \boldsymbol{b} = \delta_{ij} a^i b^j \ , \tag{2.2}$$

where the unit matrix $\delta_{ij} = \mathrm{diag}(1, 1, 1)$, and one reminds that the Einstein summation with respect to the repeated indices has been assumed. In terms of the dot-product the Euclidean distance (2.1) is expressed as

$$d^2(\boldsymbol{x}_a, \boldsymbol{x}_b) = (\boldsymbol{x}_b - \boldsymbol{x}_a) \cdot (\boldsymbol{x}_b - \boldsymbol{x}_a) \ , \tag{2.3}$$

and its always positive except when the spatial coordinates of two events coincide, in which case, $d(\boldsymbol{x}_a, \boldsymbol{x}_a) = 0$. The distance in the Newtonian physics can be formally defined between either simultaneous or non-simultaneous events. The reason why one needs to operate with the distance between non-simultaneous events comes from astrometry which studies the positions and motions of celestial bodies with respect to observer on the Earth with the help of light rays. As light propagates with finite speed [2], astrometrists must inevitably deal with the distances between the events of emission and observation of light, which are obviously not simultaneous in case of

1) Non-Cartesian coordinates can be also used in the absolute space but they do not provide a direct access to the Euclidean distance.
2) Classical physics does not postulate the speed of light as invariant but it does accept that it has a finite speed.

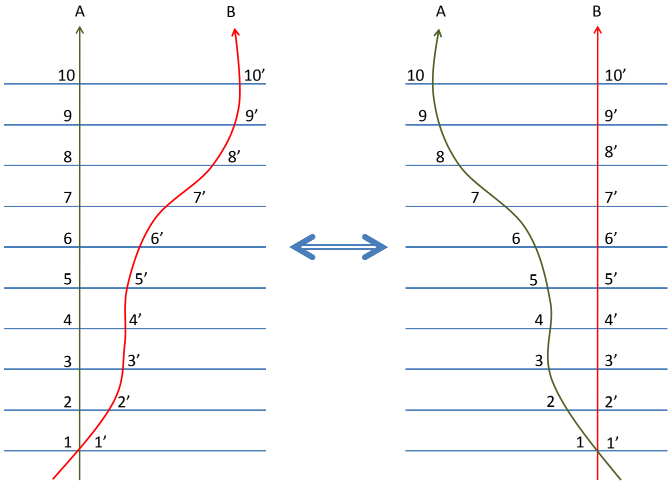

Figure 2.1 Mathematical structure of the world of the classical physics consists of the direct product of three-dimensional absolute space (shown as a stack of parallel, horizontal lines) and a one-dimensional absolute time. Depending on observer, the time axis may look like a straight or curved line connecting different events on the sequence of the space sections which represent an invariant geometric structure. Relative motion of observer A with respect to observer B is described by the reciprocal Newtonian transformation, which preserves the time difference and mutual connection of the events in the fiber-bundle assembly of the Newtonian spacetime.

observation of distant celestial objects. This is especially important for observations of the solar system bodies where the finite speed of light affects measured distances to planets, thus, changing their positions on the celestial sphere with respect to those predicted on the basis of the Newtonian mechanics – the effect known as the planetary aberration [Murray, 1983, §2.5.5].

2.1.2
The Newtonian Transformations

Classical mechanics describes the world of the Newtonian physics through vector laws acting in three-dimensional Euclidean space. There is a class of transformations between the Cartesian coordinates preserving the fiber-bundle structure of the Newtonian spacetime called the *Newtonian transformations*. Let us introduce two reference frames, S and S', with the time t and spatial Cartesian coordinates denoted as $\boldsymbol{r} = (x, y, z)$ and $\boldsymbol{r}' = (x', y', z')$ respectively. The Newtonian transformations consist of the following basic types (see Figure 2.2):

$$\text{time shifts}: \qquad t' = t + t_0\ , \qquad \boldsymbol{r}' = \boldsymbol{r}\ , \tag{2.4a}$$

$$\text{translations}: \qquad t' = t\ , \qquad \boldsymbol{r}' = \boldsymbol{r} + \boldsymbol{R}\ , \tag{2.4b}$$

$$\text{rotations}: \qquad t' = t\ , \qquad \boldsymbol{r}' = \mathcal{R} * \boldsymbol{r}\ , \tag{2.4c}$$

where t_0 is a constant time offset, $\boldsymbol{R} = \boldsymbol{R}(t)$ is an arbitrary vector function of time, $\mathcal{R} = \mathcal{R}(t)$ is an orthogonal matrix of rotation depending on time, and the sign $*$ denotes a matrix multiplication of a vector with matrix. It is straightforward to check that if one takes two simultaneous events a and b with times $t_a = t_b$, and coordinates $\boldsymbol{r}_a = (x_a, y_a, z_a), \boldsymbol{r}_b = (x_b, y_b, z_b)$ and $\boldsymbol{r}'_a = (x'_a, y'_a, z'_a), \boldsymbol{r}'_b = (x'_b, y'_b, z'_b)$ in the frames S and S' respectively, the Newtonian transformations do not change the Euclidean distance between the events, that is

$$d(\boldsymbol{x}'_a, \boldsymbol{x}'_b) = d(\boldsymbol{x}_a, \boldsymbol{x}_b)\ , \tag{2.5}$$

in both frames. The fiber-bundle structure of the Newtonian spacetime is also preserved under discrete transformations: spatial inversion, $\boldsymbol{r}' = -\boldsymbol{r}$, and time reversal, $t' = -t$. The Newtonian transformations leave invariant the structure of the Newtonian spacetime, but they preserve nor the distance between non-simultaneous events having different spatial coordinates, nor the form-invariance of the Newtonian equations of motion, thus, altering the laws of the Newtonian mechanics. This apparently leads to difficulties in the interpretation of forces and physical interactions between bodies especially under the Newtonian law of universal gravity. Therefore, it is important to find out a restricted class of the Newtonian transformations which keep the Newtonian laws invariant. These transformations are called the *Galilean transformations*.

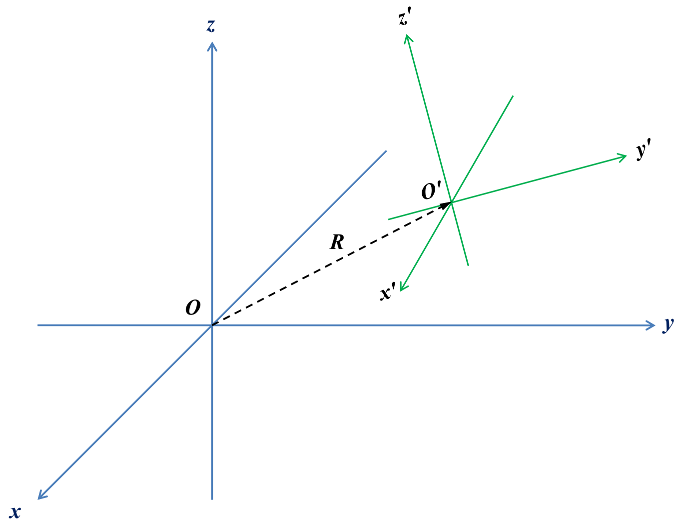

Figure 2.2 Cartesian coordinates $\boldsymbol{r} = (x, y, z)$ represent a reference frame S with the origin at point O. Cartesian coordinates $\boldsymbol{r}' = (x', y', z')$ represent a reference frame S' with the origin at point O'. Frame S' relates to frame S by the Newtonian transformations (2.4) that depend on time without further limitations.

2.1.3
The Galilean Transformations

Recall that, according to the first law of Newton, *there exist reference frames, called inertial, such that a particle at rest or with constant velocity in one inertial frame will remain at rest or have constant velocity in all inertial frames, provided the net force acting on the particle is nil.* This law may, at the same time, be regarded as a definition of inertial frames – the frames wherein a freely moving particle keeps its state of motion or rest without change. Denoting with $\boldsymbol{r}$ the spatial coordinates of the particle in an inertial frame S and with t the absolute time measured by an ideal clock in that frame, one can write the mathematical expression for the first law of Newton,

$$\frac{d^2\boldsymbol{r}}{dt^2} = 0 \,. \tag{2.6}$$

Although this expression may seem to ensue from Newton's second law with zero external forces, the first law of Newton should *not* be understood as merely a special case of the second law of Newton. The main burden of the first law is to *postulate* the existence of inertial frames of reference.

With this caveat in mind, let us explore what happens to the first law under the most general Newtonian transformation (2.4) which is written as

$$t' = t + t_0 \,, \qquad\qquad \boldsymbol{r}' = \mathcal{R} * \boldsymbol{r} + \boldsymbol{R} \,, \tag{2.7}$$

with vector $\boldsymbol{R} = \boldsymbol{R}(t)$ pointing from the origin O of frame S to the origin O' of frame S', and the orthogonal time-dependent matrix of rotation, $\mathcal{R} = \mathcal{R}(t)$, as shown in Figure 2.2. One notices that due to the absolute nature of the time, the time derivative $d/dt = d/dt'$. Differentiating the radius vector $\boldsymbol{r}'$ two times with respect to time t', and making use of equation (2.6), one arrives to the following equations

$$\boldsymbol{v}' = \boldsymbol{V} + \boldsymbol{\omega} \times (\boldsymbol{r}' - \boldsymbol{R}) + \mathcal{R} * \boldsymbol{v} \,, \tag{2.8}$$

$$\frac{d\boldsymbol{v}'}{dt'} = \dot{\boldsymbol{V}} + \dot{\boldsymbol{\omega}} \times (\boldsymbol{r}' - \boldsymbol{R}) + 2\boldsymbol{\omega} \times (\boldsymbol{v}' - \boldsymbol{V}) - \boldsymbol{\omega} \times [\boldsymbol{\omega} \times (\boldsymbol{r}' - \boldsymbol{R})] \,, \tag{2.9}$$

where $\boldsymbol{V} = d\boldsymbol{R}/dt$ is velocity of motion of the frame S' with respect to S, $\boldsymbol{v} = d\boldsymbol{r}/dt$ is velocity of motion of the particle with respect to the frame S, $\boldsymbol{v}' = d\boldsymbol{r}'/dt'$ is velocity of motion of the particle with respect to the frame S', and $\boldsymbol{\omega} = \dot{\mathcal{R}} * \mathcal{R}^{-1}$ is the angular velocity of rotation of spatial axes of the frame S' with respect to those of the frame S. The terms that appear in the right side of equation (2.9) are associated with the forces of inertia caused by the non-uniform relative motion of two frames and/or non-uniform relative rotation of their spatial axes.

The demand to keep the first law of Newton invariant with respect to the transformations between the frames imposes strong limitations on the vectors entering the Newtonian transformation: (1) the angular velocity of rotation, $\boldsymbol{\omega}$ must be equal to zero, thus, making the matrix of rotation $\mathcal{R}$ a constant orthogonal matrix, and (2) the velocity $\boldsymbol{V}$ of motion of one frame with respect to another must be constant making $\boldsymbol{R} = \boldsymbol{R}_0 + \boldsymbol{V}t$, where $\boldsymbol{R}_0$ is a constant vector of translation. These restrictions eliminate all forces of inertia from the Newtonian equations of motion and reduce the Newtonian transformations to a special sub-class called the Galilean transformations.

The set of all the Galilean transformations and their combinations constitute the Galilean group. This means the following:

(a) A sequence of any two transformations constitute another transformation of the same type.
(b) Among transformations between frames, there exists an identity transformation, the one that leaves everything unchanged.
(c) For any transformation, there exists an opposite one - such that the former followed or preceded by the latter yields the identical transformation.
(d) A sequence of any three consecutive transformations is associative.

The proof of the group's properties is pretty straightforward and involves the use of simple vector algebra [Arnold, 1995]. The Galilean group has ten primary transformations (generators): one shift of time t_0, three constant translations $\boldsymbol{R}_0$, three constant rotations that are the independent elements of the matrix $\mathcal{R}$, and three boosts which are the components of the constant velocity $\boldsymbol{V}$. The Cartesian frames connected between themselves by the Galilean transformations constitute the class of the inertial frames of reference. A standard example of the Galilean transformation is given by the frame S' moving with respect to S in a non-rotating manner so that the Cartesian axes x', y', z' of S' always stay parallel to the corresponding axes x, y, z of the frame S. Assuming the time offset $t_0 = 0$, the Galilean transformations are simplified to

$$t' = t\,, \qquad\qquad \boldsymbol{r}' = \boldsymbol{r} + \boldsymbol{R}_0 + \boldsymbol{V}t\,. \tag{2.10}$$

Form-invariance of the equations of motion of a freely-moving particle under the group of constant time offsets and spatial translations of coordinates shows that it is unimportant which point of the absolute space is taken as reference and which instant of time is taken as the initial epoch. It indicates that the absolute time and absolute space are homogeneous. Form-invariance of the equations of motion of a freely-moving particle under the group of constant rotations indicates that the absolute space is isotropic – there is no preferred direction in space. Neither Newtonian nor Galilean transformations change time which remains separated from space.

As an interesting historical aside, let us mention that the currently all-important concept of changing a reference frame had its origins in a work by Huygens [1669] who considered an elastic collision between two equally massive spheres with equal and opposite velocities, $\pm\boldsymbol{v}$. On the grounds of symmetry, Huygens pointed out that the spheres would recoil with their velocities reversed. He then imagined such a collision taking place on a boat that was itself moving at velocity $\boldsymbol{V}$ relative to the river bank. Were the collision watched by an observer standing on the bank, he would see it as a collision between spheres moving with velocities $\boldsymbol{V} - \boldsymbol{v}$ and $\boldsymbol{V} + \boldsymbol{v}$. The velocities, as seen from the bank, would be interchanged by the collision. Thus, on the basis of his consideration of the original symmetric collision, Huygens predicted the results of all collisions between these two spheres, occurring with an arbitrary relative initial velocity. Underlying Huygens' proof was another tacit assumption, the one that the mass of each object involved in a collision remained constant at any velocity.

One has to consider also the Newtonian equations of motion under interactions of particles with fields or with other particles, through fields. Therefore, the goal now is to explore if the form-invariance of the equations under the Galilean group can be extended to situations with physical interactions turned on. This check will require inspection of the transformation properties of forces included. Hence the outcome will depend upon these forces' nature and may be different for forces of different types. As the reader will see shortly, all the Galilean transformations do leave the Newtonian laws of mechanics form-invariant, provided the involved forces depend only on the differences between particle' positions and velocities. At the same time, transformations from the subgroup (2.10) of the Galilean group alter the Maxwell equations. This may make an impression that the equations of electrodynamics change when one switches from one inertial frame to another. Thus, while all the inertial frames look equally good insofar as only mechanical systems are considered, these frames may appear to be unequal in respect to electromagnetic phenomena – a dramatic paradox that once impelled Poincaré, Einstein, and other wonderful minds onto their quest for a new theory of space and time.

2.1.4
Form-Invariance of the Newtonian Equations of Motion

Let us consider a system of N interacting particles labeled by indices $A, B, C, \ldots$ running from 1 to N. The second Newton law relates acceleration, $\boldsymbol{a}_A = d^2\boldsymbol{r}_A/dt^2$, of particle A with the force, $\boldsymbol{F}_A$, exerted on the particle. Assuming that all forces of inertia are absent in the frame S, the second Newton law reads,

$$m_A \frac{d^2\boldsymbol{r}_A}{dt^2} = \boldsymbol{F}_A\left(\boldsymbol{r}_B, \boldsymbol{v}_B, t\right) , \qquad (A, B = 1, 2, \ldots, N) , \tag{2.11}$$

where m_A is mass of the particle, vectors $\boldsymbol{r}_A$ and $\boldsymbol{v}_A$ define position and velocity of the particle in the frame S, and one has assumed that the force explicitly depends on the position, velocity and the time t. If the initial conditions imposed on the position and velocity of the particle are known, its motion can be uniquely determined by solving the second-order differential equation (2.11). The result will be particle's trajectory given as a vector function of time $\boldsymbol{r}_A = \boldsymbol{r}_A(t)$. The acceleration of the particle given by the left side of the Newtonian equation of motion (2.11), is invariant with respect to the Galilean transformations. In order to keep the right side of this equation invariant, the Newtonian force $\boldsymbol{F}_A$ must obey certain conditions.

First of all, the force can not depend explicitly on time since the constant time offset t_0 must preserve the geometric image of particle's trajectory, that is if $\boldsymbol{r}_A(t)$ is solution of equation (2.11), the vector function $\boldsymbol{r}_A(t + t_0)$ must be also its solution. However, if the force depends on time, the time offset changes the force to $\boldsymbol{F}_A(\boldsymbol{r}_B, \boldsymbol{v}_B, t + t_0) \neq \boldsymbol{F}_A(\boldsymbol{r}_B, \boldsymbol{v}_B, t)$ without changing the left side of the differential equation (2.11). As a consequence, solution $\boldsymbol{r}_A(t+t_0)$ of this equation with the force $\boldsymbol{F}_A(\boldsymbol{r}_B, \boldsymbol{v}_B, t+t_0)$ will be different from the solution of the same equation with the force $\boldsymbol{F}_A(\boldsymbol{r}_B, \boldsymbol{v}_B, t)$ which contradict the principle of the Galilean invariance. Hence, one must admit that the force can not depend on the time explicitly if it obeys the Galilean principle of relativity. Not all forces satisfy this criteria. For example, dissipative forces can explicitly depend on time.

The Galilean invariance also suggests that if $\boldsymbol{r}_A(t)$ is solution of equation (2.11), the vector function $\boldsymbol{r}'_A(t) = \boldsymbol{r}_A(t) + \boldsymbol{R}_0 + \boldsymbol{V}t$ given by equation (2.10) must also satisfies the same equation. From this it follows that the Newtonian force in the inertial coordinate frame can depend only on the relative coordinates of the particles, $\boldsymbol{r}_B - \boldsymbol{r}_C$, and on their relative velocities, $\boldsymbol{v}_B - \boldsymbol{v}_C$, where coordinates and velocities of all interacting particles must be taken at the same time. Hence, the Galilean invariance also suggests that the Newtonian force of interaction is instantaneous. Thus, the most general form of the Galilean-invariant equations of motion of the interacting particles in the Newtonian physics are

$$m_A \frac{d^2\boldsymbol{r}_A}{dt^2} = \boldsymbol{F}_A\left(\boldsymbol{r}_B - \boldsymbol{r}_C, \boldsymbol{v}_B - \boldsymbol{v}_C\right) , \qquad (A, B, C = 1, 2, ..., N) . \tag{2.12}$$

In particular, the force of the universal gravitational interaction in the Newtonian physics has such Galilean-invariant structure

$$\boldsymbol{F}_A = -\sum_{B\neq A}^{N} \frac{Gm_A m_B}{|\boldsymbol{r}_A - \boldsymbol{r}_B|^3}\left(\boldsymbol{r}_A - \boldsymbol{r}_B\right) , \tag{2.13}$$

G being the universal gravitational constant.

A switch from one inertial frame S to another S' in accordance with the Galilean transformation (2.10) leaves the equations of motion (2.12) form-invariant. Indeed, after making the Galilean transformation (2.10), equation (2.12) acquires the following form

$$m_A \frac{d^2\boldsymbol{r}'_A}{dt'^2} = \boldsymbol{F}_A\left(\boldsymbol{r}'_B - \boldsymbol{r}'_C, \boldsymbol{v}'_B - \boldsymbol{v}'_C\right) , \qquad (A, B, C = 1, 2, ..., N) . \tag{2.14}$$

This equation has the same functional form as (2.12) except that it contains the coordinates and velocities of the particles measured in the new inertial frame S'. As pointed above, form-invariance of the equations of motion under the Galilean group tells us that all the inertial frames are equal, none of them being preferred in any way.

It is rather straightforward to prove that the Newtonian equations of motion are form-invariant under the noncommutative group of spatial rotations

$$t = t' , \qquad \boldsymbol{r}' = \mathcal{R} * \boldsymbol{r} , \tag{2.15}$$

$\mathcal{R}$ being an arbitrary, constant-in-time three-dimensional rotation matrix. Indeed, such a transformation does not change both the time derivative and the absolute value of the force, due to the orthogonality of the rotation matrix. Direction of the acceleration and the force will change by the same matrix of the rotation. Hence, transformation (2.15) does not affect the form of the equations of motion (2.12).

If the mechanical system of N interacting particles is closed, the Newtonian laws admit the time and/or space reversal [3] symmetry besides that with respect to the Galilean transformations. It means that the temporal evolution of the closed mechanical system does not have any naturally preferred direction. The concept of the arrow of time is applicable only to the systems which interact with their surroundings in time-irreversible manner with dissipative forces. This is one of the sources of the arrow of time emerging in classical physics. At the deepest level, though, the origin of the arrow of time should be traced to its quantum origins [Zeh, 2001].

2.1.5
The Maxwell Equations and the Lorentz Transformations

On the 8-th of December 1864, the Royal Society of London gathered for a session at which its recently elected Fellow, James Clerk Maxwell of King's College, read a presentation entitled "*A Dynamical Theory of the Electromagnetic Field*". The speaker summarized the results of Ampere, Faraday, and Gauss, and combined these with his own findings - mainly the displacement current [**?**] and the expression for the electromagnetic force - the one which is presently called the Lorentz force and which had in fact been pioneered by Maxwell [1862]. The talk, later published [Maxwell, 1865], became a major stride in the intellectual history of mankind.

Though being the intellectual breakthrough, the problem with Maxwell's work was its cumbersome mathematical form that gave the reader few chances to grasp the content of the theory. Among those who understood it, was a self-taught amateur explorer and inventor Oliver Heaviside, a chronically ill and chronically poor Londoner, who was living at his parents' place and who was destined soon to raise to prominence and comfort, and honors. In 1884, Heaviside gave Maxwell's equations a concise and elegant form, in which Maxwell himself never saw them. In current notations the Maxwell equations in vacuum are given by the set of two scalar and two vector equations [4]

$$\nabla \cdot \boldsymbol{E} = 4\pi\rho_e \,, \tag{2.16a}$$

$$\nabla \cdot \boldsymbol{B} = 0 \,, \tag{2.16b}$$

$$\nabla \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} \,, \tag{2.16c}$$

$$\nabla \times \boldsymbol{B} = \frac{4\pi}{c}\boldsymbol{j}_e + \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} \,. \tag{2.16d}$$

Here ∇ is the differential operator of spatial gradient, the symbols $\cdot$ and $\times$ between two vectors denote ordinary Euclidean dot and vector products, $\boldsymbol{E}$ and $\boldsymbol{B}$ are the

3) It does not matter whether all the three spatial axes are inverted or only one of these, because the resulting configurations are interconnected by a proper spatial rotation.

4) Electromagnetic field interacts with matter. It necessitates to distinguish in the media the strength of the field from its induction [Landau and Lifshitz, 1984]. In vacuum, the strength and induction of the field are identical.

electric and magnetic fields, ρ_e and $\boldsymbol{j}_e = \rho_e \boldsymbol{v}$ are the charge and electric current densities, $\boldsymbol{v}$ is the local velocity of motion of the charged matter distribution, and c is the universal fundamental constant defining the speed of propagation of light in vacuum [5]. By applying the operator of divergence, ∇ to both sides of equation (2.16d), with the subsequent use of equation (2.16a), one arrives at the conservation law of the electric charge,

$$\frac{\partial \rho_e}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{j}_e = 0 \,. \tag{2.17}$$

Maxwell's equations (2.16a)-(2.16d) taken along with the equation (2.17) of charge conservation form a system of eight differential equations in partial derivatives which is incomplete as the number of independent variables is ten: ρ_e, $\boldsymbol{v}$, $\boldsymbol{E}$, $\boldsymbol{B}$. To make the system complete one should introduce one more variable - the invariant density of matter $\rho^* = \rho\gamma$, where ρ is the density of matter in the comoving frame, and $\gamma = (1 - v^2/c^2)^{-1/2}$ is the Lorentz factor, and apply the relativistic equation of conservation and the equation of motion of matter

$$\frac{\partial \rho^*}{\partial t} + \boldsymbol{\nabla} \cdot (\rho^* \boldsymbol{v}) = 0 \,, \tag{2.18}$$

$$\frac{\partial (\rho^* \boldsymbol{v})}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla} (\rho^* \boldsymbol{v}) = \boldsymbol{F} \,. \tag{2.19}$$

Herein, the density of the relativistic force *per unit volume* ia

$$\boldsymbol{F} = \boldsymbol{F}_M + \boldsymbol{F}_L \,, \tag{2.20}$$

is an algebraic sum of the density of the net mechanical force $\boldsymbol{F}_M$, assumed to be known from the theory of elasticity, and

$$\boldsymbol{F}_L = \rho_e \left(\boldsymbol{E} + \frac{1}{c} \boldsymbol{v} \times \boldsymbol{B} \right) \,, \tag{2.21}$$

the volume density of the electromagnetic Lorentz force.

This increases the number of independent variables to eleven: : $\rho^*, \rho_e, \boldsymbol{v}, \boldsymbol{E}, \boldsymbol{B}$, and the number of equations goes to twelve. It looks now like the system of equations (2.16a)-(2.21) becomes over-determined. However, this is not the case because there is one degree of freedom of the Maxwell equations associated with the so-called gauge transformation of the electromagnetic field. Indeed, Maxwell's equations can be written in an alternative form, involving the electric (scalar) potential φ and magnetic (vector) potential $\boldsymbol{A} = (A^x, A^y, A^z)$. They appear as a result of the most general solution of the homogeneous Maxwell equations (2.16b) and (2.16c) written in the following form

$$\boldsymbol{E} = -\boldsymbol{\nabla}\varphi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} \,, \tag{2.22}$$

$$\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A} \,. \tag{2.23}$$

5) In fact, the universal speed defines the ultimate speed of propagation of any other fundamental field having no rest mass like gravitational field.

Simple inspection shows that after substituting these expressions back to equations (2.16b) and (2.16c), they are identically satisfied.

The gauge freedom of the Maxwell equations is associated with the freedom of transformation of the potentials

$$\boldsymbol{A}' = \boldsymbol{A} + \boldsymbol{\nabla}\chi\,, \qquad \varphi' = \varphi - \frac{1}{c}\frac{\partial\chi}{\partial t}\,, \tag{2.24}$$

where χ is an arbitrary scalar function. Electric $\boldsymbol{E}$ and magnetic $\boldsymbol{B}$ fields defined by equations (2.22), (2.23) do not depend on the choice of the gauge function χ which means that one is allowed to impose one constraint on the potentials φ, $\boldsymbol{A}$. This effectively reduces the number of independent equations (2.16a)-(2.21) from twelve to eleven, that is exactly equal to the number of independent variables. Hence, the equations admit a unique solution.

Different gauge conditions are used depending on the specific of the problem under consideration. Among them, the Lorentz gauge is the most famous, given by equation

$$\boldsymbol{\nabla}\cdot\boldsymbol{A} + \frac{1}{c}\frac{\partial\varphi}{\partial t} = 0\,. \tag{2.25}$$

Making use of the definitions (2.22), (2.23), and imposing the Lorentz gauge (2.25) reduces the Maxwell equations (2.16a) and (2.16d) to the wave equations for the potentials

$$\boldsymbol{\nabla}^2\varphi - \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = -4\pi\rho_e\,, \tag{2.26}$$

$$\boldsymbol{\nabla}^2\boldsymbol{A} - \frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} = -\frac{4\pi}{c}\boldsymbol{j}_e\,. \tag{2.27}$$

These equations admit solutions associated with electromagnetic waves propagating in vacuum with the fundamental speed c. The equation (2.25) does not fix the Lorentz gauge uniquely - the residual gauge freedom remains. It is easy to check by inspection that imposing the Lorentz gauge on both the potentials $\boldsymbol{A}$ and $\boldsymbol{A}'$ in the gauge transformation equation (2.24) leads to the homogeneous wave equation

$$\boldsymbol{\nabla}^2\chi - \frac{1}{c^2}\frac{\partial^2\chi}{\partial t^2} = 0\,, \tag{2.28}$$

imposed on function χ and that has a plenty of non-trivial solutions. Two restrictions (2.25) and (2.28) imposed on the four components of the scalar and vector potential leave only two of them independent. These two independent components are associated with the two polarization states of a freely propagating electromagnetic wave which is a solution of the wave equations (2.26), (2.27) in vacuum in the absence of matter.

An important "peculiarity" of Maxwell's equations, noticed very soon after they were published, was the lack of their form-invariance under the Galilean boosts (2.10). As it was the time when the existence of the æther and, therefore of an absolute motion with respect to a preferred frame, was considered a likely option, the lack of invariance under the Galilean boosts was not regarded as a flaw in a theory. This motivated a number of researchers to enquire if Maxwell's equations possess a more general internal symmetry extending the Galilean invariance of the second law of Newton. A decisive step in that direction was taken by Lorentz who found that Maxwell's equations in vacuum stay form-invariant when the spatial coordinates and time are transformed as

$$ct' = \gamma(ct - \beta x)\,, \quad x' = \gamma(x - c\beta t)\,, \quad y' = y\,, \quad z' = z\,, \tag{2.29}$$

and, simultaneously, the potentials are transformed as

$$\varphi' = \gamma(\varphi - \beta A_x)\,, \quad A'_x = \gamma(A_x - \beta\varphi)\,, \quad A'_y = A_y\,, \quad A'_z = A_z\,, \tag{2.30}$$

where the parameters of the transformation are

$$\beta \equiv \frac{V}{c}\,, \qquad \gamma \equiv \frac{1}{\sqrt{1-\beta^2}}\,, \tag{2.31}$$

with V being a quantity with dimensions of velocity. In the modern textbooks, this transformation is immediately followed by an explanation that V is a relative velocity of a frame S' parameterized with time t' and coordinates (x', y', z'), moving along x-axis with respect to frame S parameterized with time t and spatial coordinates (x, y, z). It should be remembered however that, having discovered this mathematical form-invariance, Lorentz did not interpret the new quantities t' and (x', y', z') as the time and Cartesian coordinates in a new frame of reference. Accordingly, he did not endow the parameter V with the meaning of a relative velocity of frames. Interpretation of Lorentz' transformations as switches between frames of reference was a major breakthrough, carried out independently by Poincaré and Einstein. Most importantly, they noticed that the Lorentz transformation of the space and time coordinates constitute a group of boost transformations generalizing the Galilean transformations (2.10).

Poincaré also corrected and extended Lorentz' mathematical results, by demonstrating that Maxwell's equations *with sources included*, too, stay form-invariant under transformations (2.29 - 2.30), provided the electric current and charge densities of the sources get simultaneously transformed as

$$c\rho'_e = \gamma(c\rho_e - \beta j_x)\,, \quad j'_x = \gamma(j_x - c\beta\rho_e)\,, \quad j'_y = j_y\,, \quad j'_z = j_z\,, \tag{2.32}$$

while the force density *per unit volume*, $\boldsymbol{F} = (F_x, F_y, F_z)$, and the instantaneous power $\boldsymbol{F}\cdot\boldsymbol{v}$, exerted by the force, get transformed via

$$\boldsymbol{F}'\cdot\boldsymbol{v}' = \gamma(\boldsymbol{F}\cdot\boldsymbol{v} - c\beta F_x)\,, \quad F'_x = \gamma\left(F_x - \frac{\beta}{c}\boldsymbol{F}\cdot\boldsymbol{v}\right)\,, \quad F'_y = F_y\,, \quad F'_z = F_z\,, \tag{2.33}$$

$\boldsymbol{v}$ and $\boldsymbol{v}'$ being the velocity of matter in the frames S and S', respectively. By associating Lorentz' formulae with transformations between frames, and by establishing their group property, Poincaré and Einstein gave the final shape to one of the toughest conundrums faced by the scholars of the early XX-th century: why are the laws of classic mechanics invariant under the Galilean group of transformations, while those of electrodynamics stay invariant under the Lorentz group?

2.2
Building the Special Relativity

2.2.1
Basic Requirements to a New Theory of Space and Time

Various experiments confirm that the Euclidean fiber-bundle construction of the spacetime manifold of the classical physics is inadequate to the realities of physical world, the mismatch becoming more noticeable the faster relative motions are reached. For example, the classical physics assumes that the life time of an unstable elementary particle is one and the same does not matter how fast the particle moves with respect to observer. However, experiments with elementary particles demonstrated elongation of time decay of rapidly moving particles, that is the higher the speed relative to the observer, the longer the observed lifetime of the particle. The necessity to explain this and other relativistic phenomena requires to build a new approach to the mathematical structure of spacetime, one that adequately describes physical phenomena measured by rapidly moving observers, though it includes the classical spacetime of events in the slow-motion limit.

Let us formulate the basic requirements to the new theory.

Requirement I. Relativistic spacetime is manifold $\mathbb{R}^4$ of events parameterized with one time and three spatial coordinates. Space $\mathbb{R}^3$ is homogeneous and isotropic at each point. Time is homogeneous.

Requirement II. Relativistic spacetime admits a special class of inertial reference frames, such that a test particle being at rest or moving with constant velocity in one inertial frame will remain at rest or have constant velocity in all inertial frames, provided the net force acting on the particle is nil.

Requirement III. Fundamental laws of physics have the same form in all inertial frames.

Requirement IV. There exist a fundamental constant c with dimension of speed which has the same numerical value in all inertial reference frames.

Requirements I and II are often taken for granted in the literature on special relativity as they are exactly the same as in the Newtonian mechanics. This is the reason why they are usually not referred to the postulates of special relativity. However, it is not self-evident that the new theory of spacetime must preserve these requirements. This is why one has spelled them out explicitly.

Requirements III is known as the first Einstein's postulate [Einstein, 1905] or the *special relativity principle*[Einstein, 1920b]. It extends the applicability of the relativistic transformations from spacetime coordinates to all independent variables entering the physical laws. For example, it demands to transform the electromagnetic potentials and the source variables in Maxwell's equations in accordance to equations (2.30), (2.32) besides the Lorentz transformation (2.29) of time and space coordinates.

Requirement IV is the modern formulation of the second Einstein's postulate: *light always propagates through vacuum at a universal invariant speed, c, which* *is independent of the state of motion of the emitting body.*. At the time, when Einstein published his paper on special relativity, the electromagnetic field was the only relativistic field known to theorists. Einstein, however, envisaged that its speed of propagation in vacuum has a fundamental meaning originating in the non-Euclidean structure of spacetime. In a letter written in 1955 to Seelig, Einstein summarized special relativity in the following words: “Its novelty was to formulate that the Lorentz transformations are of importance beyond the scope of the Maxwell equations, and concern the structure of the space and time. Another new point was the conclusion that the ‘Lorentz invariance’ is the general demand for any physical theory" [Seelig, 1956].

Other particles have been predicted that must move in vacuum with the same speed c as light. The best example is graviton - the particle of gravitational field which provides us with another physical realization of the invariant speed c [Kopeikin, 2005]. Requirement III establishes the mathematical form of equivalence between the time and space coordinates of the inertial frames, introduced in the requirement II. This equivalence is known under the name of the Lorentz transformation shown in equation (2.29) for a particular case of two frames moving one with respect to another along x axis. The Lorentz transformations put on equal mathematical footing the time and space coordinates and destroy the fiber-bundle structure of the Newtonian spacetime. Let us emphasize in the spirit of Einstein, that identification of the Lorentz transformation with a switch between inertial frames reveals a fundamental property of space and time, and not of a particular class of physical interactions. In other words, the time and space coordinates change via the Lorentz transformations in *all* laws of physics, not only in Maxwell’s equations. Therefore, the new mechanical equations of motion must be form-invariant under the Lorentz transformations, not the Galilean ones. Still, in the limit of slow motions these equations should acquire their customary Newtonian form.

It is possible to prove (and it is done later) that the fundamental speed c is the ultimate speed for any particle or field residing on the spacetime. The ultimate speed c is often identified with the speed of light c_l in vacuum. Indeed, according to Maxwell’s equations (2.16) all observers in inertial frames will measure in vacuum the same speed of light $c_l = c$, regardless of their state of motion. Nonetheless, the invariant speed c is the intrinsic property of the Minkowski spacetime alone, and as such should not be confused with the physical speed of light c_l, which is always less than c when light is propagating in a transparent substance. The fundamental speed c enters all relativistic equations of the fundamental interactions not only the Maxwell equations. Therefore, its physical meaning should be treated with taking into account the nature of the fundamental interaction [Ellis and Uzan, 2005; Kopeikin, 2004].

2.2.2
On the "Single-Postulate" Approach to Special Relativity

Some authors [Schröder, 1994; Torretti, 1983] suggest that the special relativity principle, that is Einstein’s first postulate, alone is sufficient to establish the invariance of all physical constants entering the laws of physics, including the speed of light.

The adjective "single-postulate", which is sometimes applied to such derivations is somewhat misleading, because the derivations also rely on other assumptions. Being mathematically more involved than Einstein's treatment, the "single-postulate" derivation of special relativity is omitted in most textbooks, though some authors mention this option.

An attempt to construct special relativity without using Einstein's second postulate explicitly, was first undertaken by von Ignatowski [1910], whose derivation was based on several assertions: (a) the principle of relativity is applied to electrodynamics, (b) the space is isotropic, (c) space and time and homogeneous, (d) the transformation between two inertial frames is reciprocal. Arguing that Ignatowski's proof was incomplete, Frank and Rothe [1911] came up with a more accurate derivation based on the assertions that (a) the Lorentz transformations form a group, (b) a switch between inertial frame entails only a change of the sign of the relative speed, (c) the length contraction depends only on the relative speed between the frames. Later, Pauli [1921] pointed out that the sufficiency of the "single-postulate" treatment offered by von Ignatowski [1910] and Frank and Rothe [1911] hinges on the recourse to electrodynamics. Within a treatment embracing Maxwell's theory, the second postulate is not, in fact, abandoned but replaced by the universality of the Lorentz transformations and, hence, by the presence of the invariant speed c in these transformations. Indeed the sole way of reconciling the relativity principle with electrodynamics is to say that the Lorentz transformations are universal and applicable to *all* forces, for which reason the speed of light is equated to a *universal* constant c that must show up in all fundamental equations of theoretical physics, not only in Maxwell's theory. However, within an axiomatic treatment not referring to Maxwell's equations, the second Einstein's postulate is needed.

The most optimal way of introducing the invariant speed c to the theory of spacetime was implemented by Minkowski [1908], who re-formulated special relativity in geometrical terms, by postulating the invariance of the interval, $ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$, between two infinitesimally close events under linear coordinate transformations as a fundamental property of spacetime. Minkowski's formulation entails the mathematical form of the Lorentz transformations and requires physical laws be invariant under these transformations. This formulation shed a new light on the fascinating discovery of Einstein – the existence in Nature of a fundamental constant c having the dimension of speed.

The discussion on the necessity of the second Einstein's postulate is still going on in the literature [Behera, 2002; Lee and Kalotas, 1975]. A short but very accurate historical study on the origin and role of the second postulate, with numerous references, can be found in Abiko [2003]. Review by Silagadze [2008] discusses other, deeply hidden theoretical aspects of the foundations of special relativity and its relation to non-Euclidean geometry of spaces with constant curvature.

2.2.3
The Difference in the Interpretation of Special Relativity by Einstein, Poincaré and Lorentz

Opening his original paper with the two postulates, Einstein [1905] employed them to derive the Lorentz transformations, and also demonstrated that the transformations constitute a group. This may sound like new variations of the old theme by Lorentz, who had pioneered the transformations, and Poincaré, who had formulated the relativity principle in his Saint Louis address [Poincaré, 1904], and proved the group property of the Lorentz transformations [Poincaré, 1905]. In reality, Einstein brought a crucially new element into the theory, as will be explained below.

Lorentz was convinced that the result of the Michelson-Morley experiment was negative because of "*a fortuitous compensation of opposing effects*" [Lorentz, 1909], and not because of the absence of an absolute frame as such. His opinion was shared by Poincaré, who thought that the inability of contemporary physics of that time to register æther was due to some coupling between the æther and matter. For this reason, Poincaré always drew a distinction between the "apparent" or "local" coordinates and phenomena seen by observers moving in the æther, and the "true" or "real" ones registered by observers at rest in the æther. Similarly, he argued that clocks being at rest in the æther show the "true time", while clocks moving in æther show the "apparent time". This viewpoint reserved the potential possibility for some, thitherto undiscovered, physical effect to single out a preferred frame out of all inertial frames – a possibility that has found no confirmation in the presently available experimental data on the Lorentz-violating mechanisms in particle physics and anisotropy of cosmic microwave background radiation (CMBR) [Coleman and Glashow, 1999; Ehlers and Lämmerzahl, 2006; Kostelecký and Mewes, 2007; Will, 1993].

Einstein [1905] broke with the traditions of classical physics by rejecting absolute motion as such, and by giving an equal status to the time and space coordinates measured by all inertial observers – so the notions of "true" and "apparent" time and coordinates became redundant. Within this, revolutionary thinking, the equivalence of inertial frames became a cornerstone property of the spacetime continuum, and not a concatenation of dynamical interactions allegedly preventing us from observing the absolute motion or æther. Similarly, the universality of the speed of light in vacuum became an inherent property of the spacetime, and not a combination of physical effects leading to coincidence in results of measuring one and the same c by different observers. Last, and by no means least, Einstein's treatment was extremely "economic", in that he managed to derive all the results from the two postulates (also assuming tacitly the isotropy of space and the homogeneity of spacetime) [Kobzarev, 1975].

Although, mathematically, the special relativity by Einstein and the special relativity by Poincaré were equivalent,[6] Einstein's views sharply diverged from those

6) A theory of light developed by Lorentz [1904] was, mathematically, close to but not equivalent to the treatment by Poincaré or Einstein. In particular, Lorentz' theory could not account exactly for the optical Doppler effect, nor for the stellar aberration [Miller, 1974, 1986].

of Poincaré and Lorentz, on both the relativity principle and the meaning of c. As the cleavage was interpretational, and as Poincaré and Lorentz had indisputably pioneered most of the mathematical apparatus of special relativity, one is faced with the old question as to who actually should be credited with the authorship of special relativity. Some scholars insist on Poincaré's contribution being greater than that of Einstein [Logunov, 1990, 2004; Zahar, 1989], though this opinion has not become mainstream. Whittaker [1953] went so far in his book as to credit only Poincaré and Lorentz for developing special relativity, attributing only little importance to Einstein's work. It seems, special relativity was to Whittaker just another mathematical theory, so he did not pay a due attention at the far-reaching physical consequences of Einstein's abolition of absolute motion. At the same time, a later study by Miller [1987] also revealed that Whittaker [1953, Chapter II] devoted to the history of special relativity was " fraught with substantial historical errors". Among those was Whittaker's mis-attribution to Poincaré [1900] of Einstein's, $E = mc^2$, formula. While Poincaré [1900] indeed hypothesized that the electromagnetic energy behaves as a fluid endowed with mass, the said formula was certainly derived by Einstein – it is equivalent to the antepenultimate equation in Einstein [1905]. An accurate rebuttal to Whittaker's eisegesis can be found in Gray [1995]; Holton [1973]; Miller [1981] and references therein.

Seeking a careful criterion to compare Einstein's, Poincaré's, and Lorentz' contributions to the theory of special relativity, one should approach the case within the context of the next-level theory, of which special relativity is a basic part, – general relativity. It describes the world as a curved manifold whose tangent spaces are special-relativistic spaces of events. Among the basic premises of the theory is the nonexistence of any preferred frame of reference – in general relativity all frames, inertial and noninertial, are fundamentally equivalent. From the hight of this edifice, Einstein's priority in rejecting absolute motion, as well as his priority in elevating the postulates to the level of the spacetime properties, appear to have an even greater importance than Poincaré's priority in setting out the relativity principle or in deriving the ensuing formulae. For this reason, acknowledging both Einstein and Poincaré (and, largely, Lorentz) as the creators of special relativity, Einstein is kept firmly as its first author – for the reason that his insights were critical for the development of a fundamentally-new theory of space and time. Other students of history may, of course, try to weigh Poincaré's and Einstein's inputs differently, so the case may never get settled.

2.2.4 From Einstein's Postulates to Minkowski's Space-Time of Events

Classical mechanics finds its mathematical implementation in the Newtonian space of events, the bundles of Euclidean three-dimensional world impaled on time axis as shown in Figure 2.1. The time is considered as a book-keeping parameter and is completely separated from space. Although time may be called as fourth dimension, this will gain us nothing, because in the so-assembled four-dimensional spacetime continuum the time flows everywhere at the same rate, no matter how fast the observer moves. Our goal now is to construct a relativistic ensemble of space and time in compliance with the requirements I - IV from the subsection 2.2.1. This new spacetime will be called Minkowski spacetime and denoted with $\mathcal{M}$. Dynamics of particles is expected to be described by some second-order differential equations generalizing the Newtonian equations of motion. Let us make a rough sketch of the new space of events.

2.2.4.1 ***Dimension of the Minkowski spacetime.***

The Minkowski spacetime combines space and time to a single, four-dimensional manifold (dim $\mathcal{M} = 4$), which is covered by coordinate charts x^α ($\alpha = 0, 1, 2, 3$). More advanced spacetime theories include additional dimensions, in most cases spatial Smolin [2006]. How many dimensions are needed to describe the underlying mathematical structure of the physical world is still an open question. For example, string theory predicts 10 or 26 dimensions, and M-theory introduces 11 dimensions - 10 spatial and 1 temporal. In fact, the existence of more than four dimensions can make a difference only at the sub-atomic level and goes beyond the scope of this book.

2.2.4.2 ***Homogeneity and isotropy of the Minkowski spacetime.***

The Minkowski spacetime $\mathcal{M}$ is homogeneous and isotropic at each point. It means all points are equivalent and there is no preferred direction in spacetime anywhere. Homogeneity means the spacetime has translational invariance and isotropy means it has rotational invariance. The homogeneity and isotropy of space lead to the conservation laws of linear momentum, angular momentum, and energy of an isolated physical system, in particular in N-body problem. In this sense the Minkowski spacetime manifold is similar to the Newtonian spacetime. However, the construction of the Minkowski spacetime differs significantly from the Newtonian spacetime of events which preserves spatial distances and time intervals independently. It rigidly fixes the fiber-bundle structure of the Newtonian spacetime that can be always decomposed into spatial coordinates plus a time coordinate, and the separation is absolute and valid for any observer. The drastic difference of the Minkowski spacetime is the presence of the invariant speed c which is not an element of the Newtonian spacetime. Of course, the speed of light appears already in the Newtonian physics but it has no the fundamental meaning that is given to it in the Minkowski spacetime. Existence of the invariant speed c prevents the absolute separation of time from space leading to the mathematical structure which is a pseudo-Euclidean four-dimensional world. This extends the isotropy of space to the spacetime allowing the spacetime rotations which are nothing else but the Lorentz transformations.

2.2.4.3 ***Coordinates and reference frames***

Coordinate charts are used in the spacetime manifold $\mathcal{M}$ to represent the events. Usually, Cartesian coordinates $x^\alpha = (ct, x, y, z)$ are used [7] though curvilinear coor-

7) Many textbooks assume the invariant speed $c = 1$ to simplify equations, so that time is measured in units of length. We do not use this convention.

dinates are also allowed. In the most general case a number of overlapping coordinate charts is needed to cover the entire manifold. A reference frame is identified with a Cartesian coordinate chart with a physical observer placed at its origin. Another reference frame may be identified by a second Cartesian coordinate chart with another observer at its origin. Two Cartesian coordinate charts always overlap in the Minkowski spacetime. The domain of intersection of the charts represents the region of spacetime in which both observers can measure physical quantities and coordinates of events, and compare results. The relation between the two sets of measurements is given by a non-singular coordinate transformation that allows to match numerically different but physically equivalent measurements of one and the same object (event).

If coordinates are not Cartesian, the reference frame of each observer is identified with four unit vectors $\boldsymbol{e}_\alpha$ $(\alpha = 0, 1, 2, 3)$ making up an orthonormal basis called *tetrad*. One of the vectors, $\boldsymbol{e}_0$, is associated with the direction of time and is called timelike. The remaining three vectors of the basis $\boldsymbol{e}_i$ (i=1,2,3), are orthogonal [8] to the timelike vector and reside in a three-dimensional physical space. They are called spacelike vectors. In case of Cartesian coordinates the unit vectors of the reference frame defines the direction of coordinates axes. When a vector pointing from observer toward another event is decomposed over the basis, this decomposition is interpreted as measurement of the event's coordinates made by the observer at rest in the origin of the frame associated with this particular basis.

A four-dimensional vector basis is interpreted as an inertial frame of reference if positional vector of a test particle measured by observer, remains constant or is a linear function of time under condition that the net force exerted on the particle is absent. This property is preserved by the Lorentz transformations. The Minkowski spacetime manifold can be covered with non-inertial frames as well. However, fitting non-inertial frames into the special theory of relativity theory requires more mathematical justifications, and will be explained later.

2.2.4.4 ***Spacetime interval***

In the Newtonian spacetime, the separation between two points is measured by the Euclidean distance between the spatial coordinates of the two points as defined in equation (2.1). This distance is always positive. It is also purely spatial and does not depend on velocity of a particular observer doing the measurement, if and only if, the two events are simultaneous. In Minkowski spacetime, the separation between two events is measured by the interval between the two events, which takes into account not only the spatial separation between the events, but also their temporal separation. The interval Δs between two events is defined in the Cartesian coordinates as

$$\Delta s^2 = -c^2\Delta t^2 + \Delta r^2 \,, \tag{2.34}$$

8) The concept of orthogonality is explained in section 2.3.2.

where c is the invariant speed, Δt and $\Delta r = \left(\Delta x^2 + \Delta y^2 + \Delta z^2\right)^{1/2}$ are differences between the time and space coordinates of the events [9]. Our convention for the choice of signature for Δs^2 corresponds to $(-1, +1, +1, +1)$, where -1 belongs to the time, and $+1$ does to the space. Some other books like Landau and Lifshitz [1975], reverse the signature of the interval. Our choice of signature agrees with the contemporary choice accepted in other standard textbooks on special relativity and gravitation, for example, [Hawking and Ellis, 1975; Misner *et al.*, 1973; Wald, 1984]. Remarkable property of the spacetime interval (2.34) in the Minkowski spacetime is that it remains invariant under Lorentz transformations which explains its privileged role in measurements. Equation (2.34) could be formally written in the Newtonian physics as well. However, the speed c has no property of invariance in classic physics, and for differently moving observers the interval Δs between the same two events will have different numerical values.

Spacetime intervals are separated into three essentially different types based on the sign of Δs^2. They are:

$\Delta s^2 < 0$ timelike interval,
$\Delta s^2 = 0$ null-like interval,
$\Delta s^2 > 0$ spacelike interval.

Sometimes the null-like interval is called light-like because the experiments proved that the speed of light in vacuum coincides with the invariant speed c with an enormous precision in the sense that the speed of light in vacuum depends neither on the choice of the reference frame nor on the motion of observer or the source of light. Nevertheless, saying "light-like interval" may make an erroneous impression that light has a privileged position among other fundamental fields also propagating in vacuum with the invariant speed c like gravitons. To avoid such a misleading impression, the interval $\Delta s^2 = 0$ is called null-like.

2.2.4.5 *The null cone*

In a null-like interval the spatial distance between two events is exactly balanced by the time difference between them multiplied with the fundamental speed, $\Delta r = \pm c\Delta t$. Events connected by the path of a particle that propagates with the invariant speed c [10] all have null-like separation. Given one event, all those events which follow at null-like interval with $\Delta r = +c\Delta t$ define a half of the null cone propagated to future, and all the events which preceded from a null-like interval with $\Delta r = -c\Delta t$ define a second half of the null cone propagated to past. The null cone is invariant hypersurface for each event of spacetime. Therefore, it splits the Minkowski spacetime in three distinct regions: absolute future, absolute past, and absolute remoteness as shown in Figure 2.3. Absolute future and past make the interior of the null cone while the absolute remoteness occupies its exterior part.

9) Notations like Δt^2 or Δr^2 are understood in relativity as $(\Delta t)^2$ and $(\Delta r)^2$ respectively but not as $\Delta(t)^2$ or $\Delta(r)^2$.
10) Photon, graviton, gluon, or any other massless particle.

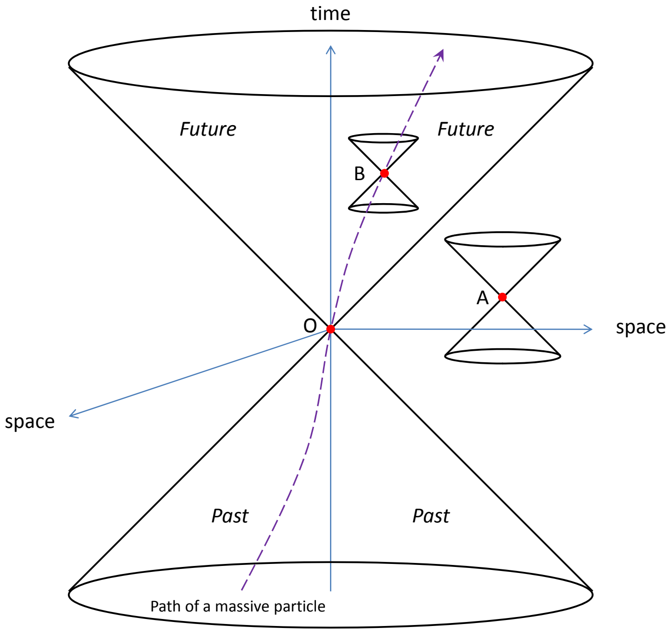

Figure 2.3 A future null cone of a single event O is the surface that a swarm of massless particles (photons, gravitons, etc.), emanating from the event O and traveling in all directions, would take through the Minkowski spacetime. A past null cone of the event O is the surface that is made of by paths of massless particles arriving to the event O from all directions in the Minkowski spacetime. Worldline of an arbitrary-moving massive particle is shown by the dashed line. It is always laying inside the null cone at each event on its worldline. Each event in the Minkowski spacetime has its own null cone.

2.2.4.6 *The proper time*

The measure of a timelike spacetime interval is described by the proper time

$$\Delta\tau = \left(\Delta t^2 - \frac{\Delta r^2}{c^2}\right)^{1/2} . \tag{2.35}$$

The proper time interval is measured by an observer with an ideal clock traveling between two infinitesimally close events in an inertial reference frame. In case of sufficiently separated events (for instance, O and B in Figure 2.3), the proper time is calculated by integration of equation (2.35) along the worldline of the clocks. The proper time defines a real number, since the interior of the square root in equation (2.35) is positive.

2.2.4.7 *The proper distance*

For pairs of event separated with a positive squared spacetime interval (for instance, O and A in Figure 2.3), the measurement of spacelike separation is called the proper distance

$$\Delta\sigma = \left(\Delta r^2 - c^2\Delta t^2\right)^{1/2} . \tag{2.36}$$

Like the proper time, the proper distance (2.36) is a real number.

2.2.4.8 *Causal relationship*

Causality is the relationship between a first event (the cause) and a second event (the effect), where the second event is a consequence of the first.

For two events, for example, O and B in Figure 2.3, separated by a timelike interval, enough time passes between them to establish a cause-effect relationship between them. For a particle traveling with a speed less than the invariant speed c, any two events on the worldline of the particle are separated by a timelike interval and occur in each other's future or past. There exists a reference frame such that the two events are observed to occur in the same spatial location, but there is no reference frame in which the two events can occur at the same time.

When two events are separated by a spacelike interval, for example, O and A in Figure 2.3, not enough time passes between their occurrences for establishing a causal relationship between them as it would require crossing the spatial distance between the two events at the speed faster than the invariant speed c, which is physically impossible. Such events are considered not to occur in each other's future or past. There exists a reference frame such that the two events are observed to occur at the same time, but there is no reference frame in which the two events can occur in the same spatial location. Besides the frames in which O precedes A (as shown in Figure 2.3) there are also frames in which A precedes O.

2.3
Minkowski Space-Time as a Pseudo-Euclidean Vector Space

2.3.1
Axioms of Vector Space

Let us designate the points of a four-dimensional manifold, $\mathcal{M}$, with upper-case Latin letters: A,B,C, etc. Each point has its own Cartesian coordinates x^α_A, x^α_B, x^α_C, etc. ($\alpha = 0, 1, 2, 3$) in an associated coordinate chart. The manifold itself has no any specific geometric structure and is not a vector space. The most simple way to endow $\mathcal{M}$ with a vector-space structure is to identify each ordered pair of points, say A and B, with a vector $\mathbf{a} = \overrightarrow{AB}$, having components, a^α, in the given coordinate chart that are defined as the differences between the coordinates of the corresponding points,

$$a^\alpha = x^\alpha_B - x^\alpha_A \; . \tag{2.37}$$

This vector can be visualized geometrically as an arrow connecting the two points and directed from point A to point B as shown in Figure 2.4. If A is any point and $\mathbf{a}$ any vector, there is one and only one point B for which $\overrightarrow{AB} = \mathbf{a}$. If C is any point not coinciding with A and B, then, one can always define a vector $\mathbf{c} = \overrightarrow{AC}$ such that it is given as an algebraic sum of two vectors,

$$\mathbf{c} = \mathbf{a} + \mathbf{b} \; , \tag{2.38}$$

where $\mathbf{b} = \overrightarrow{BC}$. Equations (2.37) and (2.38) establishes a one-to-one correspondence between points of the manifold and vectors. In what follows, vectors are always denoted by bold letters $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, etc.

The whole set of vectors defined on the manifold $\mathcal{M}$ obeys the law of addition and scalar multiplication that satisfy the set of axioms [Weyl, 1950, page 17]:

WA1 Commutativity of addition. For any two vectors $\mathbf{a}$ and $\mathbf{b}$,

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a} \; .$$

WA2 Associativity of addition. For any three vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$,

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}) \; .$$

WA3 Inverse element of addition. For any two vectors $\mathbf{a}$ and $\mathbf{b}$ there exists one and only one vector $\mathbf{x}$ such that

$$\mathbf{a} + \mathbf{x} = \mathbf{b} \; .$$

The vector $\mathbf{x} \equiv \mathbf{b} - \mathbf{a}$ is called the difference between vectors $\mathbf{a}$ and $\mathbf{b}$.

WA4 Identity element of addition. There exist a null vector $\mathbf{0}$, such that for any vector $\mathbf{a}$ one has

$$\mathbf{a} + \mathbf{0} = \mathbf{a}$$

.

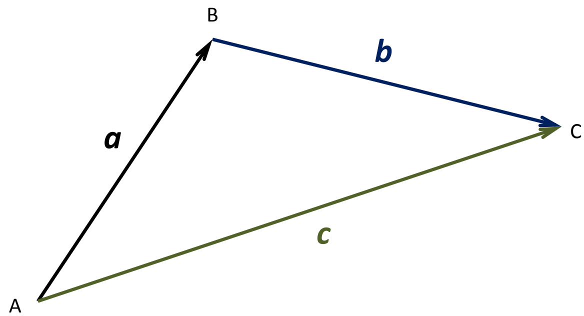

Figure 2.4 The identification of points A,B,C,... of a manifold with vectors. Vector $\mathbf{a} = \overrightarrow{AB}$, vector $\mathbf{b} = \overrightarrow{BC}$, and vector $\mathbf{c} = \overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC} = \mathbf{a} + \mathbf{b}$. Components of the vectors are defined as differences between Cartesian coordinates of the corresponding points: $a^\alpha = x^\alpha_B - x^\alpha_A$, $b^\alpha = x^\alpha_C - x^\alpha_B$, and $c^\alpha = x^\alpha_C - x^\alpha_A$.

WA5 Distributivity of scalar multiplication. For any real numbers α and β and any vector $\mathbf{a}$,

$$(\alpha + \beta)\mathbf{a} = \alpha\mathbf{a} + \beta\mathbf{a} \,.$$

WA6 Compatibility of scalar multiplication with vector multiplication. For any real numbers α and β and any vector $\mathbf{a}$,

$$\alpha(\beta\mathbf{a}) = (\alpha\beta)\mathbf{a} \,.$$

WA7 Identity element of scalar multiplication. For any vector $\mathbf{a}$ there exists a unit number 1 such that

$$1\mathbf{a} = \mathbf{a} \,.$$

WA8 Distributivity. For any real number α and two vectors $\mathbf{a}$ and $\mathbf{b}$,

$$\alpha(\mathbf{a} + \mathbf{b}) = \alpha\mathbf{a} + \alpha\mathbf{b} \,.$$

These axioms allow the ordered pairs of points on the manifold $\mathcal{M}$ be treated as vectors: one can add and subtract them, multiply them by numbers.

The structure of the vector space so defined, is local and can not be extrapolated on the entire manifold $\mathcal{M}$ in the most general case. However, if the manifold $\mathcal{M}$ is homogeneous and isotropic at each point it can be covered by a single coordinate chart. This is exactly the case of the Minkowski spacetime which is a global vector space having an affine structure which means that any pair of vectors in the Minkowski spacetime can be compared even if they are not attached to one and the same point. It establishes the relation of equality between a whole set of vectors being parallel to each other, so each vector $\mathbf{a}$ eventually becomes identified not with one ordered pair of events but with an equivalence class of pairs. The set of the axioms of the vector space may be slightly different from that given above but they are all logically interconnected so that it is not crucially important for physics. Historical aspects of the axioms of vector space and their significance in foundations of modern mathematics are discussed in a review article by Coleman and Korté [2001].

2.3.2
Dot-products and Norms

The comparison of vectors requires introducing one more operation between vectors in the vector space, which should be defined in addition to the axiom WA1-WA7. This operation maps a product of any two vectors to a real number [11)] and is called a dot-product. Mathematics tolerates a broad freedom in defining this operation because it carries a certain physical meaning related directly to calculation of the length (norm) of vectors and angles between them. In what follows, the dot-product between any two vectors, say $\mathbf{a}$ and $\mathbf{b}$, will be denoted with a dot symbol between them, $\mathbf{a}\cdot\mathbf{b}$. The properties of the dot-products in the Euclidean space and Minkowski spacetime are discussed below. For more mathematical details the reader is advised to consult the book by Penrose [2004, Chapter 18].

11) In the case of complex-valued manifolds, the dot-product can be an imaginary number. In this book one considers only real-valued manifolds.

2.3.2.1 *Euclidean space*

The Euclidean dot-product (also known as a scalar product) is the one obeying the following four axioms for any two vectors, $\boldsymbol{a}$ and $\boldsymbol{b}$, and real numbers, α and β:

EA1 The dot-product is commutative, $\boldsymbol{a}\cdot\boldsymbol{b} = \boldsymbol{b}\cdot\boldsymbol{a}$.
EA2 The dot-product is distributive, $(\boldsymbol{a}+\boldsymbol{b})\cdot\boldsymbol{c} = \boldsymbol{a}\cdot\boldsymbol{c} + \boldsymbol{b}\cdot\boldsymbol{c}$.
EA3 The dot-product is bi-linear, $(\alpha\boldsymbol{a})\cdot(\beta\boldsymbol{b}) = \alpha\beta(\boldsymbol{a}\cdot\boldsymbol{b})$.
EA4 The dot-product is positively defined $\boldsymbol{a}\cdot\boldsymbol{a} > 0$, if $\boldsymbol{a} \neq \boldsymbol{0}$, and $\boldsymbol{a}\cdot\boldsymbol{a} = 0$, if $\boldsymbol{a} = \boldsymbol{0}$.

In Newtonian mechanics, the spacetime manifold $\mathcal{M}$ can be covered with the direct product of time and space Cartesian coordinates. Vectors are defined as differences between the Cartesian coordinates of two points in the absolute space irrespectively whether they are simultaneous or not. The Euclidean dot-product between two vectors $\boldsymbol{a} = (a^i) = (a^1, a^2, a^3)$ and $\boldsymbol{b} = (b^i) = (b^1, b^2, b^3)$ $(i = 1, 2, 3)$ is

$$\boldsymbol{a}\cdot\boldsymbol{b} = a^1b^1 + a^2b^2 + a^3b^3 \ . \tag{2.39}$$

The time component of vectors is not involved to the definition of the Euclidean dot-product and, hence, does not matter.

The Euclidean norm of a vector $\boldsymbol{a}$ is denoted as $|\boldsymbol{a}|$ and is introduced as

$$|\boldsymbol{a}| \equiv (\boldsymbol{a}\cdot\boldsymbol{a})^{1/2} = \sqrt{a_1^2 + a_2^2 + a_3^2} \ , \tag{2.40}$$

that is a well-known mathematical expression. The angle θ between two vectors, $\boldsymbol{a}$ and $\boldsymbol{b}$, is defined as

$$\cos\theta = \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{|\boldsymbol{a}||\boldsymbol{b}|} \ . \tag{2.41}$$

Two vectors are called orthogonal in Euclidean space if $\cos\theta = 0$.

Directly from the axioms EA1-EA4, one can prove the following properties for any real number α and any two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$.

$$\begin{aligned}
\text{Homogeneity:}\quad & |\alpha\boldsymbol{a}| = |\alpha||\boldsymbol{a}| \ . \\
\text{Cauchy-Schwarz inequality:}\quad & \boldsymbol{a}\cdot\boldsymbol{b} \le |\boldsymbol{a}||\boldsymbol{b}| \ . \\
\text{Triangle inequality:}\quad & |\boldsymbol{a}+\boldsymbol{b}| \le |\boldsymbol{a}| + |\boldsymbol{b}| \ . \\
\text{Pythagorean theorem:}\quad & |\boldsymbol{a}|^2 + |\boldsymbol{b}|^2 = |\boldsymbol{a}+\boldsymbol{b}|^2 \ , \ \text{if } \boldsymbol{a}\cdot\boldsymbol{b} = 0 \ .
\end{aligned}$$

The Euclidean norm (2.40) is the common-sense distance that stays invariant under spatial rotations. If two vectors are formed by making use of spatial coordinates of two simultaneous events, the Euclidean norm is also invariant with respect to time offsets, space translations, and the Galilean transformations. This remains valid for any vector in the Newtonian vector space irrespectively of its physical nature - be it a vector of acceleration, force, velocity, and so on. However, this classical construction fails to describe the actual physics of ultra-relativistic particles and fields, which physically meaningful properties are invariant under the Lorentz transformations, not the Galilean ones. To ensure compatibility of relativistic physics with mathematics of a vector space requires new definition of the dot-product by sacrificing axiom EA4. This leads us to the Lorentzian (also called pseudo-Euclidean) dot-product and norm in the Minkowski spacetime.

2.3.2.2 *Pseudo-Euclidean space*

Real-valued dot-products of an arbitrary sign are inherent to pseudo-Euclidean spaces. They are called Lorentzian and their properties differ from the Euclidean ones by excluding the axiom EA4. Instead of this axiom one accepts a broader possibility admitting that the dot-product is indefinite. More specifically, the Lorentzian dot-product in the Minkowski spacetime is defined between two four-dimensional vectors, **a** and **b**, as a quadratic form

$$\mathbf{a} \cdot \mathbf{b} = -a^0 b^0 + a^1 b^1 + a^2 b^2 + a^3 b^3 \,, \tag{2.42}$$

where $a^\alpha = (a^0, a^1, a^2, a^3)$ and $b^\alpha = (b^0, b^1, b^2, b^3)$ are the Cartesian components of vectors **a** and **b** respectively. Traditionally, the components of vectors with index 0, that is a^0 and b^0, are associated with time axis, and the components with indices $1, 2, 3$, that is a^i and b^i, are associated with space. Time components of vectors are apparently involved to the definition of the Lorentzian dot-product as contrasted to the scalar product of two vectors in Euclidean space.

The Lorentzian norm of a four-vector **a** will be denoted as $||\mathbf{a}||$, and it is related to the Lorentzian dot-product by relationship

$$||\mathbf{a}||^2 = \mathbf{a} \cdot \mathbf{a} \,, \tag{2.43}$$

where the right side can be either positive, or negative, or nil. Since components a^α of each vector **a** are identified with the difference between coordinates of its origin and tip, according to equation (2.37), the value of $||\mathbf{a}||^2$ is closely associated with the concept of the Minkowski interval Δs introduced above in section 2.2.4.4. In fact, the quantity $||\mathbf{a}||^2$ for a vector **a** is equal to the interval

$$||\mathbf{a}||^2 = -\left(x^0_B - x^0_A\right)^2 + \left(x^1_B - x^1_A\right)^2 + \left(x^2_B - x^2_A\right)^2 + \left(x^3_B - x^3_A\right)^2 \,. \tag{2.44}$$

The numerical value of $||\mathbf{a}||^2$ depends on the type of the corresponding interval. There are three different cases which split the Minkowski spacetime in three Lorentz-invariant domains filled by timelike vectors **a** with $||\mathbf{a}||^2 < 0$; spacelike vectors **b** with $||\mathbf{b}||^2 > 0$, and null-like vectors **c** having $||\mathbf{c}|| = 0$. We emphasize that though $||\mathbf{c}||$ of a null-like vector **c** is zero, it does not mean, as it was in the Euclidean space, that the vector itself is zero, $\mathbf{c} \neq \mathbf{0}$. The timelike domain of the Minkowski spacetime is called the absolute future and absolute past , the null-like hypersurface is called the null cone, and the spacelike domain is called the absolute remoteness. The three domains are not sub-spaces of the Minkowski spacetime because the sum of any two vectors can be any kind: timelike, null-like or spacelike, independently of the type of the vectors involved.

The Lorentzian norm of a four vector **a** is defined by the following equations:

$$\begin{aligned} ||\mathbf{a}|| &= \sqrt{-\mathbf{a} \cdot \mathbf{a}} && \text{timelike vector} \,, \\ ||\mathbf{a}|| &= 0 && \text{null-like vector} \,, \\ ||\mathbf{a}|| &= \sqrt{\mathbf{a} \cdot \mathbf{a}} && \text{spacelike vector} \,. \end{aligned}$$

The Lorentzian norm keeps the property of homogeneity, $||\alpha\mathbf{a}|| = |\alpha|||\mathbf{a}||$ for any four-vector $\mathbf{a}$ and a real number α. However, all other properties are either violated or reversed. Let us consider, for example, the triangle inequality for two timelike vectors, $\mathbf{a}$ and $\mathbf{b}$ such that their sum is also a timelike vector $\mathbf{c} = \mathbf{a} + \mathbf{b}$, which norm $||\mathbf{c}||^2 = ||\mathbf{a}||^2 + ||\mathbf{b}||^2 - 2\mathbf{a} \cdot \mathbf{b}$. However, timelike vectors obey the Cauchy-Schwarz inequality: $\mathbf{a} \cdot \mathbf{b} \le ||\mathbf{a}||||\mathbf{b}||$, and, hence, $||\mathbf{c}||^2 \ge ||\mathbf{a}||^2 + ||\mathbf{b}||^2 + 2\mathbf{a} \cdot \mathbf{b} = (||\mathbf{a}|| + ||\mathbf{b}||)^2$. The triangle inequality for timelike vectors read

$$||\mathbf{a} + \mathbf{b}|| \ge ||\mathbf{a}|| + ||\mathbf{b}|| \,, \tag{2.45}$$

which is opposite to the triangle inequality in Euclidean space. The reason for this is that timelike vectors in the Minkowski world measure intervals of time. The clock, which traveled along a straight line (vector $\mathbf{c}$) from point A to B always shows larger time interval than the clock moving from A to B along a broken line given by vectors $\mathbf{a}$ and $\mathbf{b}$. This is the essence of the "twin paradox" of special relativity.

Two four-vectors, $\mathbf{a}$ and $\mathbf{b}$, are called *orthogonal* if their Lorentzian dot-product $\mathbf{a} \cdot \mathbf{b} = 0$. One notices that a timelike vector can be orthogonal only to a spacelike vector; a null-like vector can be orthogonal to another null like-vector or to a spacelike vector, and a spacelike vector can be orthogonal to any four-vector.

2.3.3
The vector basis

The Minkowski spacetime of events is a four-dimensional vector space acting transitively which means that the Minkowski spacetime is a homogenous and isotropic affine space all points of which are equivalent and there is no preferred direction. Hence, there always exist four linearly-independent vectors, $\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$, such that any other vector $\mathbf{a}$ can be represented as a linear combination

$$\mathbf{a} = a^0\boldsymbol{e}_0 + a^1\boldsymbol{e}_1 + a^2\boldsymbol{e}_2 + a^3\boldsymbol{e}_3 \,, \tag{2.46}$$

where four real numbers (a^0, a^1, a^2, a^3) are called the components of vector $\mathbf{a}$. For the space is affine, the vector decomposition (2.46) is valid at any point of the Minkowski spacetime.

In what follows, the Greek indices $\alpha, \beta, \gamma, ...$ are used to denote the components of four-dimensional vectors and to numerate the basis vectors. The Greek indices run through the values 0,1,2,3. Spatial components of vectors will be denoted with the Roman letters $i, j, k, ...$ which run through the values 1,2,3. The Roman indices will be also used to denote the spacelike basis vectors. Thus, the vector $\mathbf{a}$ from equation (2.46) has four components $a^\mu = (a^0, a^1, a^2, a^3)$ each of which is multiplied with a basis vector $\boldsymbol{e}_\mu = (\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$. We shall also use Einstein's summation rule assuming that the dummy repeated Greek indices imply summation from 0 to 3, and the dummy repeated Roman indices imply summation from 1 to 3. The repeated indices are called dummy indices since their notation can be chosen arbitrary. For example, equation (2.46) can be written down in several equivalent forms

$$\mathbf{a} = \sum_{\alpha=0}^{3} a^\alpha \boldsymbol{e}_\alpha \equiv a^\alpha \boldsymbol{e}_\alpha = a^\beta \boldsymbol{e}_\beta = a^\mu \boldsymbol{e}_\mu = \ldots , \tag{2.47}$$

that have an identical meaning. Notice that each vector $\boldsymbol{e}_\beta$ of the basis $\boldsymbol{e}_\beta$ can be decomposed with respect to this basis like any other vector $\mathbf{a}$. Coefficients of this decomposition are given by the unit matrix δ^α_β because of equality $\boldsymbol{e}_\beta = \delta^\alpha_\beta \boldsymbol{e}_\alpha$ where the index β numerates the basis vectors, and the index α numerates its components. The conclusion is that a basis vector $\boldsymbol{e}_\beta$ for each index β fixed, has components δ^α_β in the same basis, that is

$$\boldsymbol{e}_0 = (1, 0, 0, 0) , \quad \boldsymbol{e}_1 = (0, 1, 0, 0) , \quad \boldsymbol{e}_2 = (0, 0, 1, 0) , \quad \boldsymbol{e}_3 = (0, 0, 0, 1) . \tag{2.48}$$

This does not mean that the vectors $\boldsymbol{e}_\alpha$ are unit vectors as the norm of the vector has not yet been fully specified (see section 2.3.4 for more detail)..

Any four linearly-independent vectors form a *basis* in the Minkowski spacetime that can be used for decomposition (2.46) of vectors [12]. Standard practice is to chose a basis consisting of one timelike, and three spacelike vectors. However, it is possible to include into the basis a vector or vectors that are null. Indeed, it is easy to pick up, for an arbitrary timelike vector $\mathbf{a}$ and a spacelike vector $\mathbf{b}$, such real numbers α, β that the linear combination $\mathbf{c} = \alpha\mathbf{a} + \beta\mathbf{b}$ will satisfy $\mathbf{c} \cdot \mathbf{c} = 0$ while $\mathbf{c}$ being non-vanishing. The null vectors are also called isotropic, and the set of all these vectors form a hypersurface in the Minkowski spacetime called *null cone*. Four null vectors define a null-cone basis (null tetrad) that is a basic ingredient of the Newman-Penrose formalism [Frolov, 1979; Penrose, 1968] employed frequently in studying physics of black holes, in string theory, and in other applications dealing with electromagnetic and gravitational radiation as well as with motion of ultra-relativistic particles [Penrose, 2004]. Proposals have been made to use the null tetrads for the purposes of relativistic navigation [Bini *et al.*, 2008; Coll and Pozo, 2006].

The basis, $\boldsymbol{e}_\alpha$, in the Minkowski spacetime and a corresponding decomposition (2.46) are not unique. There is an infinite number of bases and corresponding decompositions of an arbitrary vector $\mathbf{a}$. A new basis $\boldsymbol{e}_{\alpha'}$ (primed indices $\alpha', \beta', \ldots = 0, 1, 2, 3$) is connected to the original one, $\boldsymbol{e}_\alpha$, by a linear transformation

$$\boldsymbol{e}_{\alpha'} = \Lambda^\beta{}_{\alpha'} \boldsymbol{e}_\beta , \tag{2.49}$$

where $\Lambda^\alpha{}_{\beta'}$ are coefficients of decomposition of the basis vector $\boldsymbol{e}_{\alpha'}$ with respect to the basis $\boldsymbol{e}_\alpha$ according to equation (2.49). These coefficients are nothing else but

12) A "basis" does not mean a set of unit vectors. The term refers exclusively to a minimal linearly-independent set of vectors that spans the entire affine space. Moreover, the ability to find a basis for a vector space is a general result in the theory of vector spaces, and has nothing to do with whether it is the Minkowski spacetime or not.

components of vector $\boldsymbol{e}_{\alpha'}$ in the basis $\boldsymbol{e}_\alpha$,

$$\begin{aligned}
\boldsymbol{e}_{0'} &= \left(\Lambda^0{}_{0'}\,,\Lambda^1{}_{0'}\,,\Lambda^2{}_{0'}\,,\Lambda^3{}_{0'}\right)\,,\\
\boldsymbol{e}_{1'} &= \left(\Lambda^0{}_{1'}\,,\Lambda^1{}_{1'}\,,\Lambda^2{}_{1'}\,,\Lambda^3{}_{1'}\right)\,,\\
\boldsymbol{e}_{2'} &= \left(\Lambda^0{}_{2'}\,,\Lambda^1{}_{2'}\,,\Lambda^2{}_{2'}\,,\Lambda^3{}_{2'}\right)\,,\\
\boldsymbol{e}_{3'} &= \left(\Lambda^0{}_{3'}\,,\Lambda^1{}_{3'}\,,\Lambda^2{}_{3'}\,,\Lambda^3{}_{3'}\right)\,.
\end{aligned} \tag{2.50}$$

The $\Lambda^\alpha{}_{\beta'}$ is called the *matrix of transformation*. Elements of the matrix of transformation are real numbers which are independent of each other. Hence, in the most general case, the matrix of transformation has 16 independent components in the Minkowski spacetime. Notice also that the indices of the matrix of transformation belong to different bases. If the bases coincide, the matrix of transformation apparently becomes the unit matrix: $\Lambda^\beta{}_\alpha = \delta^\beta_\alpha$ in accordance with equation (2.48).

Decomposition of vector $\mathbf{a}$ generated by the basis $\boldsymbol{e}_{\alpha'}$ will be different from the one generated by the basis $\boldsymbol{e}_\alpha$ in the sense that equation (2.46) acquires in the new basis the following form

$$\mathbf{a} = a^{0'}\boldsymbol{e}_{0'} + a^{1'}\boldsymbol{e}_{1'} + a^{2'}\boldsymbol{e}_{2'} + a^{3'}\boldsymbol{e}_{3'}\,, \tag{2.51}$$

with the vector components $a^{\alpha'} = (a^{0'}, a^{1'}, a^{2'}, a^{3'})$ being different from those $a^\alpha = (a^0, a^1, a^2, a^3)$. The left sides of equations (2.46) and (2.51) represent one and the same geometric object, the vector $\mathbf{a}$, and they must be equal. It corresponds to the statement that vectors are geometric objects being independent of the choice of coordinates and bases. Hence, the primed components of the vector $\mathbf{a}$ in the new basis $\boldsymbol{e}_{\alpha'}$ are connected to its unprimed components in the old basis $\boldsymbol{e}_\alpha$, by equation

$$a^\alpha = \Lambda^\alpha{}_{\beta'}a^{\beta'}\,, \tag{2.52}$$

which follows immediately from equations (2.47), (2.49) and (2.51).

The basis $\boldsymbol{e}_{\alpha'}$ can be transformed back to the basis $\boldsymbol{e}_\alpha$ with the help of the matrix $\Lambda^{\alpha'}{}_\beta$ of the inverse linear transformation that is defined by equations

$$\Lambda^{\alpha'}{}_\beta\Lambda^\beta{}_{\gamma'} = \delta^{\alpha'}_{\gamma'}\,, \qquad \Lambda^\alpha{}_{\beta'}\Lambda^{\beta'}{}_\gamma = \delta^\alpha_\gamma\,, \tag{2.53}$$

where $\delta^{\alpha'}_{\gamma'}$ and δ^α_γ are the unit matrices referred to the primed and unprimed bases correspondingly. Making use of the inverse matrix of transformation, one can write the inverse transformation

$$\boldsymbol{e}_\alpha = \Lambda^{\beta'}{}_\alpha\boldsymbol{e}_{\beta'}\,, \qquad a^{\alpha'} = \Lambda^{\alpha'}{}_\beta a^\beta\,, \tag{2.54}$$

which can be easily checked by inspection.

One has noticed already that the Newtonian mechanics can be also formulated in terms of a four-dimensional spacetime of events having a fiber bundle structure (see Figure 2.1). Because of this structure the decomposition of a vector $\mathbf{a}$ in two different bases of the Newtonian spacetime has the same time component of the vector in both bases. The fiber bundle structure of the Newtonian spacetime is too rigid and is abandoned in the Minkowski spacetime to make it compatible with physics of ultra-relativistic motion. It makes a drastic difference for the time component of a vector if it becomes dependent on the choice of the basis and participates in the linear transformations on the same footing as its spatial components. Because each basis corresponds to a physical observer, time in the Minkowski spacetime flows differently for different observers.

2.3.4
The metric tensor

The most natural way to describe the property of an arbitrary set of the basis vectors $\boldsymbol{e}_\alpha$ in the Minkowski spacetime is to introduce a 4×4 matrix made of the dot-products of the vectors

$$g_{\alpha\beta} \equiv \boldsymbol{e}_\alpha \cdot \boldsymbol{e}_\beta \ . \tag{2.55}$$

The matrix components $g_{\alpha\beta}$ are real numbers which form a new mathematical object in vector space, called the metric tensor. The metric tensor is used for measuring the metric properties of spacetime (time, angles, lengths, etc.) and for rising and lowering indices of tensors (see sections 2.4.4.3 and 3.3.8.1). For example, the dot product between two vectors **a** and **b** is given in terms of their components and the metric tensor as

$$\mathbf{a} \cdot \mathbf{b} = (a^\alpha \boldsymbol{e}_\alpha) \cdot (b^\beta \boldsymbol{e}_\beta) = a^\alpha b^\beta (\boldsymbol{e}_\alpha \cdot \boldsymbol{e}_\beta) = a^\alpha b^\beta g_{\alpha\beta} \ . \tag{2.56}$$

According to Axiom EA1, the metric tensor is symmetric, $g_{\mu\nu} = g_{\nu\mu}$ and, thus, has only ten algebraically-independent components. The metric tensor $g_{\alpha\beta}$ has its inverse, denoted as $g^{\alpha\beta}$ and defined by the condition

$$g^{\alpha\beta} g_{\beta\gamma} = \delta^\alpha_\gamma \ , \tag{2.57}$$

where the repeated Greek indices assume the Einstein summation from 0 to 3, and $\delta^\mu_\nu = \mathrm{diag}(1, 1, 1, 1)$ is the unit matrix also known as the Kronecker symbol.

Since the metric tensor is formed as a dot product of two basis vectors, it is straightforward to write down the law of transformation of its components. Indeed, picking up a new basis $\boldsymbol{e}_{\alpha'}$, one defines the components of the metric tensor in this basis as

$$g_{\alpha'\beta'} \equiv \boldsymbol{e}_{\alpha'} \cdot \boldsymbol{e}_{\beta'} \ . \tag{2.58}$$

Substituting the law of transformation of the vector bases (2.49) into equation (2.58) and using definition (2.55), one obtains the law of transformation of the components of the metric tensor

$$g_{\alpha'\beta'} = \Lambda^\mu{}_{\alpha'} \Lambda^\nu{}_{\beta'} g_{\mu\nu} \ , \tag{2.59}$$

and the inverse transformation is obtained with the inverse transformation matrix

$$g_{\alpha\beta} = \Lambda^{\mu'}{}_\alpha \Lambda^{\nu'}{}_\beta g_{\mu'\nu'} \ . \tag{2.60}$$

The basis vectors that have the unit norm and are mutually-orthogonal are instrumental in mathematical work and for physical applications. Such bases are called *orthonormal*. One notices that the null-cone bases can not be orthonormal since if four null vectors were orthogonal, decomposition of any timelike or spacelike vector **a** with respect to such a null-basis would yield a null vector, in contradiction with the original nature of the vector **a**. We shall mostly consider orthonormal bases consisting of three spacelike vectors, $\boldsymbol{e}_i$, (i=1,2,3) and one timelike vector $\boldsymbol{e}_0$. The Lorentzian norm of the basis vectors satisfy to the following condition

$$-||\boldsymbol{e}_0||^2 = ||\boldsymbol{e}_1||^2 = ||\boldsymbol{e}_2||^2 = ||\boldsymbol{e}_3||^2 = 1 \,, \tag{2.61}$$

and the Lorentzian dot-product

$$\boldsymbol{e}_\mu \cdot \boldsymbol{e}_\nu = 0 \,, \qquad (\mu \neq \nu). \tag{2.62}$$

In the orthonormal basis the components of the metric tensor assume the most simple, diagonal form $g_{\alpha\beta} = \eta_{\alpha\beta} = \mathrm{diag}(-1, 1, 1, 1)$, or

$$\eta_{\alpha\beta} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} . \tag{2.63}$$

Let us remark that in an inertial frame of reference the metric $g_{\mu\nu}$ should not be necessarily read as (2.63). This is because the concept of the inertial frame of reference does not say anything about the orthogonality of the basis so that in the most general case its vectors can be oriented arbitrary with respect to each other. Nevertheless, orthonormal bases are used for measuring purposes in many practical applications. It is always possible to build an orthogonal basis in the Minkowski spacetime. Indeed, if one starts from a non-orthogonal basis $\boldsymbol{e}_\alpha$, and apply a linear transformation to a basis $\boldsymbol{e}_{\alpha'}$ consisting of the product of a (pseudo)-orthogonal matrix $O^\beta{}_{\alpha'}$ with a diagonal matrix $D^\gamma{}_\beta$, one can always transform the non-diagonal metric tensor $g_{\alpha\beta}$ to a diagonal form $g_{\alpha'\beta'} = \eta_{\alpha'\beta'}$. This means that the new basis

$$\boldsymbol{e}_{\alpha'} = O^\beta{}_{\alpha'} D^\gamma{}_\beta \boldsymbol{e}_\gamma \,, \tag{2.64}$$

is orthogonal and represents the normalized eigenvectors of the metric $g_{\alpha\beta}$.

The signature of the metric is defined as the sum of its diagonal components that is also called a trace of the tensor . The Minkowski metric $\eta_{\alpha\beta}$ has the signature +2 that remains the same in all acceptable bases [13)], according to a signature-stability theorem by Silvester [Schouten, 1954]. Though the signature of the metric tensor in the Minkowski spacetime is invariant it does not mean that the metric has always the form shown in equation (2.63). Mathematically, it is possible to build orthonormal bases in which the metric may have a diagonal form $\eta_{\alpha\beta} = \mathrm{diag}(1, -1, 1, 1)$, or $\eta_{\alpha\beta} = \mathrm{diag}(1, 1, -1, 1)$ or $\eta_{\alpha\beta} = \mathrm{diag}(1, 1, 1, -1)$. Nevertheless, only the bases in which the metric is reduced to $\eta_{\alpha\beta} = \mathrm{diag}(-1, 1, 1, 1)$ can be associated with physical observations.

13) One assumes that transformation of the basis does not involve reflections and dilatations of coordinate axes.

2.3.5 The Lorentz group

2.3.5.1 *General properties*

There are linear transformations (2.49) that leave the basis of the Minkowski spacetime orthogonal. Not every linear transformation will guarantee fulfilment of equations (2.61), (2.62) or (2.63). The entire set of matrices which preserve the orthonormality of bases, form a continuous group of transformations known as the *Lorentz group*. It will turned out that a subset of linear transformations from this group coincides with the Lorentz transformations (2.29) from electrodynamics. However, the existence of the group of the Lorentz transformations in the Minkowski spacetime does not rely upon a specific propagation property of any physical field and relates exclusively to the symmetries of the spacetime endowed with the fundamental speed c and the null cone. On the other hand, any physical theory that pretends to be consistent with relativity must be Lorentz-invariant, that is its equations must admit the freedom of the Lorentz transformations.

The Lorentz group is denoted with $O(1,3)$ which is a standard notation for a generalized orthogonal group. The condition of orthogonality of the bases is effectively reduced to the condition of form-invariance of the Minkowski metric, which means that the Minkowski metric must have the same diagonal form (2.63) in any orthogonal basis (frame of reference). Equation of transformation (2.59) entails

$$\eta_{\alpha'\beta'} = \Lambda^{\mu}{}_{\alpha'}\Lambda^{\nu}{}_{\beta'}\eta_{\mu\nu}\ . \tag{2.65}$$

Inserting the values of $\eta_{\mu\nu} = \mathrm{diag}(-1,1,1,1)$ and $\eta_{\alpha'\beta'} = \mathrm{diag}(-1,1,1,1)$, one sees that the matrix $\Lambda^{\mu}{}_{\alpha'}$ is pseudo-orthogonal, that is, it satisfies

$$-\Lambda^{0}{}_{\alpha'}\Lambda^{0}{}_{\beta'} + \Lambda^{i}{}_{\alpha'}\Lambda^{i}{}_{\beta'} = \begin{cases} 0\,, & \text{if} \quad \alpha' \neq \beta'\,, \\ -1\,, & \text{if} \quad \alpha' = \beta' = 0\,, \\ 1\,, & \text{if} \quad \alpha' = \beta' = 1,2,3\,, \end{cases} \tag{2.66}$$

where the repeated spatial indices assume the Einstein summation from 1 to 3. In case of $\alpha' = \beta' = 0$ equation (2.66) reads

$$\left(\Lambda^{0}{}_{0'}\right)^2 = 1 + \Lambda^{i}{}_{0'}\Lambda^{i}{}_{0'}\ , \tag{2.67}$$

wherefrom the absolute value of the matrix element

$$|\Lambda^{0}{}_{0'}| \geq 1\ . \tag{2.68}$$

Moreover, equation (2.67) assumes that the set of all transformations parameterized with matrices $\Lambda^{\mu}{}_{\alpha'}$ will be disconnected into two sub-sets: those with positive and with negative value of the matrix element $\Lambda^{0}{}_{0'}$. We also conclude from the diagonal form of the Minkowski metric that the determinant of the Minkowski metric, $\det\left[\eta_{\alpha\beta}\right] = \det\left[\eta_{\mu'\nu'}\right] = -1$. This, along with equation (2.65), demands that the determinant of the matrix of transformation

$$\det\left[\Lambda^{\mu}{}_{\alpha'}\right] = \pm 1\ , \tag{2.69}$$

which indicates to one more division to two other sub-sets: the one with a positive and an other with a negative determinant of the matrix.

Thus, the entire Lorentz group of transformations, $O(1,3)$, is split into four topologically separated pieces:

$SO^+(1,3)$ - proper orthochronous transformations with $\Lambda^0_{0'} > 0$, $\det[\Lambda^\mu{}_{\alpha'}] = 1$. It consists of those Lorentz transformations that preserve the orientation of spatial axes (spatial reflections are not allowed) and direction of time.

$\mathbb{P}SO^+(1,3)$ - improper orthochronous transformations with $\Lambda^0_{0'} > 0$, $\det[\Lambda^\mu{}_{\alpha'}] = -1$. It consists of the Lorentz transformations that preserve the direction of time but change the orientation of spatial axes to opposite.

$\mathbb{PT}SO^+(1,3)$ - the proper non-orthochronous transformations with $\Lambda^0_{0'} < 0$, $\det[\Lambda^\mu{}_{\alpha'}] = 1$. It consists of the Lorentz transformations that change both the direction of time and orientation of spatial axes.

$\mathbb{T}SO^+(1,3)$ - the improper non-orthochronous transformations with $\Lambda^0_{0'} < 0$, $\det[\Lambda^\mu{}_{\alpha'}] = -1$. It consists of the Lorentz transformations that change the direction of time but leaves the orientation of spatial axes the same.

The latter three sets can be obtained from the first one by means of the operation of inversion of the time, $\mathbb{T}$, and that of the spatial axes, $\mathbb{P}$. It explains the notations for these three sub-groups: $\mathbb{P}SO^+(1,3)$, $\mathbb{PT}SO^+(1,3)$, $\mathbb{T}SO^+(1,3)$. Only the proper orthochronous subgroup $SO^+(1,3)$ is used in astronomy. The other sub-groups are relevant to the particle physics.

One sees that a subset of the Lorentz transformations is characterized by whether or not its elements reverse the orientation of space and/or time. Transformations which reverse either the orientation of time or space (but not both), have determinant -1; the ones that reverse neither or both have determinant $+1$.

The sub-set $SO^+(1,3)$ makes to a sub-group called the restricted Lorentz group. None of the other three sets $\mathbb{P}SO^+(1,3)$, $\mathbb{PT}SO^+(1,3)$, and $\mathbb{T}SO^+(1,3)$ is a sub-group, because none of these contains the identity element - the unit matrix $\delta_{\mu\nu}$. However, each of these sub-sets makes a sub-group together with $SO^+(1,3)$. More specifically,

– $\mathbb{P}SO^+(1,3)$ and $SO^+(1,3)$ constitute the sub-group $O^+(1,3)$ of orthochronous (i.e., preserving the direction of time) Lorentz transformations,
– $\mathbb{PT}SO^+(1,3)$ and $SO^+(1,3)$ constitute the subgroup $SO(1,3)$ of proper (i.e., having $\det[\Lambda^\mu{}_{\alpha'}] = 1$) Lorentz transformations,
– $\mathbb{T}SO^+(1,3)$ and $SO^+(1,3)$ constitute the subgroup of orthochorous (i.e., preserving the sign of the volume of space) Lorentz transformations.

2.3.5.2 *Parametrization of the Lorentz group*

Each orthonormal basis $\boldsymbol{e}_\alpha$ in the Minkowski spacetime is identified with a Cartesian coordinate chart x^α, where coordinates x^α of an event P are considered as components of a position vector $\mathbf{r} = x^\alpha \boldsymbol{e}_\alpha$ of this event with respect to the origin of the coordinates. Another Cartesian coordinate chart, $x^{\alpha'}$ is associated with an orthonormal basis $\boldsymbol{e}_{\alpha'}$. Position vector of the same event P is $\mathbf{r}' = x^{\alpha'}\boldsymbol{e}_{\alpha'}$, and it is related to $\mathbf{r}$ by a constant translation vector $\mathbf{a} = a^\alpha \boldsymbol{e}_\alpha$, which yields

$$\mathbf{r} = \mathbf{a} + \mathbf{r}' , \tag{2.70}$$

or in the coordinate components

$$x^\alpha = a^\alpha + \Lambda^\alpha{}_{\beta'} x^{\beta'} . \tag{2.71}$$

This is the most general transformation between the Cartesian coordinates in the Minkowski spacetime [14], which is called the Poincaré transformation. The group of the Lorentz transformations correspond to the case when the translation vector $\mathbf{a} = 0$. Equation (2.71) is simplified to

$$x^{\alpha'} = \Lambda^{\alpha'}{}_\mu x^\mu , \tag{2.72}$$

where one has used the inverse matrix $\Lambda^{\alpha'}{}_\mu$.

It can be demonstrated that the restricted Lorentz group $SO^+(1,3)$ can be generated by a direct product of three independent transformations - a Lorentz boost and two three-dimensional rotation of spatial axes. More specifically, each element Λ of $SO^+(1,3)$ (indices of the matrix Λ have been omitted) can be always written down as a product of three matrices: $\Lambda = R_1 L R_2$, with L being the matrix of the Lorentz boost and R_1 and R_2 being the orthogonal matrices of spatial rotations [Richtmyer, 1982]. The Lorentz boost generalizes the Galilean transformation of classical mechanics, and represents transformation from static to a moving frame which depends in special relativity on both time and space coordinates of one of the three coordinate planes: (t, x^1), (t, x^2), or (t, x^3). Hence, there are three independent boosts L_x, L_y and L_z each of which is parameterized with a parameter, let say, $\theta_x, \theta_y, \theta_z$ respectively. The pure spatial rotations R_1 and R_2 are parameterized with three Euler angles. Therefore, the restricted Lorentz group $SO^+(1,3)$ is characterized by six parameters. A proverbial example is a boost transformation L_x describing motion of frame S' with respect to S in positive direction of x^1 axis with velocity V, and leaving two of the spatial coordinates, for example, x^2 and x^3, untouched. The matrix of such a boost is parameterized by one parameter only θ_x (which is denoted below as θ), and the matrix of the boost is

$$L_x(\theta) \equiv \begin{bmatrix} \Lambda^{0'}{}_0 & \Lambda^{0'}{}_1 & \Lambda^{0'}{}_2 & \Lambda^{0'}{}_3 \\ \Lambda^{1'}{}_0 & \Lambda^{1'}{}_1 & \Lambda^{1'}{}_2 & \Lambda^{1'}{}_3 \\ \Lambda^{2'}{}_0 & \Lambda^{2'}{}_1 & \Lambda^{2'}{}_2 & \Lambda^{2'}{}_3 \\ \Lambda^{3'}{}_0 & \Lambda^{3'}{}_1 & \Lambda^{3'}{}_2 & \Lambda^{3'}{}_3 \end{bmatrix} = \begin{bmatrix} \cosh\theta & -\sinh\theta & 0 & 0 \\ -\sinh\theta & \cosh\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} , \tag{2.73}$$

Although this parametrization through the boost parameter θ is not obligatory, its strong side is that the so-chosen parameter is additive like axial rotations in the same

14) Curvilinear coordinates can be also used in the Minkowski spacetime but relationship between them is not given by the linear transformation.

plane in the Euclidean geometry. Indeed, making two consecutive boosts, $L_x(\theta_1)$ and $L_x(\theta_2)$, and using formulae for the hyperbolic functions,

$$\sinh(\theta_1+\theta_2) = \sinh\theta_1\cosh\theta_2+\sinh\theta_2\cosh\theta_1\,, \tag{2.74}$$

$$\cosh(\theta_1+\theta_2) = \cosh\theta_1\cosh\theta_2+\sinh\theta_2\sinh\theta_1\,, \tag{2.75}$$

one gets for two successive boost transformations in the same (t, x) plane,

$$L_x(\theta_2)L_x(\theta_1) = L_x(\theta_2+\theta_1)\,. \tag{2.76}$$

In terms of the coordinates $x^\alpha = (ct, x, y, z)$ and $x^{\alpha'} = (ct', x', y', z')$, the boost transformation (2.73) explicitly reads as

$$ct' = ct\cosh\theta - x\sinh\theta\,,\qquad x' = -ct\sinh\theta + x\cosh\theta\,,\qquad y' = y\,,\qquad z' = z\,. \tag{2.77}$$

Similar equations can be written down for the Lorentz boosts in (t, y) and (t, z) planes.

The boost parameter θ is also called *rapidity* and is directly related to the velocity of one basis with respect to another. Suppose that one observer in the inertial frame S' keeps a particle at rest, $dx'/dt' = 0$, and a second observer in the inertial frame S sees it moving with velocity $dx/dt = V$ in positive direction of x axis of the inertial frame $x^\alpha = (ct, x, y, z)$. Differentiating equations (2.77) and using the definitions of the particle's velocity in the two frames, one can easily conclude that

$$\tanh\theta = \frac{V}{c}\,, \tag{2.78}$$

which yields

$$\sinh\theta = \frac{V/c}{\sqrt{1-\frac{V^2}{c^2}}}\,,\qquad \cosh\theta = \frac{1}{\sqrt{1-\frac{V^2}{c^2}}}\,. \tag{2.79}$$

Parametrization (2.79) of the boost with physical velocity V make equations (2.77) look exactly as a Lorentz transformation (2.29) from the electrodynamics

$$ct' = \gamma(ct-\beta x)\,,\qquad x' = \gamma(x-\beta ct)\,,\qquad y' = y\,,\qquad z' = z\,, \tag{2.80}$$

where

$$\beta \equiv V/c\,,\qquad \gamma \equiv \frac{1}{\sqrt{1-\beta^2}}\,. \tag{2.81}$$

The Lorentz transformation changes directions of the coordinate axes (ct', x') with respect to axes (ct, x). The new axes (ct', x') make the angle $\beta = V/c$ with the old axes (t, x) as shown in Figure 2.5 The introduction of parameter V elucidates that the boost transformation (2.80) that leaves the Minkowski metric tensor form-invariant is equivalent to the Lorentz transformation (2.29) of the coordinates in the Maxwell equations. This explains why the name of the Lorentz group was given to the group $O(1, 3)$.

Equation (2.80) indicates that none physical signal can travel at a speed faster than the fundamental speed c. According to Einstein and Minkowski it appears in the

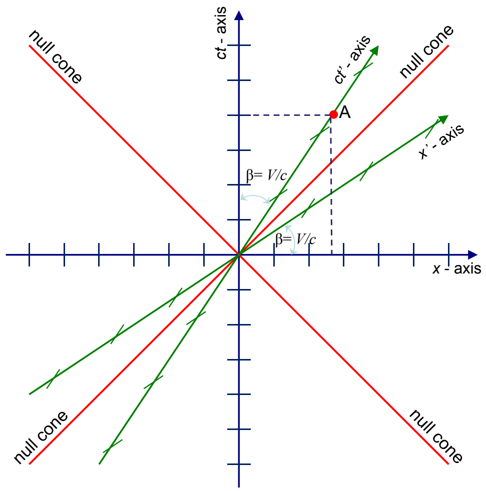

Figure 2.5 Minkowski diagram of the Lorentz transformation (2.80) from coordinates (ct, x) representing an inertial reference frame S, to coordinates (ct', x') representing an inertial reference frame S' moving with respect to S with speed V in positive direction of x-axis. The larger velocity V the closer are the axes (ct', x') to the null cone. Axis ct is marked with some arbitrary-chosen unit of time. Corresponding marks of the same unit of time plotted on ct' axis are obtained by drawing lines parallel to x' axis. The unit of time in moving frame is longer than that in the static frame showing that time in the moving frame runs slower than in the static frame (count, for example, the number of time units for the event A in both frames). The diagram uses the geometrized system of units in which $c = 1$.

Lorentz transformations independently of the Maxwell equations. Einstein's understanding of motion was in the spirit of Leibniz: by rejecting the concept of absolute motion, Einstein granted himself an opportunity to set the relativity principle as a fundamental property of the spacetime, not a consequence of matter-æther interactions preventing us from observing æther as was hypothesized by Lorentz and Poincaré. Within Einstein's approach, equations for *all* physical fields should transform in accordance with the Lorentz group not because of an exclusive property of a particular field but because of a pseudo-Euclidean nature of the spacetime. Accordingly, the invariant speed c emerging in the Lorentz transformation (2.80) is a fundamental constant for *all* types of physical interactions setting the ultimate speed of their propagation.

The matrix (2.73) of the boost transformation in the (t, x) plane can be generalized for arbitrary spatial direction of the relative motion of two frames. Assuming that there is no rotation of the spatial axes, the boost transformation $x^{\alpha'} = \Lambda^{\alpha'}{}_{\beta} x^{\beta}$ is given by the following matrix

$$\Lambda^{0'}{}_0 = \gamma\,, \tag{2.82a}$$

$$\Lambda^{i'}{}_0 = -\gamma\beta^i\,, \tag{2.82b}$$

$$\Lambda^{0'}{}_i = -\gamma\beta^i\,, \tag{2.82c}$$

$$\Lambda^{i'}{}_j = \delta_{ij} + (\gamma - 1)\frac{\beta^i\beta^j}{\beta^2}\,, \tag{2.82d}$$

where $\beta^i = V^i/c$ is the velocity of the boost in the direction $n^i = V^i/V$. This transformation describes motion of the primed basis, $\boldsymbol{e}_{\alpha'}$, with velocity V^i as seen from the unprimed basis $\boldsymbol{e}_\alpha$. From the point of view of the observer at rest in the primed basis, the unprimed basis moves with opposite velocity $-V^i$. It means that the matrix of the inverse boost transformation $\Lambda^{\alpha}_{\mu'}$ is obtained from equations (2.82) after replacing $V^i \to -V^i$. It is easy to check that the matrix (2.82) is reduced to the familiar form (2.73) in case of motion along x-axis where $V^i = (V, 0, 0)$.

Let us also notice that the boost matrix (2.82) is symmetric but the general Lorentz transformation matrix need not be symmetric. It is also worth mentioning that the set of matrices of the pure Lorentz boosts (2.82) does not represent a sub-group of the Lorentz group. This is because two consecutive Lorentz transformations performed in two different directions are not equivalent to a single Lorentz boost as oppose to two Lorentz boosts in the same direction. Two Lorentz boosts in different directions are equivalent to the product of a pure Lorentz boost and a spatial rotation. This phenomenon was discovered by L.H. Thomas in connection to his study of the atomic energy levels. The spatial rotation of the inertial frame caused by the successive Lorentz transformations in different directions is called the *Thomas precession*.

2.3.6
The Poincaré Group

Spatial translations and the Lorentz group form a broader group named after Poincaré. The Poincaré group is the largest one, under which the equations of special relativity are form-invariant [Kim and Noz, 1986]. This is also a group under which the Maxwell equations stay form-invariant though the latter admits more general class of transformations due to the presence of a conformal symmetry of the equations intimately associated with the absence of photon's rest mass [Fushchich and Nikitin, 1987]. The Poincaré group is also important in quantum mechanics as the values of mass and spin of an individual elementary particle correspond to irreducible representations of the Poincaré group as was first noted by Wigner [1939].

The Poincaré group is parameterized with 10 parameters among which 6 are parameters of the Lorentz group (boosts and rotations) and 4 parameters are translations shown in equation (2.71) as components of a constant vector $\mathbf{a} = a^\alpha \boldsymbol{e}_\alpha$. The translations act transitively on the bundle of future timelike directions of the Minkowski spacetime. It means that the vector $\mathbf{a}$ is really determined up to a translation along the timelike worldline. Indeed, let us consider an inertial frame S with basis $\boldsymbol{e}_\alpha$ and coordinates x^α and another inertial frame S' with basis $\boldsymbol{e}_{\alpha'}$ and coordinates $x^{\alpha'}$ which origin moves with respect to S along a timelike worldline in the direction defined by a four-velocity vector $\mathbf{u} = u^\alpha \boldsymbol{e}_\alpha$ where $u^\alpha = (\gamma, \gamma\beta^i) = \Lambda^\alpha{}_{0'}$. Let A be an event which is determined by a vector $\mathbf{r}$ in the frame S. This vector can be represented as an algebraic sum of two vectors in equation (2.70) or (2.71) where vector $\mathbf{a}$ characterizes position of the origin of frame S' at the time $t' = 0$. However, equation (2.71) can be written down more explicitly as

$$x^\alpha = a^\alpha + c\Lambda^\alpha{}_{0'} t' + \Lambda^\alpha{}_{j'} x^{j'} \ . \tag{2.83}$$

Defining a new vector of translation, $\mathbf{z} = z^\alpha \boldsymbol{e}_\alpha$, having components

$$z^\alpha(t') = a^\alpha + c\Lambda^\alpha{}_{0'} t' = a^\alpha + u^\alpha t' \ , \tag{2.84}$$

one presents equation (2.83)

$$x^\alpha = z^\alpha(t') + \Lambda^\alpha{}_{j'} x^{j'} \ , \tag{2.85}$$

showing coordinates of the same event A as an algebraic sum of a vector of the origin of the moving frame S' taken at the time t', and a purely spatial vector $\mathbf{x}' = \Lambda^\alpha{}_{j'} x^{j'}$. Algebraic operations corresponding to the Poincaré transformation are shown in Figure 2.6.

Presentation of coordinates of event A in the form of equation (2.85) corresponds to the construction of the inertial reference frame of an observer being equipped with an ideal clock and rulers for measuring time and length, and moving along a straight line along the direction being determined by its four-velocity $\mathbf{u} = u^\alpha \boldsymbol{e}_\alpha$. Coordinate time t' in the reference frame S' corresponds to the proper time τ of observer measured with the help of the ideal clock. Spatial coordinates $x^{\beta'}$ of an event A represent in the frame S' the proper distance ℓ from the observer to the event measured with the help of the ideal ruler: $\ell = \sqrt{x'^2 + y'^2 + z'^2}$. Because the basis $\boldsymbol{e}_{\alpha'}$ of the moving inertial observer is related to the matrix of the Lorentz transformation $\Lambda^\alpha{}_{b'}$ by equation (2.50), one can recast the Poincaré transformation (2.85) to the following, equivalent form

$$\mathbf{x} = \mathbf{z} + \boldsymbol{e}_{j'} x^{j'} \ , \tag{2.86}$$

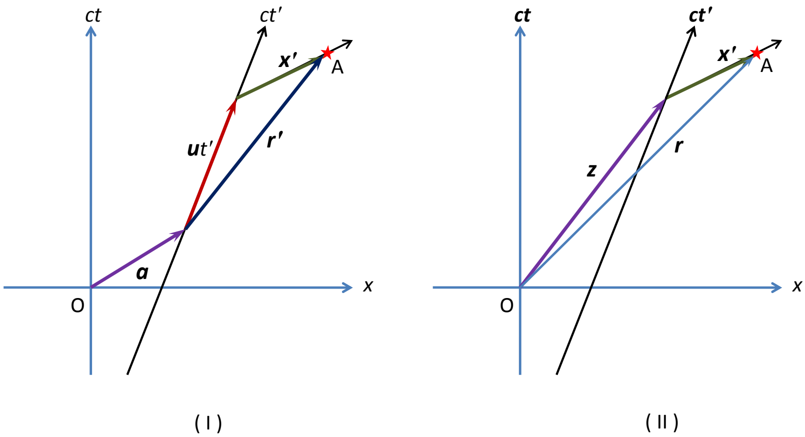

Figure 2.6 Transitive action of the Poincaré group along straight timelike lines allows us to represent the inertial coordinates x^α of event A as an algebraic sum of two vectors, $\mathbf{z} = \mathbf{a} + \mathbf{u}t'$, and $\mathbf{x}'$. Vector $\mathbf{z}$ is obtained by a translation along the worldline of motion of the frame S' at the distance $\mathbf{u}t'$, where $\mathbf{u}$ is four-velocity of motion of the frame S' with respect to S. The equivalence of presentation $\mathbf{a} + \mathbf{r}' = \mathbf{a} + (\mathbf{u}t' + \mathbf{x}') = (\mathbf{a} + \mathbf{u}t') + \mathbf{x}' = \mathbf{z} + \mathbf{x}'$ follows from the axiom WA2 of associativity of vector addition.

where

$$\mathbf{z} = \mathbf{a} + \mathbf{u}\tau \,, \tag{2.87}$$

is a four-vector pointing out from the origin of the inertial frame S to the origin of the inertial frame S' of the moving observer, $\mathbf{u}$ is the constant four-velocity of the observer.

2.4 Tensor Algebra

The goal of this section is to provide an elementary primer on contravariant and covariant tensors that provides a sufficiently rigorous mathematical explanation of their properties in vector space - the tensor algebra.

2.4.1 Warming up in three dimensions: scalars, vectors, what next?

Let us start out with several illustrative examples of mathematical objects from three-dimensional Euclidean space. The most simple object is a scalar. When a quantity is scalar - its numerical value at a given event on the manifold, by definition, does not depend upon the choice of coordinates. It means that if one has a coordinate chart x^α and a scalar function $F(x^\alpha)$, it will transform to another function $F'(\ )$ after changing coordinates $x^{\alpha'} = x^{\alpha'}(x^\beta)$ in such a way that its numerical value will remain the same,

$$F'(x^{\alpha'}) = F(x^\alpha)\ , \tag{2.88}$$

Examples are the thermodynamical quantities like density, temperature and pressure. More complicated example is a norm of a vector which does not depend on the choice of coordinates and/or vector basis.

Vectors are slightly more complicated than scalars. A good example of vectors is the second law of Newton, $m\boldsymbol{a} = \boldsymbol{F}$, where m is the inertial mass of particle, $\boldsymbol{a} = (a^i)$ - its acceleration, and $\boldsymbol{F} = (F^i)$ - the force exerted on the mass ($i = 1, 2, 3$). Although in different three-dimensional coordinates the components of force F^i will be expressed with different triples of numbers, these triples are connected to one another by a coordinate transformation law that makes the force F^i a vector. One can imagine the vector as an arrow attached to the particle along with the coordinate basis visualized as a rectangular corner. Nothing prohibits us to take another coordinate bases rotated with respect to the first one. Since the two bases have different orientations, the projections of the vector on their axes will be different. However, the force F^i is a physical object existing by itself, no matter what coordinates are employed to express the numerical values of its components. This establishes the law of transformation between the three components of the vector. The same can be said about the vector of acceleration a^i: its projections will be transforming from one coordinate basis to another in a similar way as the projections of force F^i. Despite these changes, a^i is a vector existing independently from the bases employed. Taken that in Newton's mechanics the mass m is a scalar, one draws a conclusion that in the Newtonian mechanics the law $ma^i = F^i$ is vectorial, and that it has the same physical content in any coordinates.

Another example of vectors is delivered by the Ohm law in a linear isotropic material, in which the electric field $\boldsymbol{E} = (E^i)$ and the current density $\boldsymbol{J} = (J^i)$ are interconnected in coordinates x^i via

$$J^i = \sigma E^i \ , \tag{2.89}$$

where σ is the material's conductivity. Both E^i and J^i are regarded as spatial vectors insofar as the bases are Cartesian and only rotational transformations are concerned. In other words, under a three-dimensional spatial rotation of the coordinate basis,

$$x^{i'} = R^{i'}{}_k x^k \ , \tag{2.90}$$

where $R^{i'}{}_j$ is the orthogonal matrix, both J^i and E^i should transform as vectors. In the unprimed coordinates, x^i, their projections are related to one another through equation (2.89) while in the primed coordinates, $x^{i'}$, they will assume different values [15)]

$$E^{i'} = R^{i'}{}_k E^k \ , \tag{2.91}$$
$$J^{i'} = R^{i'}{}_k J^k \ , \tag{2.92}$$

but still will be related in a similar way

$$J^{i'} = \sigma' E^{i'} \ , \tag{2.93}$$

where $\sigma' = \sigma$ because it is a scalar in isotropic material and does not depend on the coordinate choice.

Do one has more complicated geometric objects? To answer this question, let us extend the law (2.89) to a material, which is no longer isotropic. In coordinates x^i, the law will read

$$J^i = \sigma^i{}_k E^k \ , \tag{2.94}$$

where $\sigma^i{}_k$ is a 3×3 matrix characterizing the conductivity of the material in various directions. If relationship (2.94) pretends to be a physical law, it must have the same form in any other coordinates $x^{i'}$

$$J^{i'} = \sigma^{i'}{}_{k'} E^{k'} \ . \tag{2.95}$$

Assuming that the primed and unprimed coordinates are related by the rotation (2.90) and accounting for the transformation laws (2.91), (2.92) for the electric field and current, one should come to the conclusion that the components of the matrix of conductivity σ^i_k should transform in a prescribed way

$$\sigma^{i'}{}_{k'} = R^{i'}{}_p R^q_{k'} \sigma^p{}_q \ , \tag{2.96}$$

where the matrix $R^q_{k'}$ is the inverse of the matrix $R^{q'}_k$. Any geometric object with two indices that transforms like the conductivity matrix is called tensor of a second rank. Its transformation law contains two matrices of transformation $R^{i'}{}_p$ of the coordinate basis. Notice that the transformation law for vectors contains one transformation

15) One reminds the Einstein summation rule is applied: repeated indices assumes summation over corresponding values.

matrix, and that for scalars does not contain the transformation matrix at all. Hence, vectors are called tensors of the first rank, and scalars are tensors of zero rank.

The latter example prompts us to think that there are tensors of higher rank that are more complicated geometric objects expressing some other physical properties in a coordinate-independent way. It is reasonable to think that a tensor of rank n has n indices, and its transformation law under a smooth coordinate transformation, $x^{i'} = x^{i'}(x)$, must be given by a product of n matrices of transformation

$$\Lambda^{i'}{}_{j} = \frac{\partial x^{i'}}{\partial x^{j}} , \tag{2.97}$$

or their inverse. The key point about application of tensors in physics is that they are *not* just matrices but geometric objects residing on manifold and possessing certain physical properties; while considered in different coordinate charts (or bases) these objects are expressed by different matrix components connected to one another by a linear transformation. It may be useful to notice that tensors make only one class of geometric objects capable of universal describing the physical laws of nature. There are other objects, called spinors, that complement tensors. They are employed in particle physics to describe properties of particles with half-integer spin, and shall not be considered in this book as it goes beyond its scope.

2.4.2 Covectors

2.4.2.1 *Axioms of covector space*

We shall introduce tensors in the Minkowski spacetime starting from covectors which are defined as linear functions on a real vector space whose elements

$$\mathbf{a} = a^{\mu}\boldsymbol{e}_{\mu} , \quad \mathbf{b} = b^{\mu}\boldsymbol{e}_{\mu} , \quad \mathbf{c} = c^{\mu}\boldsymbol{e}_{\mu} , \ldots \tag{2.98}$$

are four-dimensional vectors (also called contravariant tensors of the first rank) having contravariant components a^{μ}, b^{μ}, c^{μ}, etc. with respect to a basis $\boldsymbol{e}_{\mu}$ in this space. Let us consider a set of linear functions $\tilde{\boldsymbol{\alpha}}$, $\tilde{\boldsymbol{\beta}}$, $\tilde{\boldsymbol{\gamma}}$, etc., mapping each vector into a scalar. Such functions are called covectors, or linear 1-forms, or covariant tensors of the first rank. They obey the following axioms:

VA1 For any two covectors $\tilde{\boldsymbol{\alpha}}$ and $\tilde{\boldsymbol{\beta}}$, there always exists a third covector denoted $\tilde{\boldsymbol{\alpha}} + \tilde{\boldsymbol{\beta}}$, such that its action on an arbitrary vector $\mathbf{a}$ gives a scalar that is a sum of $\tilde{\boldsymbol{\alpha}}(\mathbf{a})$ and $\tilde{\boldsymbol{\beta}}(\mathbf{a})$. Mathematically,

$$(\tilde{\boldsymbol{\alpha}} + \tilde{\boldsymbol{\beta}})(\mathbf{a}) = \tilde{\boldsymbol{\alpha}}(\mathbf{a}) + \tilde{\boldsymbol{\beta}}(\mathbf{a}) .$$

This equation makes it clear that $\tilde{\boldsymbol{\alpha}} + \tilde{\boldsymbol{\beta}} = \tilde{\boldsymbol{\beta}} + \tilde{\boldsymbol{\alpha}}$.

VA2 Covectors obey the associative law

$$\left[(\tilde{\boldsymbol{\alpha}} + \tilde{\boldsymbol{\beta}}) + \tilde{\boldsymbol{\gamma}}\right](\mathbf{a}) = \left[\tilde{\boldsymbol{\alpha}} + (\tilde{\boldsymbol{\beta}} + \tilde{\boldsymbol{\gamma}})\right](\mathbf{a})$$

VA3 For any covector $\tilde{\alpha}$, there exists a negative covector denoted $-\tilde{\alpha}$, such that their values on any argument $\mathbf{a}$ are opposite

$$(-\tilde{\alpha})(\mathbf{a}) = -\left[\tilde{\alpha}(\mathbf{a})\right] \; .$$

VA4 There exists a zero covector $\tilde{\mathbf{0}}$, i.e., a function mapping any vector $\mathbf{a}$ to zero

$$\tilde{\mathbf{0}}(\mathbf{a}) = 0 \; .$$

VA5 For any covector $\tilde{\alpha}$ and any real numbers a and b

$$\left[(a+b)\tilde{\alpha}\right](\mathbf{a}) = (a\tilde{\alpha})(\mathbf{a}) + (b\tilde{\alpha})(\mathbf{a}) \; .$$

VA6 For any real numbers a and b and a covector $\tilde{\alpha}$, there exists a covector denoted $ab\tilde{\alpha}$, such that for any argument $\mathbf{a}$

$$a(b\tilde{\alpha})(\mathbf{a}) = (ab\tilde{\alpha})(\mathbf{a}) \; .$$

This gives us an opportunity to define an operation of multiplying a covector by a number.

VA7 For any covector $\tilde{\alpha}$ there exists a unit number 1 such that

$$1\tilde{\alpha}(\mathbf{a}) = \alpha(\mathbf{a}) \; .$$

VA8 Covectors satisfies the following distributive law

$$\left[a(\tilde{\alpha} + \tilde{\boldsymbol{\beta}})\right](\mathbf{a}) = (a\tilde{\alpha})(\mathbf{a}) + (a\tilde{\boldsymbol{\beta}})(\mathbf{a}) \; ,$$

for any covectors $\tilde{\alpha}$ and $\tilde{\boldsymbol{\beta}}$, any numbers a, b, and any vector argument $\mathbf{a}$.

The linearity of covectors means that if one takes an argument of any covector $\tilde{\alpha}$ in the form of a linear combination of vectors, the following equality

$$\tilde{\alpha}(a\mathbf{a} + b\mathbf{b}) = a\tilde{\alpha}(\mathbf{a}) + b\tilde{\alpha}(\mathbf{b}) \; , \tag{2.99}$$

must be valid *by definition* for arbitrary vectors $\mathbf{a}$, $\mathbf{b}$, and arbitrary real numbers a, b. All in all, covectors (linear forms) obey exactly the same set of axioms WA1–WA8 for vector space. Hence covectors make a vector space of their own, a space that is called dual to the original vector space.

Vector are visualized geometrically as an arrow connecting two events. Because covectors are linear functions of vectors their geometric image is to be complimentary to the arrow (straight line). It is achieved by representing each covector as a set of parallel, continuously numbered hyperplanes as shown in Figure 2.7. Covectors are geometric objects that allow us to “measure" vectors in the absence of any other geometric structures on the manifold.

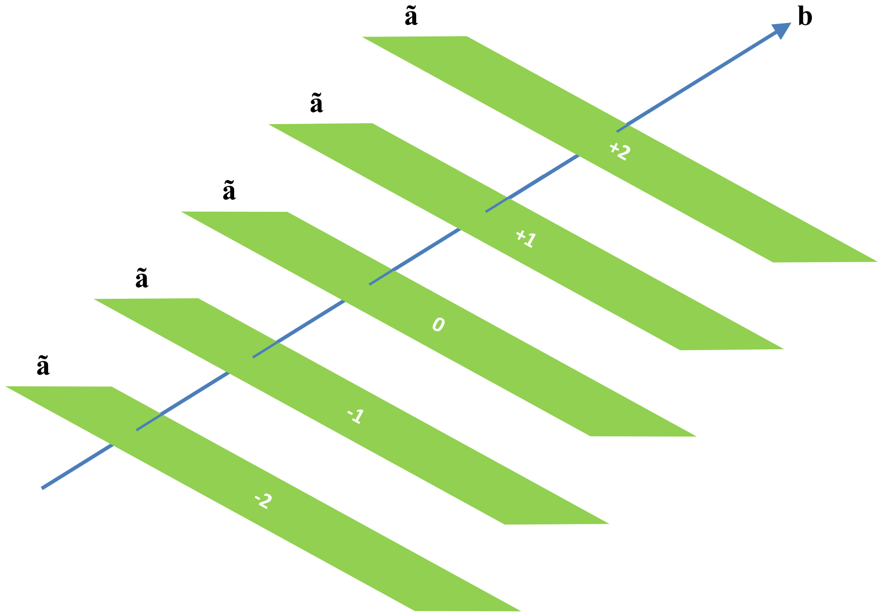

Figure 2.7 Covector $\tilde{\mathbf{a}}$ is shown as a set of parallel planes each having a number. "Direction" of the covector is defined by the orientation and the numbering order of the planes $(..., -2, -1, 0, +1, +2, ...)$. The value of the covector on a vector $\boldsymbol{b}$ is denoted as $\tilde{\mathbf{a}}(\boldsymbol{b})$ that is geometrically equal to the number of planes crossed by the arrow corresponding to the given vector.

2.4.2.2 *The basis in the covector space*

Action of a covector $\tilde{\alpha}$ on an arbitrary vector, $\mathbf{a} = a^\mu \boldsymbol{e}_\mu$, can be written down as

$$\tilde{\alpha}(\mathbf{a}) = \tilde{\alpha}(a^\mu \boldsymbol{e}_\mu) = a^\mu \tilde{\alpha}(\boldsymbol{e}_\mu) \ , \tag{2.100}$$

where one has applied equation (2.99). Let us introduce the numbers

$$\alpha_\mu \equiv \tilde{\alpha}(\boldsymbol{e}_\mu) \ , \tag{2.101}$$

that define how covector $\tilde{\alpha}$ acts on the basis vectors $\boldsymbol{e}_\mu$. These numbers do *not* depend upon the choice of a particular vector argument in equation (2.100) and their set fully defines the covector $\tilde{\alpha}$, i.e., fully determines how $\tilde{\alpha}$ acts on an arbitrary vector $\mathbf{a}$. Equation (2.100) can be re-written in the form

$$\tilde{\alpha}(\mathbf{a}) = a^\mu \alpha_\mu \ , \tag{2.102}$$

which is called a *transvection* of a vector $\mathbf{a}$ with a covector $\tilde{\alpha}$.

Components a^μ of the vector $\mathbf{a} = a^\mu \boldsymbol{e}_\mu$ are real numbers. Hence, they can be also obtained by taking a value of some covector $\tilde{\omega}^\mu$ on vector $\mathbf{a}$,

$$\tilde{\omega}^\mu(\mathbf{a}) = a^\mu \ . \tag{2.103}$$

Making use of this definition, one can write equation (2.102) as follows

$$\tilde{\alpha}(\mathbf{a}) = a^\mu \alpha_\mu = \alpha_\mu \tilde{\omega}^\mu(\mathbf{a}) \ , \tag{2.104}$$

where α_μ is given by equation (2.101) and is independent from the choice of the argument $\mathbf{a}$, which means that for any covector one has

$$\tilde{\alpha} = \alpha_\mu \tilde{\omega}^\mu \ . \tag{2.105}$$

One concludes that covectors $\tilde{\omega}^\mu$ play the role of a basis in the covector space in the sense that any covector can be expanded with respect to this basis as shown in the above equation.

The basis covectors $\tilde{\omega}^\mu$ and vectors $\boldsymbol{e}_\nu$ are complementary to each other in the sense that [16)]

$$\tilde{\omega}^\mu(\boldsymbol{e}_\nu) = \delta^\mu_\nu \ . \tag{2.106}$$

This algebraic equation has a non-degenerate determinant and can be uniquely solved for $\tilde{\omega}^\mu$ as soon as *all* basis vectors $\boldsymbol{e}_\nu$ are defined. Opposite statement is also valid. For this reason, the two bases, $\tilde{\omega}^\mu$ and $\boldsymbol{e}_\nu$, are called *dual*. Due to this duality of vectors and covectors one can also define vectors as linear functions on covectors.

16) Equation (2.106) does not imply that the covector basis is orthogonal and the basis covectors $\tilde{\omega}^\mu$ have a unit norm.

2.4.2.3 *Duality of covectors and vectors*

The covector space exists independently of the vector space. The only connection between the two spaces is established by the law of correspondence between the dual bases. However, if the metric tensor is given on the manifold, it can be used in order to establish a relationship (isomorphism) between vectors and covectors. Let us show how to build the covector space from the vector space with the help of the metric tensor.

To this end, let us fix some vector $\mathbf{a} = a^\mu \boldsymbol{e}_\mu$ and consider its dot-product with an arbitrary vector $\mathbf{b} = b^\nu \boldsymbol{e}_\nu$ treated as a variable

$$\mathbf{a} \cdot \mathbf{b} = (a^\mu \boldsymbol{e}_\mu) \cdot (b^\nu \boldsymbol{e}_\nu) = a^\mu b^\nu (\boldsymbol{e}_\mu \cdot \boldsymbol{e}_\nu) = a^\mu b^\nu g_{\mu\nu} \,, \tag{2.107}$$

where the metric tensor $g_{\mu\nu}$ was introduced in accordance with equation (2.55). Notice that because the vector basis $\boldsymbol{e}_\mu$ is chosen arbitrary, the components $g_{\mu\nu}$ of the metric tensor are not making the diagonal matrix of the Minkowski metric $\eta_{\mu\nu}$.

Now, let us define a covector $\tilde{\mathbf{a}}$, which value on an arbitrary vector $\mathbf{b}$ is defined in terms of a dot product

$$\tilde{\mathbf{a}}(\mathbf{b}) \equiv \mathbf{a} \cdot \mathbf{b} \,, \tag{2.108}$$

where vector $\mathbf{a}$ is fixed. Obviously, the operation (2.108) is a covector (linear form) as it acts on a variable vector $\mathbf{b}$ and provides a number $a^\mu b^\nu g_{\mu\nu}$ as seen from equation (2.107). This linear form can be expanded over the basis covectors $\tilde{\omega}^\mu$ defined via equation (2.106)

$$\tilde{\mathbf{a}} = a_\mu \tilde{\omega}^\mu \,, \tag{2.109}$$

where the covariant components a_μ of the linear form are not arbitrary but obtained from the contravariant components a^μ of the vector $\mathbf{a}$. More specifically,

$$a_\mu = \tilde{\mathbf{a}}(\boldsymbol{e}_\mu) = \mathbf{a} \cdot \boldsymbol{e}_\mu = a^\nu \boldsymbol{e}_\nu \cdot \boldsymbol{e}_\mu = a^\nu g_{\mu\nu} \,. \tag{2.110}$$

Thus, by introducing a dot-product in a vector space, one automatically creates a dual vector space of covectors. The existence of a metric tensor establishes a one-to-one correspondence between a vector $\mathbf{a} = a^\nu \boldsymbol{e}_\nu$ and a covector $\tilde{\mathbf{a}} = a_\mu \tilde{\omega}^\mu$. One sees that the so-called covariant components of a vector are, in fact, tightly connected with its contravariant components via the metric tensor. Opposite statement is also valid, if one knows covariant components a_μ of a linear form, the corresponding vector components a^μ can be obtained by making use of the contravariant metric tensor

$$a^\mu = g^{\mu\nu} a_\nu \,. \tag{2.111}$$

This operation of transvection of the metric tensor with components of a covector is called the operation of rising index. Corresponding operation in equation (2.110) is called the operation of lowering index.

The isomorphism between the dual spaces allows us to define the dot-product between two covectors and the norm of a covector in the manner being similar to those concepts in the vector space. More specifically, one introduces a covector basis $\tilde{\omega}^\mu$ and defines a dot-product between two basis covectors in terms of the contravariant components of the metric tensor

$$\tilde{\omega}^\mu \cdot \tilde{\omega}^\nu = g^{\mu\nu} . \tag{2.112}$$

The dot-product between two covectors, $\tilde{\mathbf{a}} = a_\mu \tilde{\omega}^\mu$ and $\tilde{\mathbf{b}} = b_\mu \tilde{\omega}^\mu$, is obtained as a transvection of the components of the two covectors with the contravariant metric tensor,

$$\tilde{\mathbf{a}} \cdot \tilde{\mathbf{b}} = (a_\mu \tilde{\omega}^\mu) \cdot (b_\nu \tilde{\omega}^\nu) = a_\mu b_\nu (\tilde{\omega}^\mu \cdot \tilde{\omega}^\nu) = a_\mu b_\nu g^{\mu\nu} . \tag{2.113}$$

Due to the duality between vectors and covectors, the dot-product between the two covectors is exactly equal to the dot-product between the dual vectors, $\mathbf{a} = a^\mu \boldsymbol{e}_\mu$ and $\mathbf{b} = b^\mu \boldsymbol{e}_\mu$, because

$$\mathbf{a} \cdot \mathbf{b} = a^\mu b^\nu g_{\mu\nu} = a^\mu b_\mu = a_\mu b_\nu g^{\mu\nu} = \tilde{\mathbf{a}} \cdot \tilde{\mathbf{b}} . \tag{2.114}$$

The covector basis is called orthonormal, if the basis covectors are subject to the condition

$$\tilde{\omega}^\mu \cdot \tilde{\omega}^\nu = \eta^{\mu\nu} , \tag{2.115}$$

where $\eta^{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$ is the contravariant Minkowski metric tensor. If one chooses an orthonormal vector basis $\boldsymbol{e}_\alpha$, the corresponding dual basis of covectors $\tilde{\omega}^\mu$ will be orthonormal due to the correspondence between the two bases established by equations (2.106) and (2.114). An opposite statement is also valid.

2.4.2.4 *The transformation law of covectors*

Equations (2.106) and (2.49) allow us to derive the law of transformation of the basis covectors. Because equation (2.106) is valid in arbitrary basis, one obtains in the primed basis

$$\tilde{\omega}^{\mu'}(\boldsymbol{e}_{\nu'}) = \delta^{\mu'}_{\nu'} . \tag{2.116}$$

Substituting equation (2.49) to this equation and taking into account that

$$\delta^{\mu'}_{\nu'} = \Lambda^{\mu'}{}_\alpha \Lambda^\alpha{}_{\nu'} , \tag{2.117}$$

one obtains that the primed and unprimed bases of the covectors are related by transformation

$$\tilde{\omega}^{\mu'} = \Lambda^{\mu'}{}_\alpha \tilde{\omega}^\alpha . \tag{2.118}$$

Transformation of the components α_μ of a covector $\tilde{\alpha} = \alpha_\mu \tilde{\omega}^\mu$ is obtained from the condition that the covector is a covariant geometric object that does not depend on the choice of basis, hence,

$$\tilde{\alpha} = \alpha_\mu \tilde{\omega}^\mu = \alpha_{\mu'} \tilde{\omega}^{\mu'} . \tag{2.119}$$

Making use of equation (2.118) yields the law of transformation between the covector components,

$$\alpha_{\mu'} = \Lambda^{\nu}{}_{\mu'}\alpha_{\nu} \ . \tag{2.120}$$

This law of transformation is the same as for the basis vectors $\boldsymbol{e}_\alpha$ as shown by comparison with equation (2.49).

2.4.3
Bilinear forms

As a next step, let us consider a linear function $\mathbb{B}$ of two vectorial arguments such that its value on these arguments is

$$\mathbb{B}(\mathbf{a},\mathbf{b}) = \mathbb{B}(a^\mu \boldsymbol{e}_\mu, b^\nu \boldsymbol{e}_\nu) = a^\mu b^\nu \mathbb{B}(\boldsymbol{e}_\mu, \boldsymbol{e}_\nu) \ . \tag{2.121}$$

This function is called bilinear form or covariant tensor of the second rank because it has two vector arguments. Just as in the above case of linear forms, one introduces the basis bilinear forms formed as a tensor product of two basis covectors, $\tilde{\omega}^\mu \otimes \tilde{\omega}^\nu$, where the sign $\otimes$ denotes the tensor product. The bilinear basis form acts on two vectors $\mathbf{a}$, $\mathbf{b}$ in accordance with definition

$$\left[\tilde{\omega}^\mu \otimes \tilde{\omega}^\nu\right](\mathbf{a},\mathbf{b}) \equiv \tilde{\omega}^\mu(\mathbf{a})\ \tilde{\omega}^\nu(\mathbf{b}) = a^\mu b^\nu \ , \tag{2.122}$$

that is a matrix composed of the products of components of the two vectors. One has to emphasize that not every bilinear form can be represented as a tensor product of two covectors.

We define covariant components, $B_{\mu\nu}$, of an arbitrary bilinear form $\mathbb{B}$ as follows

$$B_{\mu\nu} \equiv \mathbb{B}(\boldsymbol{e}_\mu, \boldsymbol{e}_\nu) \ , \tag{2.123}$$

which shows how the form acts on the basis vectors producing the numbers $B_{\mu\nu}$. Equations (2.122) and (2.123) allow one to rewrite equation (2.121) as

$$\mathbb{B}(\mathbf{a},\mathbf{b}) = a^\mu b^\nu B_{\mu\nu} = B_{\mu\nu}\left[\tilde{\omega}^\mu \otimes \tilde{\omega}^\nu\right](\mathbf{a},\mathbf{b}) \ . \tag{2.124}$$

Since vectors $\mathbf{a}$ and $\mathbf{b}$ are arbitrary, they can be dropped out of the argument of the bilinear form yielding

$$\mathbb{B} = B_{\mu\nu}\tilde{\omega}^\mu \otimes \tilde{\omega}^\nu \ . \tag{2.125}$$

It demonstrates that the tensor products, $\tilde{\omega}^\mu \otimes \tilde{\omega}^\nu$, of the covectors indeed constitute a basis in the space of all possible bilinear forms $\mathbb{B}$.

A dot-product of two vectors is an example of a bilinear form $\mathbb{G}$ that is the metric tensor (see section 2.3.4). Indeed, it acts on two vectors $\mathbf{a}$, $\mathbf{b}$ to generate a number. If these two vectors are the basis vectors, the value of the bilinear form on them yields the components of the metric tensor,

$$g_{\mu\nu} = \mathbb{G}(\boldsymbol{e}_\mu, \boldsymbol{e}_\nu) \equiv \boldsymbol{e}_\mu \cdot \boldsymbol{e}_\nu \ . \tag{2.126}$$

This way, a metric tensor $\mathbb{G}$ can always be identified with a bilinear form

$$\mathbb{G} = g_{\mu\nu}\tilde{\omega}^{\mu} \otimes \tilde{\omega}^{\nu} , \tag{2.127}$$

the expansion valid in an arbitrary basis.

2.4.4
Tensors

2.4.4.1 *Definition of tensors as linear mappings*

Continuing this process, one can construct a covariant tensor $\mathbb{T}$ of an arbitrary rank n as a multi-linear function mapping n vectors $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$ to a number. This number is obtained in an arbitrary chosen basis as a result of the following definition

$$\mathbb{T}(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n) = T_{\nu_1\nu_2...\nu_n} a_1^{\nu_1} a_2^{\nu_2} ... a_n^{\nu_n} , \tag{2.128}$$

where the values of indices $i = 1, 2, ..., n$, the values of indices $\nu_i = 0, 1, 2, 3$, $a_i^{\nu_i}$ denote the ν_i-th component of the i-th vector $\mathbf{a}_i$, and $T_{\nu_1\nu_2...\nu_n}$ are the components of the covariant tensor. The basis form in the linear space of covariant tensors of rank n is a tensor product of n basis covectors and is denoted as $\tilde{\omega}^{\nu_1} \otimes ... \otimes \tilde{\omega}^{\nu_n}$. The basis form is defined through equation

$$\left[\tilde{\omega}^{\nu_1} \otimes ... \otimes \tilde{\omega}^{\nu_n}\right](\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n) \equiv \tilde{\omega}^{\nu_1}(\mathbf{a}_1)\tilde{\omega}^{\nu_2}(\mathbf{a}_2)...\tilde{\omega}^{\nu_n}(\mathbf{a}_n) = a_1^{\nu_1} a_2^{\nu_2} ... a_n^{\nu_n} . \tag{2.129}$$

Similarly, one can build a contravariant tensor $\mathbb{H}$ of an arbitrary rank m as a multi-linear function mapping m covectors $\tilde{\alpha}^1, \tilde{\alpha}^2, ..., \tilde{\alpha}^m$ to a number

$$\mathbb{H}(\tilde{\alpha}^1, \tilde{\alpha}^2, ..., \tilde{\alpha}^m) = H^{\mu_1\mu_2...\mu_m} \tilde{\alpha}^1_{\mu_1} \tilde{\alpha}^2_{\mu_2} ... \tilde{\alpha}^m_{\mu_m} , \tag{2.130}$$

where the values of indices $p = 1, 2, ..., m$, the values of indices $\mu_p = 0, 1, 2, 3$, $\tilde{\alpha}^p_{\mu_p}$ signifies the μ_p-th component of the p-th covector $\tilde{\alpha}^p$, and $T^{\mu_1\mu_2...\mu_m}$ are components of the contravariant tensor. The basis tensors in the linear space of contravariant tensors of rank m is a tensor product of m basis vectors and is denoted as $\boldsymbol{e}_{\mu_1} \otimes ... \otimes \boldsymbol{e}_{\mu_m}$. It is defined as

$$\left[\boldsymbol{e}_{\mu_1} \otimes ... \otimes \boldsymbol{e}_{\mu_m}\right](\tilde{\alpha}^1, \tilde{\alpha}^2, ..., \tilde{\alpha}^m) \equiv \tilde{\alpha}^1(\boldsymbol{e}_{\mu_1})...\tilde{\alpha}^m(\boldsymbol{e}_{\mu_m}) = \alpha^1_{\mu_1} ... \alpha^m_{\mu_m} . \tag{2.131}$$

This process can be continued further on, into the realm of multi-linear functions, which act on both vectors and covectors and produce numbers. Such mappings are called mixed tensors. An m times contravariant and n times covariant mixed tensor $\mathbb{W}$ is called the tensor of type (m, n). The rank of such tensor is $m + n$. It maps m forms and n vectors to a number according to the rule

$$\mathbb{W}(\tilde{\alpha}^1, \tilde{\alpha}^2, ..., \tilde{\alpha}^m; \mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n) = W^{\mu_1\mu_2...\mu_m}_{\nu_1\nu_2...\nu_n} \alpha^1_{\mu_1} \alpha^2_{\mu_2} ... \alpha^m_{\mu_m} a_1^{\nu_1} a_2^{\nu_2} ... a_n^{\nu_n} , \tag{2.132}$$

where $W^{\mu_1\mu_2...\mu_m}_{\nu_1\nu_2...\nu_n}$ are components of the tensor with respect to the basis chosen. Evidently, the basis tensors $\boldsymbol{e}_{\mu_1} \otimes ... \otimes \boldsymbol{e}_{\mu_m} \otimes \tilde{\omega}^{\nu_1} \otimes ... \otimes \tilde{\omega}^{\nu_n}$ defined through a relationship

$$\begin{aligned}\left[\boldsymbol{e}_{\mu_1} \otimes ... \otimes \boldsymbol{e}_{\mu_m} \otimes \tilde{\omega}^{\nu_1} \otimes ... \otimes \tilde{\omega}^{\nu_n}\right](\tilde{\alpha}^1, \tilde{\alpha}^2, ..., \tilde{\alpha}^m; \mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n) \quad &\equiv \\ \boldsymbol{e}_{\mu_1}(\tilde{\alpha}^1)...\boldsymbol{e}_{\mu_m}(\tilde{\alpha}^m)\tilde{\omega}^{\nu_1}(\mathbf{a}^1)...\tilde{\omega}^{\nu_n}(\mathbf{a}^n) = \alpha^1_{\mu_1} ... \alpha^m_{\mu_m} a_1^{\nu_1} ... a_n^{\nu_n} \quad &,\end{aligned} \tag{2.133}$$

constitute a basis in the linear space of such tensors in the sense that any tensor $\mathbb{W}$ of type (m, n) can be represented as a linear combination

$$\mathbb{W} = W^{\mu_1\mu_2\ldots\mu_m}_{\nu_1\nu_2\ldots\nu_n} \boldsymbol{e}_{\mu_1} \otimes \ldots \otimes \boldsymbol{e}_{\mu_m} \otimes \tilde{\boldsymbol{\omega}}^{\nu_1} \otimes \ldots \otimes \tilde{\boldsymbol{\omega}}^{\nu_n} \ . \tag{2.134}$$

2.4.4.2 *Transformations of tensors under a change of the basis*

Tensor of any type is considered in mathematical physics as a geometric object which does not depend on the choice of basis, since the tensor describes measurable properties either some physical object or field which exists independently of any mathematical scaffolding. This idea of the invariance of tensors with respect to the choice of the basis leads to a natural derivation of the law of transformation of tensor components. Let us consider a tensor $\mathbb{W}$ that is m times contravariant and n times covariant, that is (m, n) type. This tensor must not change when one transforms the basis in accordance with equations (2.49) and (2.118), that is irrespectively of the basis the following equations always hold

$$\begin{aligned} \mathbb{W} &= W^{\mu_1\mu_2\ldots\mu_m}_{\nu_1\nu_2\ldots\nu_n} \boldsymbol{e}_{\mu_1} \otimes \ldots \otimes \boldsymbol{e}_{\mu_m} \otimes \tilde{\boldsymbol{\omega}}^{\nu_1} \otimes \ldots \otimes \tilde{\boldsymbol{\omega}}^{\nu_n} \\ &= W^{\mu'_1\mu'_2\ldots\mu'_m}_{\nu'_1\nu'_2\ldots\nu'_n} \boldsymbol{e}_{\mu'_1} \otimes \ldots \otimes \boldsymbol{e}_{\mu'_m} \otimes \tilde{\boldsymbol{\omega}}^{\nu'_1} \otimes \ldots \otimes \tilde{\boldsymbol{\omega}}^{\nu'_n} \ . \end{aligned} \tag{2.135}$$

The only way to satisfy this equation is to have the components of the tensor $\mathbb{W}$ changing in accordance with the rule

$$W^{\mu'_1\mu'_2\ldots\mu'_m}_{\nu'_1\nu'_2\ldots\nu'_n} = W^{\alpha_1\alpha_2\ldots\alpha_m}_{\beta_1\beta_2\ldots\beta_n} \Lambda^{\mu'_1}_{\alpha_1} \Lambda^{\mu'_2}_{\alpha_2} \ldots \Lambda^{\mu'_m}_{\alpha_m} \Lambda^{\beta_1}_{\nu'_1} \Lambda^{\beta_2}_{\nu'_2} \ldots \Lambda^{\beta_n}_{\nu'_n} \ . \tag{2.136}$$

In what follows, coordinate bases will play an important role in calculations. In case of arbitrary coordinates x^α the basis vector $\boldsymbol{e}_\alpha$ is identified with a partial derivative along the coordinate axis x^α, and the basis covector is a differential taken along the same axis. More precisely,

$$\boldsymbol{e}_\alpha = \frac{\partial}{\partial x^\alpha} \ , \qquad \tilde{\omega}^\alpha = dx^\alpha \ . \tag{2.137}$$

If one introduces a new coordinate chart $x^{\alpha'}$ connected to the old coordinates by a coordinate transformation $x^{\alpha'} = x^{\alpha'}(x^\beta)$, one has the direct and inverse transformation matrices defined by equations

$$\Lambda^{\mu'}_{\alpha} = \frac{\partial x^{\mu'}}{\partial x^\alpha} \ , \qquad \Lambda^{\mu}_{\alpha'} = \frac{\partial x^{\mu}}{\partial x^{\alpha'}} \ . \tag{2.138}$$

and the matrices are orthogonal to each other,

$$\Lambda^{\mu'}_{\alpha} \Lambda^{\alpha}_{\nu'} = \delta^{\mu'}_{\nu'} \ , \tag{2.139}$$

which is easily checked by inspection. Then, the transformation law (2.136) assumes the form

$$W^{\mu'_1\mu'_2\ldots\mu'_m}_{\nu'_1\nu'_2\ldots\nu'_n}(x') = W^{\alpha_1\alpha_2\ldots\alpha_m}_{\beta_1\beta_2\ldots\beta_n}(x) \frac{\partial x^{\mu'_1}}{\partial x^{\alpha_1}} \frac{\partial x^{\mu'_2}}{\partial x^{\alpha_2}} \cdots \frac{\partial x^{\mu'_m}}{\partial x^{\alpha_m}} \frac{\partial x^{\beta_1}}{\partial x^{\nu'_1}} \frac{\partial x^{\beta_2}}{\partial x^{\nu'_2}} \cdots \frac{\partial x^{\beta_n}}{\partial x^{\nu'_n}} \ , \tag{2.140}$$

where the primed components $W^{\mu'_1\mu'_2\ldots\mu'_m}_{\nu'_1\nu'_2\ldots\nu'_n}(x')$ of the tensor $\mathbb{W}$ are calculated at the point of the spacetime manifold $\mathcal{M}$ having coordinates $x^{\alpha'}$ and the unprimed components $W^{\mu_1\mu_2\ldots\mu_m}_{\nu_1\nu_2\ldots\nu_n}(x')$ are calculated at the point having coordinates x^{α}. It is important to understand that the transformation of tensor under the change of bases takes place in one and the same event (one and the same point of the manifold). This point, however, has different values of coordinates.

There exists several mathematical operations which allows us to make new tensors from one already existed. Among them, one distinguishes algebraic and differential operations. Let us focus in the rest of this section on the algebraic operations.

2.4.4.3 ***Rising and lowering indices of tensors***

A new tensor can be obtained by changing type of the tensor. Specifically, covariant indices of tensors can be transformed to contravariant and vice versa. This procedure is defined with the help of the metric tensor $g_{\mu\nu}$. Let us assume that one has a tensor $\mathbb{W}$ of type (m, n) with components $W^{\mu_1\mu_2\ldots\mu_m}_{\nu_1\nu_2\ldots\nu_n}$. Then, one can lower any index by contracting it with the metric tensor $g_{\mu\nu}$. It defines a tensor $\mathbb{H}$ of another type. The proof that $\mathbb{H}$ is easy since this operation preserves the tensor law of transformation from one basis to another.

For example, if one lowers a contravariant index μ_k, the new tensor will be of type $(m-1, n+1)$, which components are defined by the rule

$$H^{\mu_1\mu_2\ldots\mu_{k-1}\mu_{k+1}\ldots\mu_m}_{\nu_1\nu_2\ldots\nu_k\alpha_k\nu_{k+1}\ldots\nu_n} \equiv g_{\alpha_k\mu_k} W^{\mu_1\mu_2\ldots\mu_k\ldots\mu_m}_{\nu_1\nu_2\ldots\nu_n} , \tag{2.141}$$

where the repeated Greek indices indicate summation from 0 to 3. This procedure can be repeated one more time in application to another contravariant index, thus, leading to a tensor of type $(m-2, n+2)$, and so on. The simplest example, is delivered by converting a vector a^μ to a covector $a_\mu = g_{\mu\nu}a^\nu$. In many cases the components of the new tensor $\mathbb{H}$ with the index risen, are denoted by the same letter as the original tensor $\mathbb{W}$. One has used in this example a different notation for components of the new tensor $\mathbb{H}$ as, strictly speaking, it belongs to the linear space being different from that in which the original tensor $\mathbb{W}$ was defined. What really matters, however, is not the letter used to denote the components but the number of covariant and contravariant indices present in the component of the tensor. The most evident example, is the covariant and contravariant metric tensors which components are denoted by $g_{\alpha\beta}$ and $g^{\alpha\beta}$ correspondingly. These tensors are clearly different in an arbitrary-chosen basis though the letter used for their notation is the same. Only in the orthonormal bases the metric tensor becomes the Minkowski metric tensor which has the covariant and contravariant components *formally* coinciding with each other. This notational subtlety should be taken into account when working with tensors of different types.

The indices of the tensor $W^{\mu_1\mu_2\ldots\mu_m}_{\nu_1\nu_2\ldots\nu_n}$ can be risen with the help of the contravariant metric tensor. For example, if one rises a covariant index ν_k, the new tensor will be of type $(m+1, n-1)$ defined by the rule

$$V^{\mu_1\mu_2\ldots\mu_k\beta_k\mu_{k+1}\ldots\mu_m}_{\nu_1\nu_2\ldots\nu_{k-1}\nu_{k+1}\ldots\nu_n} \equiv g^{\beta_k\nu_k} W^{\mu_1\mu_2\ldots\mu_m}_{\nu_1\nu_2\ldots\nu_k\ldots\nu_n} , \tag{2.142}$$

where the repeated Greek indices indicate summation from 0 to 3. The simplest example, is delivered by converting a covector a_μ to a vector $a^\mu = g^{\mu\nu}a_\nu$.

Of course, in case of the orthonormal bases the metric tensor used for rising and lowering indices is the Minkowski metric $\eta_{\mu\nu} = \mathrm{diag}(-1,1,1,1)$ or $\eta^{\mu\nu} = \mathrm{diag}(-1,1,1,1)$ that can help to clarify the difference between the covector components a_μ and its contra-variant counterpart which is important in the Minkowski spacetime and absent in the Euclidean spaces. Indeed, one can easily check that in the Minkowski spacetime, $a_0 = \eta_{0\nu}a^\nu = -a^0$, and $a_i = \eta_{i\nu}a^\nu = a^i$. In other words, the time components of vectors and covectors have opposite signs in the Minkowski spacetime.

2.4.4.4 ***Contraction of tensor indices***

Tensor $\mathbb{W}$ of type (m,n) can be converted to a new tensor $\mathbb{T}$ of type $(m-1, n-1)$ by applying operation of summation with respect to one covariant and one contravariant indices. This operation preserves the tensor law of transformation and is called *transvection* or contraction with respect to a pair of indices. Transvection works for any pair of indices of different types. However, it is not correct to apply it for a pair of indices of the same type. For example, one can define a new tensor by contracting a covariant index ν_n with a contravariant index μ_n of the tensor $\mathbb{W}$. In this case, both indices participating in the operation of transvection, are denoted by the same letter, so that one has

$$T^{\mu_1\mu_2\ldots\mu_{m-1}}_{\nu_1\nu_2\ldots\nu_{n-1}} \equiv W^{\mu_1\mu_2\ldots\mu_{m-1}\mu_n}_{\nu_1\nu_2\ldots\nu_{n-1}\mu_n} , \tag{2.143}$$

where one has contracted the last covariant index with the last contravariant index of the tensor $\mathbb{W}$ and used the letter μ_n to denote the contracted indices. Because transvection is a tensor operation, the particular notation for the contracted indices does not play any role - such indices are called *dummy indices*.

If the metric tensor is defined on the manifold, the operation of contraction of indices can be also applied to a pair of covariant or contravariant indices of the tensor $\mathbb{W}$. In this case the tensor changes its type from (m,n) to $(m, n-2)$, or from the type (m,n) to $(m-2, n)$ correspondingly. For example, by applying the contravariant metric tensor one can contract two last sub-indices of the tensor $\mathbb{W}$, thus, obtaining another tensor $\mathbb{S}$ having components defined by

$$S^{\mu_1\mu_2\ldots\mu_{m-1}\mu_n}_{\nu_1\nu_2\ldots\nu_{n-2}} \equiv g^{\nu_{n-1}\nu_n} W^{\mu_1\mu_2\ldots\mu_{m-1}\mu_n}_{\nu_1\nu_2\ldots\nu_{n-1}\nu_n} . \tag{2.144}$$

This operation can be applied as many times as required in a particular problem.

2.4.4.5 ***Tensor equations***

All laws of fundamental physics are formulated in the form of tensor equations. A tensor equation equates one tensor, let say $\mathbb{A}$, to another, denoted as $\mathbb{B}$, and looks like

$$\mathbb{A} = \mathbb{B} . \tag{2.145}$$

Any tensor equation must have the following properties:

1) Both tensors, $\mathbb{A}$ and $\mathbb{B}$, must be of the same type.
2) Both tensors, $\mathbb{A}$ and $\mathbb{B}$, must be taken at the same event (point) of the spacetime manifold.

If these two conditions are satisfied, the tensor equation (2.145) holds in any basis (reference frame) and coordinate chart. As a consequence, a tensor being equal to zero [17] in one frame must be equal to zero in any other frame. This follows directly from the transformation law (2.140).

Significant part of tensor equations include the operation of differentiation of tensors. A partial derivative from a tensor along a basis vector $\boldsymbol{e}_\alpha$ is denoted with ∂_α. In a coordinate basis, the vector $\boldsymbol{e}_a = \partial/\partial x^\alpha$, is directed along x^α axis. In this case,

$$\partial_\alpha \equiv \frac{\partial}{\partial x^\alpha} \,. \tag{2.146}$$

In the Minkowski spacetime and in the inertial coordinates, the partial derivatives behave as tensors. In other words, the partial derivatives in two vector bases, $\boldsymbol{e}_\alpha$ and $\boldsymbol{e}_{\alpha'}$, are related by the linear transformation

$$\partial_{\alpha'} = \Lambda^\beta{}_{\alpha'}\partial_\beta \,, \qquad \partial_\alpha = \Lambda^{\beta'}{}_\alpha \partial_{\beta'} \,. \tag{2.147}$$

Hence, a partial derivative, ∂_α, makes from a tensor $\mathbb{W}$ of type (n, m) another tensor of type $(m, n+1)$

$$W^{\mu_1\mu_2\ldots\mu_m}_{\nu_1\nu_2\ldots\nu_n} \Longrightarrow \partial_\alpha W^{\mu_1\mu_2\ldots\mu_m}_{\nu_1\nu_2\ldots\nu_n} \,. \tag{2.148}$$

Index of a partial derivative in the Minkowski spacetime can be risen with the Minkowski metric tensor $\eta^{\mu\nu}$

$$\partial^\alpha \equiv \eta^{\alpha\beta}\partial_\beta \,. \tag{2.149}$$

A multiple partial derivative of l-th order is denoted with a symbol

$$\partial_{\alpha_1\alpha_2\ldots\alpha_l} \equiv \partial_{\alpha_1}\partial_{\alpha_2}\ldots\partial_{\alpha_l} \,. \tag{2.150}$$

Sometimes, a comma is used with a coordinate index following it, in order to denote a partial derivative so that two notations for the partial derivative are equivalent

$$\partial_\alpha W^{\mu_1\mu_2\ldots\mu_m}_{\nu_1\nu_2\ldots\nu_n} \equiv W^{\mu_1\mu_2\ldots\mu_m}_{\nu_1\nu_2\ldots\nu_n,\alpha} \,. \tag{2.151}$$

Partial derivative from the tensor product of two tensors, $\mathbb{A} \otimes \mathbb{B}$, obeys the Leibnitz chain rule

$$\partial_a\,(\mathbb{A} \otimes \mathbb{B}) = (\partial_a\mathbb{A}) \otimes \mathbb{B} + \mathbb{A} \otimes (\partial_a\mathbb{B}) \,. \tag{2.152}$$

These rules of operation with the partial derivatives must be extended to keep the tensor character of physical equations in the case of general relativity when the gravity plays an essential role.

17) It means that all components of the tensor in the given basis are zero.

2.5 Kinematics

This section describes kinematic properties of particles moving in the Minkowski spacetime with arbitrary velocities with the only limitation that they are less than or equal to the fundamental speed c. Hypothetical particles like tachyon moving a priory with the speed larger than c, will not be considered. We shall use the inertial reference frames to discuss mathematical relationships between various physical quantities. Each inertial reference frame will be identified with an orthonormal basis built from the Cartesian coordinates covering the entire Minkowski spacetime. We shall denote three-dimensional vectors by *italic bold* letters, like $\boldsymbol{a} = \left(a^i\right) = \left(a^1, a^2, a^3\right)$, $\boldsymbol{b} = \left(b^i\right) = \left(b^1, b^2, b^3\right)$. Three-dimensional dot product between two spatial vectors, for example, $\boldsymbol{a}$ and $\boldsymbol{b}$, will be denoted as $\boldsymbol{a} \cdot \boldsymbol{b} = \delta_{ij} a^i b^j = a^1 b^1 + a^2 b^2 + a^3 b^3$. Cross-product between two spatial vectors will be denoted as $\boldsymbol{a} \times \boldsymbol{b} = (\boldsymbol{a} \times \boldsymbol{b})^i = \varepsilon_{ijk} a^i b^j$, where ε_{ijk} is the fully anti-symmetric symbol of Levi-Civita, such that $\varepsilon_{123} = +1$, and $\varepsilon_{ijk} = 0$, if $i \neq j \neq k$.

2.5.1 The proper frame of observer

Distance between two arbitrary events A and B in the Minkowski spacetime is called interval. If the event A has coordinates $x^\alpha_A = (ct_A, x_A, y_A, z_A) = (ct_A, \boldsymbol{x}_A)$, and the event B has coordinates $x^\alpha_B = (ct_B, x_B, y_B, z_B) = (ct_B, \boldsymbol{x}_B)$, the interval between the two events is

$$s^2_{AB} = -c^2 (t_A - t_B)^2 + (\boldsymbol{x}_A - \boldsymbol{x}_B)^2 \ , \tag{2.153}$$

which can be negative positive or equal to zero depending on the position of the event B with respect to the null cone with vertex at the event A (see section 2.2.4.4 and Figure 2.3). When two events are separated by infinitesimally small coordinate increments $dx^\alpha = (dx^0, dx^1, dx^2, dx^3) = (cdt, dx, dy, dz)$ the interval

$$ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2 \ , \tag{2.154}$$

which can be also written down in index notations as

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu \ , \tag{2.155}$$

where $\eta_{\mu\nu}$ is the Minkowski metric (2.63) in the Cartesian coordinates of the inertial frame S. If one chooses another inertial frame S' covered with the Cartesian coordinates $x^{\alpha'} = (x'^0, x'^1, x'^2, x'^3) = (ct', x', y', z')$, the interval between the same events becomes

$$ds'^2 = \eta_{\mu\nu} dx'^\mu dx'^\nu = -c^2 dt'^2 + dx'^2 + dy'^2 + dz'^2 \ . \tag{2.156}$$

According to the principle of relativity the numerical value of the interval between two events does not depend on the choice of coordinates and is preserved when one goes from one inertial frame to another, that is

$$ds'^2 = ds^2 \ . \tag{2.157}$$

Each physical observer is associated with its own inertial reference frame, let say, S'. The observer is placed at the origin of the frame that is equipped with an orthonormal vector basis consisting of four vectors, $\boldsymbol{e}_{\alpha'} = (\boldsymbol{e}_{0'}, \boldsymbol{e}_{1'}, \boldsymbol{e}_{2'}, \boldsymbol{e}_{3'})$, which are the unit vectors along four axes of the Cartesian coordinates $x^{\alpha'} = (x^{0'}, x^{1'}, x^{2'}, x^{3'})$. The coordinate time $x^{0'} = ct'$ of the observer's frame coincides with the proper time τ of the observer that is measured with the help of an ideal clock. Though. the ideal clock does not exist in nature, the belief is that, in principle, one can build clocks with the rate approximating the ideal time with an arbitrary accuracy limited only by the laws of quantum mechanics. Current practical realization of the ideal clock is achieved with the atomic clocks (see Chapter 9.3 for further discussion).

Spatial coordinate axes $x^{i'} = (x^{1'}, x^{2'}, x^{3'})$ of the observer are assumed to be equipped with the ideal rigid rulers that are used to measure distances in space from observer to any other event being simultaneous (that is having the same value of the time coordinate in the observer's frame) with the event of the measurement undertaken by the observer. Since the observer does not move with respect to its own frame S', the primed spatial coordinates stays unchanged and equal to zero $x^{i'} = 0$; only the time coordinate t' of the observer changes. The proper time $\tau = t'$ of the observer is related to the interval, measured in the proper reference frame S', in the following way

$$c^2 d\tau^2 = c^2 dt'^2 = -ds'^2 \ . \tag{2.158}$$

Due to the principle of relativity and equation (2.157), the proper time τ of the observer does not depend on the choice of the frame. In any inertial frame S covered with coordinates x^α, the proper time of observer is defined by differential equation

$$c^2 d\tau^2 = -ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 \ , \tag{2.159}$$

where the spatial coordinates must be known as functions of the coordinate time t, that is $x^i = x^i(t)$, along observer's worldline. The proper time τ is defined only between two events separated by timelike interval for which $ds^2 < 0$. It can not be defined for the events separated by the null-like, $ds^2 = 0$, or spacelike intervals $ds^2 > 0$.

Spatial coordinates $x^{i'}$ of some event measured by observer in the proper reference frame S' are used to measure the proper distance ℓ from the observer to the event under consideration. Since the event must be simultaneous with the instant of measurement undertaken by the observer, the interval of the coordinate time $dt' = 0$, and the proper distance is defined by equation

$$d\ell^2 = \delta_{i'j'} dx^{i'} dx^{j'} = ds'^2 \ . \tag{2.160}$$

If two events are separated by a finite distance, this equation should be integrated to get the proper distance ℓ. In any other inertial frame S covered with coordinates

x^α, the proper distance is defined by making use of the principle of relativity and equation (2.157),

$$d\ell^2 = ds^2 = dx^2 + dy^2 + dz^2 - c^2 dt^2 \,. \tag{2.161}$$

In the new frame the proper distance ℓ must be calculated with taking into account that the two events are not simultaneous in this frame, so that $dt \neq 0$. The proper distance can be defined only for two events separated by spacelike, $ds^2 > 0$, interval.

2.5.2
Four-velocity and four-acceleration

Let us consider a particle moving in the inertial frame S along a worldline $x^\mu(t)$ which is parameterized by the coordinate time t. Three-dimensional, coordinate velocity of the particle $v^i = dx^i/dt$, and its three-dimensional coordinate acceleration $a^i = dv^i/dt$. Proper time τ along the worldline of the particle relates to the coordinate time by equation (2.159). Because three-velocity of the particle is well-defined, the proper time of the particle becomes

$$d\tau = dt\sqrt{1 - \frac{\boldsymbol{v}^2}{c^2}} \,, \tag{2.162}$$

where $\boldsymbol{v} = (v^i)$.

Four-velocity $u^\mu = (u^0, u^i)$ of the particle is defined as

$$u^\mu \equiv \frac{dx^\mu}{d\tau} = \left(c\frac{dt}{d\tau}, \frac{dx^i}{d\tau}\right) = c\gamma_{\rm v}\left(1, \beta^i_{\rm v}\right) \,, \tag{2.163}$$

where $\beta^i_{\rm v} = v^i/c$, and $\gamma_{\rm v} = dt/d\tau = (1 - \boldsymbol{\beta}^2_{\rm v})^{1/2}$ is the Lorentz factor of the particle corresponding to the Lorentz boosts from the static frame S to the moving proper frame of the particle, $\boldsymbol{\beta}_{\rm v} = (\beta^i_{\rm v})$. The proper time is invariant and the Cartesian coordinates x^α transform as a four-vector under the Lorentz transformation. Hence, four-velocity makes a vector in the Minkowski spacetime that is tangent to the worldline of the particle. Since the worldline of the particle is always passing in the interior of the null cone (see Figure 2.3), the four-velocity is a timelike vector. Indeed, simple calculation shows that

$$u^\mu u_\mu = \eta_{\mu\nu} u^\mu u^\nu = -c^2 < 0 \,. \tag{2.164}$$

Four-acceleration w^μ of the particle is defined as

$$w^\mu \equiv \frac{du^\mu}{d\tau} = u^\beta \partial_\beta u^\mu \,, \tag{2.165}$$

where $\partial_\beta = \partial/\partial x^\beta$. This definition clearly shows that four-acceleration is a vector in the Minkowski spacetime. Taking into account that four-velocity of the particle depends only on time, one can rewrite previous equation in the following form

$$w^\mu = \gamma_{\rm v} \frac{du^\mu}{dt} \,, \tag{2.166}$$

that does not depend explicitly on the proper time making it more convenient for subsequent calculations.

Taking the ordinary time derivative in equation (2.166) one obtains the four-acceleration $w^\mu = (w^0, w^i)$ expressed in terms of three-dimensional coordinate acceleration, $\boldsymbol{a} = (a^i) = (a^1, a^2, a^3)$, and velocity, $\boldsymbol{v} = (v^i) = (v^1, v^2, v^3)$, of the particle

$$w^0 = \gamma_{\rm v}^4(\boldsymbol{\beta}_{\rm v} \cdot \boldsymbol{a}) \,, \tag{2.167}$$

$$w^i = \gamma_{\rm v}^2\left[a^i + \frac{\gamma_{\rm v}^2 - 1}{v^2}(\boldsymbol{v} \cdot \boldsymbol{a})v^i\right] . \tag{2.168}$$

In the rest frame of the particle, its four-velocity $u^\mu = (c, 0, 0, 0)$, while w^μ reduces to $w^\mu = (0, A^1, A^2, A^3)$, where $A^i = (A^1, A^2, A^3)$ is the proper acceleration Rindler [1960a] experienced by the particle in its rest frame [18]. Because the four-dimensional acceleration is a vector in the Minkowski spacetime, its Lorentzian norm, $w^\mu w_\mu = \eta_{\mu\nu} w^\mu w^\nu$, is a scalar that does not depend on a particular choice of the frame of reference. The most simple is to calculate it in the rest frame of the particle, which yields

$$w^\mu w_\mu = \boldsymbol{A}^2 \,. \tag{2.169}$$

Because $\boldsymbol{A}^2 > 0$, the four-acceleration is a spacelike vector. Making use of equations (2.167) and (2.168) for calculating $w_\mu w^\mu$, one derives a relationship between the measured value of the proper acceleration of the particle and the coordinate value of three-acceleration,

$$\boldsymbol{A}^2 = \gamma_{\rm v}^4\left[\boldsymbol{a}^2 + \gamma_{\rm v}^2(\boldsymbol{\beta}_{\rm v} \cdot \boldsymbol{a})^2\right] . \tag{2.170}$$

Final remark is that four-velocity and four-acceleration are orthogonal to each other. This can be seen immediately by doing calculation of the dot product, $u_\mu w^\mu$, in the rest frame of the particle. Another way around is to use the covariant definition (2.165) of four-acceleration. It yields,

$$u_\mu w^\mu = \eta_{\mu\nu} u^\mu \frac{du^\nu}{d\tau} = \frac{1}{2}\frac{d(u^\mu u_\mu)}{d\tau} = 0 \,, \tag{2.171}$$

as a consequence of equation (2.164).

2.5.3
Transformation of velocity

Let us consider two inertial reference frame, S and S' covered with Cartesian coordinates $x^\mu = (x^0, x^i) = (ct, \boldsymbol{x})$ and $x^{\mu'} = (x^{0'}, x^{i'}) = (ct', \boldsymbol{x}')$ respectively. Let frame S' move uniformly with respect to S with a constant velocity $\boldsymbol{V} = (V^i)$. Let a particle move with a three-dimensional velocity $\boldsymbol{v} = (v^i) = dx^i/dt$ with respect to frame S, and with that $\boldsymbol{v}' = (v^{i'}) = dx^{i'}/dt'$ with respect to S'. One is looking for the relationship connecting velocities $\boldsymbol{v}$ and $\boldsymbol{v}'$ of the particle in the two frames. The most

18) The rest-frame acceleration $\boldsymbol{A} = (A^i)$ of the particle can be measured by accelerometer. In gravimetric measurements the accelerometer is called the absolute gravimeter (see Chapter 8).

straightforward way to derive this relationship is to apply the transformation law of the four-velocity of the particle

$$u^{\mu'} = \Lambda^{\mu'}{}_{\nu} u^{\nu} , \tag{2.172}$$

where the matrix $\Lambda^{\mu'}{}_{\nu}$ of the Lorentz boost is given in equation (2.82). The Lorentz transformation between the three-velocities, $v^{i'} = u^{i'}/u^{0'}$ and $v^i = u^i/u^0$, is derived with the help of equation (2.172) that yields

$$v^{i'} = \frac{\Lambda^{i'}{}_0 + \Lambda^{i'}{}_j v^j}{\Lambda^{0'}{}_0 + \Lambda^{0'}{}_j v^j} . \tag{2.173}$$

Substituting the matrices of the Lorentz boost, given by expressions (2.82), in this equation, one arrives at

$$\boldsymbol{v}' = \frac{\boldsymbol{v} - \gamma \boldsymbol{V} + (\gamma - 1)V^{-2}(\boldsymbol{V} \cdot \boldsymbol{v})\boldsymbol{V}}{\gamma(1 - c^{-2}\boldsymbol{v} \cdot \boldsymbol{V})} , \tag{2.174}$$

where $\gamma = (1 - V^2/c^2)^{1/2}$. Projecting this equation on the direction of motion of frame S' with respect to S, one obtains

$$1 + \frac{\boldsymbol{v}' \cdot \boldsymbol{V}}{c^2} = \frac{1 - \dfrac{V^2}{c^2}}{1 - \dfrac{\boldsymbol{v} \cdot \boldsymbol{V}}{c^2}} . \tag{2.175}$$

This equation tells us that in the limit of ultra-relativistic motion, when $V \to c$, the velocity $\boldsymbol{v}'$ of the particle in the frame S' is almost opposite to the velocity $\boldsymbol{V}$. If one had many particles moving chaotically with arbitrary directed velocities in the frame S, their motions in the frame S' moving with ultra-relativistic speed with respect to frame S, would be strongly collimated along the direction of vector $\boldsymbol{V}$.

Formula (2.174) describes the relative velocity of the particle with respect to a moving frame S', and it was pioneered by Poincaré (1906). Inverse transformation yields the relativistic law of addition of velocities

$$\boldsymbol{v} = \frac{\boldsymbol{v}' + \gamma \boldsymbol{V} + (\gamma - 1)V^{-2}(\boldsymbol{V} \cdot \boldsymbol{v}')\boldsymbol{V}}{\gamma(1 + c^{-2}\boldsymbol{v}' \cdot \boldsymbol{V})} . \tag{2.176}$$

As can be seen the inverse transformation can be obtained from the direct transformation (2.174) by replacing $\boldsymbol{V} \to -\boldsymbol{V}$, and $\boldsymbol{v} \to -\boldsymbol{v}'$. This can be proved by doing direct calculations with the Lorentz transformation matrices, or by making use of the principle of relativity which states that all frames are equivalent and, hence, motion of particles must be described by the same equations with the corresponding accounting for the direction of the relative velocity between the two inertial frames.

In case of one-dimensional motion, when velocities $\boldsymbol{v}'$ and $\boldsymbol{V}$ are parallel and have the same direction, equation (2.176) is reduced to a more simple form,

$$v = \frac{v' + V}{1 + c^{-2}v'V} . \tag{2.177}$$

This equation tells us that the fundamental speed c is indeed the ultimate speed of motion of particles existing in nature. Indeed, one can easily confirm that in case of the asymptotic approaching of the relative velocity v' of the particle to the speed c, equation (2.177) tells us that the particle's speed v with respect to the static frame, also approaches c asymptotically from below but can never reach nor exceed it. Particles like photons, move with the speed c in any frame irrespectively of the speed V of relative motion between the frames. Equation (**??**) also tells us that the physical speed of light does not depend on the speed of its source, but only if the source *moves in vacuum*. If the source of light moves through a transparent medium, where the physical speed of light is less than the fundamental speed c, the speed of light observed by a static observer depends on the motion of the source. This can be easily seen from equation (2.177) where one has to consider the source of light moving with the speed V, and the speed of light emitted by the source in its rest frame is $v' = c' < c$. The speed of light v measured in the static frame depends on the speed V of the source of light.

The absolute magnitude of the relative velocity $\boldsymbol{v}'$ is obtained by the squaring equation (2.176) which results in

$$\boldsymbol{v}'^2 = \frac{(\boldsymbol{v}-\boldsymbol{V})^2 - c^{-2}\,(\boldsymbol{v}\times\boldsymbol{V})^2}{(1-c^{-2}\boldsymbol{v}\cdot\boldsymbol{V})^2}\,, \tag{2.178}$$

demonstrating that velocities $\boldsymbol{v}$ and $\boldsymbol{V}$ enter this equation symmetrically. Another useful relationship between the Lorentz factors of the moving particle in different frames is

$$\frac{1-\dfrac{\boldsymbol{v}'^2}{c^2}}{1-\dfrac{\boldsymbol{v}^2}{c^2}} = \frac{1-\dfrac{\boldsymbol{V}^2}{c^2}}{\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{V}}{c^2}\right)^2}\,, \tag{2.179}$$

which is a consequence of equation (2.178). It is instructive to compare equation (2.179) with equation (2.175) to see various transformation symmetries existing between velocities $\boldsymbol{v}$, $\boldsymbol{V}$ and $\boldsymbol{v}'$.

Let us explore the transformation property of the relative velocity. To this end, let us make a Lorentz transformation from the inertial frame S to a frame S'' moving with respect to S with velocity $\boldsymbol{W}$. The particle's velocity $\boldsymbol{v}$ will change to $\boldsymbol{v}''$, and velocity $\boldsymbol{V}$ will transform to $\boldsymbol{V}''$ in accordance with the Lorentz transformations

$$\boldsymbol{v}'' = \frac{\boldsymbol{v} - \Gamma\boldsymbol{W} + (\Gamma-1)W^{-2}(\boldsymbol{W}\cdot\boldsymbol{v})\boldsymbol{W}}{\Gamma\,(1-c^{-2}\boldsymbol{v}\cdot\boldsymbol{W})}\,, \tag{2.180}$$

$$\boldsymbol{V}'' = \frac{\boldsymbol{V} - \Gamma\boldsymbol{W} + (\Gamma-1)W^{-2}(\boldsymbol{W}\cdot\boldsymbol{V})\boldsymbol{W}}{\Gamma\,(1-c^{-2}\boldsymbol{V}\cdot\boldsymbol{W})}\,, \tag{2.181}$$

where the Lorentz factor $\Gamma = (1-W^2/c^2)^{1/2}$. It is tedious but straightforward to demonstrate that

$$\frac{(\boldsymbol{v}''-\boldsymbol{V}'')^2 - c^{-2}\,(\boldsymbol{v}''\times\boldsymbol{V}'')^2}{(1-c^{-2}\boldsymbol{v}''\cdot\boldsymbol{V}'')^2} = \frac{(\boldsymbol{v}-\boldsymbol{V})^2 - c^{-2}\,(\boldsymbol{v}\times\boldsymbol{V})^2}{(1-c^{-2}\boldsymbol{v}\cdot\boldsymbol{V})^2}\,, \tag{2.182}$$

which proves that the magnitude of the relative velocity given by equation (2.178) is a relativistic invariant which does not depend on the choice of the inertial frame of reference. In other words, if one has two particles moving with respect to an inertial reference frame with two different velocities, the magnitude of their relative velocity does not depend on which particular frame is used for its calculation.

If one considers two particles moving with infinitesimally close values of their velocities, say $\boldsymbol{V}$ and $\boldsymbol{v} = \boldsymbol{V} + d\boldsymbol{\xi}$, the magnitude of the relative velocity denoted as $d\zeta^2 \equiv \boldsymbol{v}'^2$ will be expressed through equation (2.178) as follows,

$$d\zeta^2 = \gamma^2 \left[d\boldsymbol{\xi}^2 + \gamma^2 (\boldsymbol{\beta} \cdot d\boldsymbol{\xi})^2 \right] , \tag{2.183}$$

where $\gamma = (1 - \beta^2)^{-1/2}$, and $\boldsymbol{\beta} = \boldsymbol{v}/c$. This expression coincides with the interval between two infinitesimally close points in the Lobachevsky space of three-dimensional hyperbolic geometry [Fock, 1964, §17] expressed in terms of coordinates $\boldsymbol{\xi} = (\xi^1, \xi^2, \xi^3)$. Hyperbolic geometry was discovered and extensively explored irrespectively of special relativity by János Bolyai and Nikolai Lobachevsky, after whom it is named [Lobachevsky, 1999]. Equation (2.183) establishes a tight connection between the hyperbolic geometry and the space of relative velocities of particles in special relativity.

2.5.4
Transformation of acceleration

One is looking for the relationship connecting accelerations $\boldsymbol{a} = d\boldsymbol{v}/dt$ and $\boldsymbol{a}' = d\boldsymbol{v}'/dt'$ of a particle in two reference frames, S and S' introduced in the previous subsection. The most straightforward way to derive this relationship is to differentiate the transformation law (2.174) of the three-velocity of the particle by making use of the chain rule

$$\boldsymbol{a}' = \frac{d\boldsymbol{v}'}{dt}\frac{dt}{dt'} . \tag{2.184}$$

Time derivative

$$\frac{dt}{dt'} = \frac{u^0}{u^{0'}} = \frac{u^0}{\Lambda^{0'}{}_0 u^0 + \Lambda^{0'}{}_j u^j} = \frac{1}{\gamma\left(1 - c^{-2}\boldsymbol{v} \cdot \boldsymbol{V}\right)} , \tag{2.185}$$

where one has used the matrix $\Lambda^{\mu'}{}_\nu$ given in equation (2.82). When taking time derivative $d\boldsymbol{v}'/dt$ one should differentiate in equation (2.174) velocity $\boldsymbol{v}$ only, because velocity $\boldsymbol{V}$ between the two frames is constant.

After completing the calculations and reducing similar terms, one obtains

$$\boldsymbol{a}' = \frac{\boldsymbol{a}}{\gamma^2 s^2} - \frac{\gamma - 1}{\gamma^3 s^3}\frac{(\boldsymbol{a} \cdot \boldsymbol{V})\boldsymbol{V}}{V^2} + \frac{(\boldsymbol{a} \cdot \boldsymbol{V})\boldsymbol{v}}{\gamma^2 s^3 c^2} , \tag{2.186}$$

where one has introduced notation

$$s \equiv 1 - \frac{\boldsymbol{v} \cdot \boldsymbol{V}}{c^2} . \tag{2.187}$$

Formula for the inverse transformation of the acceleration is obtained from equations (2.186), (2.187) by the replacement: $\boldsymbol{v} \to \boldsymbol{v}'$, $\boldsymbol{a} \to \boldsymbol{a}'$, $\boldsymbol{V} \to -\boldsymbol{V}$.

It is clear from equation (2.186) that if the particle moves with a constant acceleration $\boldsymbol{a}$ in one frame, then the acceleration $\boldsymbol{a}'$ in another reference frame will, in general, be a function of time. This happens because the velocity $\boldsymbol{v}$ entering equation (2.186) is not constant in the presence of the acceleration. This conclusion is drastically different from that in the Newtonian mechanics where acceleration is invariant with respect to the Galilean transformations.

Formula (2.186) allows us to establish a relationship between the proper acceleration, $\boldsymbol{A}$, experienced by the particle in its rest frame, and the coordinate acceleration, $\boldsymbol{a}$, of the particle in the laboratory frame S. In the rest frame of the particle its instantaneous velocity $\boldsymbol{v}' = 0$. Equation (2.176) yields, then, $\boldsymbol{v} = \boldsymbol{V}$. Substituting this result to equation (2.186) gives for the instantaneous acceleration, $\boldsymbol{A}$, of the particle

$$\boldsymbol{A} = \gamma_{\rm v}^2 \left[\boldsymbol{a} + \frac{\gamma_{\rm v} - 1}{v^2} (\boldsymbol{v} \cdot \boldsymbol{a}) \boldsymbol{v} \right] , \tag{2.188}$$

where $\gamma_{\rm v} = (1 - v^2/c^2)^{-1/2}$. This equation should not be confused with equation (2.168) which defines the spatial components $\boldsymbol{w} = (w^i)$ of four-acceleration in terms of acceleration $\boldsymbol{a}$. The components $\boldsymbol{w}$ do not coincide with the instantaneous acceleration $\boldsymbol{A}$ in any other frame but the rest frame of the particle. Indeed, since four-acceleration is a four-vector, its spatial components in the frame S are transformed to the rest frame as $w^i = \Lambda^i{}_{\alpha'} w^{\alpha'} = \Lambda^i{}_{j'} A^{j'}$, or in three dimensional notations

$$\boldsymbol{w} = \boldsymbol{A} + \frac{\gamma_{\rm v} - 1}{v^2} (\boldsymbol{v} \cdot \boldsymbol{A}) \boldsymbol{v} , \tag{2.189}$$

which makes clear that $\boldsymbol{w} \neq \boldsymbol{A}$ except the case when $\boldsymbol{v} = 0$.

Squaring equation (2.188) gives again equation (2.169) as expected. If velocity $\boldsymbol{v}$ of the particle is parallel to its coordinate acceleration $\boldsymbol{a}$ in frame S, it varies in magnitude only,

$$\boldsymbol{A} = \gamma_{\rm v}^3 \boldsymbol{a} , \tag{2.190}$$

where $\gamma_{\rm v}$ is a function of time that can be determined by integration of equation (2.190) if the proper acceleration $\boldsymbol{A}$ is known. On the other hand, if the velocity of the particle varies only in direction, that is $\boldsymbol{v}$ is orthogonal to $\boldsymbol{a}$, then $\boldsymbol{v} \cdot \boldsymbol{a} = 0$, and

$$\boldsymbol{A} = \gamma_{\rm v}^2 \boldsymbol{a} , \tag{2.191}$$

with $\gamma_{\rm v}$ being a constant Lorentz factor. These equations are used in the physics of particle's accelerators.

2.5.5
Dilation of time

Equation (2.162) clearly indicates that in special relativity the proper time τ of observer goes slower than the coordinate time t by the factor $\gamma = \sqrt{1 - \beta^2}$. It means that the interval of time between two events separated by a timelike vector and measured by a static observer will be always longer than that measured by a moving observer. This statement is illustrated in Figure 2.5 where one event is denoted by letter A and another event is taken at the origin of two inertial frames, which is the same for both coordinates, that is the events with coordinates ($t' = 0, x' = 0$) and ($t = 0, x = 0$), are identical. The event A is lying on the worldline of the moving observer. Figure 2.5 demonstrates that the time interval to the event A from the origin of the coordinates is longer for a static observer, who measures time t, than for a moving observer, who measures time t'. This statement can be also confirmed by a simple calculation. Interval of time between the origin and the event A in static frame S is t_A. The interval of time between the origin and the same event A in moving frame S', has the following coordinates

$$ct'_A = \gamma(ct_A - \beta x_A) \,, \tag{2.192}$$

$$x'_A = \gamma(x_A - \beta ct_A) \,, \tag{2.193}$$

where one has used the Lorentz transformation (2.80) with $\beta = V/c$. Spatial coordinate x_A of the event A relates to time t_A by the straight worldline of the origin of the frame S', that is $x_A = Vt_A$, which is a consequence of the condition $x'_A = 0$, and equation (2.193). Substituting $x_A = Vt_A$ to the time transformation (2.192) yields

$$t'_A = t_A\sqrt{1-\beta^2} < t_A \,. \tag{2.194}$$

One sees that time interval t'_A measured by a moving observer is shorter than time t_A measured by the static observer by the factor of $\sqrt{1-\beta^2} < 1$, an effect called time dilation. It is due to this effect that one can study many species of elementary particles produced in colliders or in cosmic-ray showers. Many of the particles live extremely short time in their rest frame, but when they are moving with respect to us with ultra-relativistic velocity, the time of their decay in the laboratory frame of terrestrial observer grows proportionally to the Lorentz factor γ, in accordance with equation (2.194). Hence, the lifespan of the particle elongates.

2.5.6
Simultaneity and synchronization of clocks

Simultaneity of two events can be established by a static observer by sending light signals to the events equipped with mirrors. In real practice, the mirrors can be replaced with radio transponders or other equivalent technique. Having reflected from mirrors, the signals return to the observer. If these signals are emitted at one instant of time, and return to the observer without time lag between them, the observer says that the events of reflection from the mirrors are simultaneous. It is important, though, to realize that the same two events would *not* be simultaneous to another observer moving with respect to the first one along the line between the two events. This experiment is similar to Einstein's train thought (Gedanken) experiment. Thus, the concept of simultaneity is not absolute, and depends on the reference frame of observer. For the first time, this fact was pointed out by Einstein [1905] though Lorentz and Poincaré were aware of it but their interpretation was based on the assumption of existence of æther.

Despite the relative nature of simultaneity, special relativity does not prohibit to unambiguously synchronize a set of clocks by making use of light signals, provided the clocks are moving uniformly and light propagates in vacuum. This procedure is called Einstein's synchronization of clocks [19]. Let us consider, first, the Einstein synchronization of clocks for static observers with the first observer located at the origin of some inertial frame of reference S with clock measuring time t and spatial coordinates (x, y, z). Let the observer emit a light signal at time t_O in the positive direction of x-axis. The signal reaches another static observer at point A having coordinates $(x_A, 0, 0)$, and is immediately reflected back to the first static observer at instant of time t_A. The reflected light signal is received by the first static observer at instant of time t_B as shown in Figure 2.8.

Light propagates in vacuum with invariant speed c on the hypersurface of null cone that is defined in the Cartesian coordinates $x^\alpha = (ct, x, y, z)$ by equation

$$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2 = 0 \,, \tag{2.195}$$

for infinitesimally close events. In case of the inertial reference frame, this equation can be easily integrated to obtain the time of propagation of light from one event to another. The times of emission, reflection, and reception of the light signal are related by simple equations

$$t_A = t_O + \frac{L_{OA}}{c} \,, \tag{2.196}$$

$$t_B = t_A + \frac{L_{AB}}{c} \,, \tag{2.197}$$

where $L_{OA} = x_A$ and $L_{AB} = |-x_A| = x_A$ are distances between points O and A, and A and B, correspondingly. Subtracting equation (2.197) from equation (2.196) and taking into account that for the static observer, the distance $L_{OA} = L_{AB}$, one obtains

$$t_A = \frac{1}{2}\left(t_O + t_B\right) \,. \tag{2.198}$$

The static observer located at the origin of frame S defines the instant of time

$$t_C = t_O + \frac{1}{2}\left(t_B - t_O\right) \,, \tag{2.199}$$

which corresponds to the event C on observer's worldline. Einstein's synchronization procedure postulates that the two events A and C are simultaneous in the static frame S. Indeed, simple inspection shows that the instant of time t_C of the observer who is always located at the origin of frame S coincides with the instant of time t_A of another static observer in this frame, whose worldline passes through event A at the time of the reflection of the signal: $t_C = t_A$. The clocks of the two observers are called synchronized in this case, and the events A and C are simultaneous in the inertial frame S (see Figure 2.9 which complements Figure 2.8).

19) In fact, Einstein's synchronization can be performed with any signal traveling with the fundamental speed c. Light is given a privilege only because it is the most convenient in practical realization of the synchronization procedure.

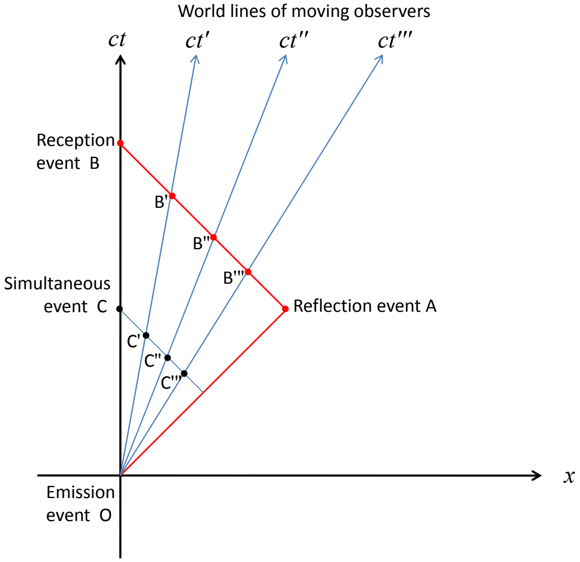

Figure 2.8 The Minkowski diagram explaining the Einstein's method of clock synchronization. The static observer emits light at point O toward an observer located at point A, and receives it back at point B after it gets reflected from A. Moving observers, which are passing through the point O at the instant of emission of the light signal with different velocities, will receive the light signal at points B', B", B"', depending on the speed of their motion. The events C, C', C", C"' lying on the past null cone passing through the event C, are simultaneous with the event A for the corresponding moving observers. These events of simultaneity are determined geometrically by equalities of intervals: |OC'|=|C'B'|, |OC"|=|C"B"|, and so on, in accordance with equation (2.200).

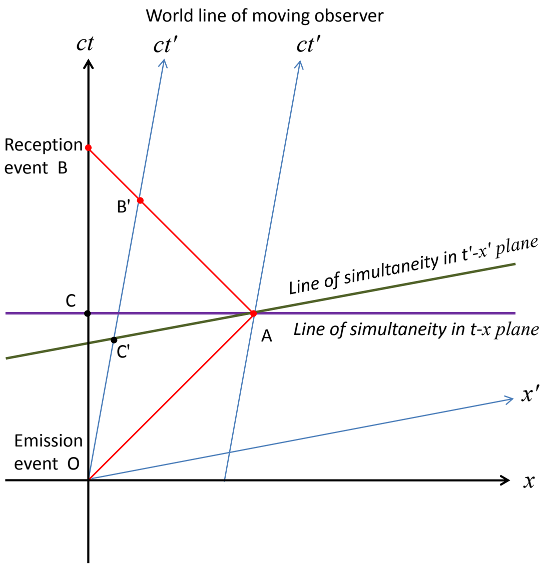

Figure 2.9 There is no absolute simultaneity in special relativity. For a given event A, the Einstein procedure of synchronization of clocks by exchange of light signals, leads to different lines of simultaneity in two frames, S and S', moving with respect to each other. Event C is simultaneous with event A for observer being at rest in the origin of frame S. However, for observer being at rest in the origin of the moving frame S', the event C' is simultaneous with A. All events are identical to those shown in Figure 2.8.

Because light propagates in vacuum with the invariant speed c that has the same numerical value in all reference frames, one can extend the Einstein rule of synchronization of clocks and the special-relativistic concept of simultaneity to arbitrarily moving frames. For example, let us consider a frame S' moving in a positive direction of x-axis of a static frame S. Time t' in the frame S' is measured by clocks of observers being at rest with respect to this frame and the spatial coordinates of the moving frame are denoted (x', y', z'). Let one observer is located at the origin of the frame S' and he sends a light signal to a second observer, who is located on x' axis at a constant distance x'_A, at the instant of time t'_O as the observer in the static frame S, that is $t'_O = t_O$. The second observer in frame S' receives the light signal at the event A, and send it immediately back at the time t'_A to the first moving observer. He will receive the reflected light signal at the time $t'_{B'}$ at the event B' that corresponds to the point of intersection of the worldline of the moving observer (that is ct' axis) with the future null cone having vertex at the event A, as shown in Figure 2.8. By definition, the instant of time $t'_{C'}$ in the frame S', which is simultaneous with the event A, is defined for the observer located at the origin of S' frame by the Einstein rule

$$t'_{C'} = t'_O + \frac{1}{2}\left(t'_{B'} - t'_O\right) . \tag{2.200}$$

It can be checked that the event $t'_{C'} = t'_A$ by doing calculations in the frame S'. Figure 2.9 visualizes the concept of simultaneity in the moving inertial frame S'. The lines of simultaneity are parallel to x'-axis, which corresponds to the line of simultaneity $t' = 0$. The same procedure can be applied to a set of differently moving observers to synchronize their clocks with the observer A. The events that are simultaneous with the event A will be lying on a null hypersurface shown in Figure 2.8 as a null line passing through points C, C', C'', C''', etc.

Einstein's procedure of synchronization of clocks can be applied globally to a set of inertial observers being at rest or moving uniformly with respect to each other. The word 'inertial' turns out to be a key to success. Observers associated with non-inertial frames of reference (rotating, accelerating) or experienced the presence of gravitational field may not be able either to synchronize of clocks with the Einstein procedure or to keep the Einstein synchronization intact. In particularly, Einstein's synchronization of clocks along a closed loop turns out to be impossible in rotating frames of reference. Clocks in accelerated frames after having been synchronized will not be able to maintain the synchronization as the time goes on [Zel'manov and Agakov, 1989].

2.5.7 Contraction of length

Consider a rigid rod [20] aligned along the x-axis of inertial frame S, and moving along it with velocity V. One is going to compare the length of the rod in two frames: static frame S covered with coordinates (ct, x, y, z), and a frame S' covered with coordinates (ct', x', y', z'), in which the rod is at rest. In the rest frame of the rod its endpoints have coordinates $(ct'_A, x'_A, 0, 0)$ and $(ct'_B, x'_B, 0, 0)$, where, in fact, $t'_A = t'_B$ because the length of the rod is measured in its rest frame on the spatial hypersurface of simultaneous events. One stipulates that $x'_B > x'_A$. The difference between spatial coordinates of the rod in frame S',

$$\ell = x'_B - x'_A \,, \tag{2.201}$$

is the physical length of the rod, also called the proper length.

In the static frame the spatial coordinates of the endpoints of the rod are $(ct_A, x_A, 0, 0)$ and $(ct_B, x_B, 0, 0)$ respectively. The difference between spatial coordinates of the rod in frame S,

$$L = x_B - x_A \,, \tag{2.202}$$

is called the coordinate length of the rod in the frame S. The spatial coordinates of the endpoints in two frames are connected by the Lorentz transformation (2.80)

$$ct_A = \gamma\,(ct'_A - \beta x'_A) \,, \qquad x_A = \gamma\,(x'_A - \beta ct'_A) \,, \tag{2.203a}$$

$$ct_B = \gamma\,(ct'_B - \beta x'_B) \,, \qquad x_B = \gamma\,(x'_B - \beta ct'_B) \,, \tag{2.203b}$$

where $\beta = V/c$, $\gamma = 1/\sqrt{1-\beta^2}$. Let us assume that an observer in frame S measures the coordinate length of the moving rod such that $t_A = t_B$. The proper length ℓ of the rod relates to its coordinate length L as follows

$$L = \gamma\,[\ell - \beta c(t'_B - t'_A)] \,, \tag{2.204}$$

where the time difference $t'_B - t'_A \neq 0$, because the endpoints of the rod are not lying in this measurement on the line of simultaneity in the frame S'. From the measurement condition, $t_A = t_B$, one finds from the Lorentz time transformation equations

$$t'_B - t'_A = \frac{\beta}{c}\ell \,. \tag{2.205}$$

Substituting this result in equation (2.204) one obtains,

$$L = \ell\sqrt{1-\beta^2} < \ell \,, \tag{2.206}$$

20) The word "rigid" means that the body is made of sufficiently inelastic material. There are no absolutely rigid bodies in special relativity. If such body existed it would allow propagation of sound waves inside it with speed larger than c that contradicts special relativity principles.

because the inverse Lorentz factor $\sqrt{1-\beta^2} < 1$. Hence, the coordinate length L of a moving rod is always smaller than the proper length ℓ of the rod in the direction of its motion, an effect called the *length contraction*. The spatial size of the rod in the directions orthogonal to the direction of its motion does not change. The proper length ℓ is longer than that measured in any moving inertial frame. As the speed of the rod approaches the fundamental speed c, its coordinate length L gets minuscule in the direction of motion.

The procedure of measuring length described in this section is idealized and should not be interpreted too straightforward. Practical measurement of the size of moving bodies is always done by making use of light rays emitted by the body towards observer who is making a photographic snapshot of the body or uses other technical means. Though the coordinate size of the body is reduced in the direction of its motion, it does not mean that the body will *look* squeezed. For example, photographic picture of a moving sphere does not reveal any Lorentz contraction of its spherical shape because the effect of the contraction is compensated by another relativistic effect, known as the *aberration of light*. Discussion of how the body moving with a relativistic speed will look like, can be found in [Nowojewski *et al.*, 2004; Smorodinskiǐ and Ugarov, 1972] or in the problem book by Batygin and Toptygin [1978].

It will be also erroneous to think that the Lorentz contraction causes appearance of internal stresses inside the body in its own rest frame. If such additional stresses appeared it would violate the principle of relativity so that an observer comoving with the body would be able to measure its velocity with respect to another inertial reference frame from the internal experiments conducted in the rest frame of the body. This scenario is unphysical and impossible. Nevertheless, the Lorentz contraction is a solid element of the mathematical structure of special (and general) theory of relativity. Calculation of relativistic equations of motion of extended bodies requires careful accounting for the length contraction in performing volume integration to avoid appearance of spurious, non-existing effects in the orbital motion of the bodies. We discuss this question in Chapter 6.3 in more detail.

2.5.8
Aberration of light

Let us now consider the effect of aberration of light, that is the change in the apparent position of the light source (star) as seen by a moving observer. The effect is produced by the Lorentz transformation from a static frame to the frame of the observer and for this reason, it is independent of the distance between observer and the source of light.

Let a light particle (photon) be emitted by a source of light and propagate in vacuum in the inertial reference frame S along a straight line in the direction of a unit, three-dimensional vector $\boldsymbol{k}$. Let us consider another inertial reference frame S' moving with respect to frame S with velocity $\boldsymbol{V}$. In the frame S', the photon propagates along a straight line in the direction of the unit vector $\boldsymbol{k}'$ which is different from vector $\boldsymbol{k}$. Numerical value of the speed of light propagating in vacuum is equal to the fundamental speed c that is the same in any frame. It does not depend on the velocity of the source of light either. Thus, the velocities of light in frames S and S' are given by vectors, $\boldsymbol{v} = c\boldsymbol{k}$ and $\boldsymbol{v}' = c\boldsymbol{k}'$, respectively. One can use the law of transformation of velocities (2.174) in order to connect the directions of propagation of light ray in both frames. It yields the aberration of light equation,

$$\boldsymbol{k}' = \frac{\boldsymbol{k} + (\gamma - 1)V^{-2}(\boldsymbol{V}\cdot\boldsymbol{k})\boldsymbol{V} - \gamma c^{-1}\boldsymbol{V}}{\gamma(1 - c^{-1}\boldsymbol{k}\cdot\boldsymbol{V})} . \tag{2.207}$$

Let us consider now two observers, one observer is at rest with respect to frame S and another observer is at rest with respect to frame S'. The observer at frame S sees light arriving from direction $\boldsymbol{n} = -\boldsymbol{k}$, while in the frame S' the observer sees light arriving from the direction defined by the unit vector $\boldsymbol{n}' = -\boldsymbol{k}'$. Hence, the equation of the aberration of light becomes [Kovalevsky and Seidelmann, 2004]

$$\boldsymbol{n}' = \frac{\boldsymbol{n} + (\gamma - 1)V^{-2}(\boldsymbol{V}\cdot\boldsymbol{n})\boldsymbol{V} + \gamma c^{-1}\boldsymbol{V}}{\gamma(1 + c^{-1}\boldsymbol{n}\cdot\boldsymbol{V})} . \tag{2.208}$$

This equation depends only on the relative velocity $\boldsymbol{V}$ of the two observers (the two frames S and S'), and it does not depend on the velocity of the source of light that is totally irrelevant in the interpretation of the phenomenon of aberration of light. It is worth emphasizing that if one's observation performed in one frame can not be compared with observations made in an other frame, moving with respect to the first one, the aberration of light becomes undetectable. It means that the aberration of light *can not* determine the presence of an “absolute" frame of reference or æther. For example, the solar system moves with respect to the center of the Milky Way with the speed approximately equal to $V = 220$ km/s. Due to this motion positions of all stars are displaced by the aberration of light at some constant angle - the effect known as the *secular aberration* [Kovalevsky and Seidelmann, 2004]. It can be calculated from theory but it can not be observed until the speed is considered as constant because astronomers in the solar system can not compare their measurement with the directions to stars measured by an external observer being at rest with respect to the center of the galaxy [Hagihara, 1933]. Only the gradual change in the velocity of the solar system caused by its acceleration towards the center of the Milky Way allows to see the presence of the secular aberration that appears as a particular pattern in the proper motions of stars [Kopeikin and Makarov, 2006].

Vectors $\boldsymbol{n}$ and $\boldsymbol{n}'$ can be decomposed in two components - parallel and orthogonal to velocity $\boldsymbol{V}$. We denote $\cos\theta = (\boldsymbol{n}\cdot\boldsymbol{V})/V$, and $\sin\theta = |\boldsymbol{n}\times\boldsymbol{V}|/V$ in frame S, and $\cos\theta' = (\boldsymbol{n}'\cdot\boldsymbol{V})/V$, and $\sin\theta' = |\boldsymbol{n}'\times\boldsymbol{V}|/V$ in frame S'. Projecting the aberration-of-light equation (2.208) onto the corresponding directions yields the following trigonometric relationships for the angles θ and θ'

$$\cos\theta' = \frac{\cos\theta + \dfrac{V}{c}}{1 + \dfrac{V}{c}\cos\theta}\,, \tag{2.209}$$

$$\sin\theta' = \frac{\sin\theta\sqrt{1 - \dfrac{V^2}{c^2}}}{1 + \dfrac{V}{c}\cos\theta}\,. \tag{2.210}$$

In the slow-motion limit, when $V \ll c$, these equations can be expanded into Taylor series with respect to the small parameter V/c. Denoting, $\Delta\theta = \theta - \theta'$, one derives the angular displacement due to aberration,

$$\sin\Delta\theta = \frac{V}{c}\sin\theta - \frac{V^2}{4c^2}\sin 2\theta + \frac{V^3}{4c^3}\sin 2\theta\cos\theta + \ldots\,, \tag{2.211}$$

or, after additional expansion of the left side of this equation with respect to $\Delta\theta$,

$$\Delta\theta = \frac{V}{c}\sin\theta - \frac{V^2}{4c^2}\sin 2\theta + \frac{V^3}{6c^3}\sin\theta\left(1 + 2\cos^2\theta\right) + \ldots\,. \tag{2.212}$$

The first (linear) term is the classic aberration of light, known yet from the time of Newton's theory of light, and the other terms are special-relativistic corrections. In case, of the annual aberration of light caused by the orbital motion of the Earth around the barycenter of the solar system with speed $V = 30$ km/s, the classic term reaches magnitude of 20 arcseconds, the quadratic term is about 0.5 milliarcsecond (mas), and the third term [21] is nearly 0.05 microarcsecond (μas), being currently unacceptable for observations.

One can notice that as the speed V approaches the speed of light c the angle θ' measured by the moving observer is approaching to zero for almost all light rays except for a tiny fraction lying very closely to the direction of $\theta = -\pi$

$$\tan\theta' = \sqrt{1 - \frac{V^2}{c^2}}\tan\frac{\theta}{2} + \ldots\,. \tag{2.213}$$

This effect is similar to lining up the velocities of particles in the frame moving ultra-relativistically as discussed in equation (2.175). One astronomical consequence of this is that if all stars in the sky were distributed uniformly over the celestial sphere, the observer moving with ultra-relativistic speed would see the stars displaced toward the point on the sky in the direction of its motion. In the limit of $V \to c$, the entire stellar sky of the observer would shrink to a single bright point embracing all stars in the sky. It is really impossible to use stars for navigating the spaceship moving with such an ultra-relativistic speed!

It may look like that measuring the anisotropy in the distribution of stars caused by a uniform and rectilinear motion of observer could be used, at least in principle,

21) Similar term in [Equation 6.28]Kovalevsky and Seidelmann [2004] has a typo.

for determination of the “absolute" speed of this motion. However, it requires a prior information about the star’ distribution obtained independently by an other observer that does not move with respect to the stars. Hence, again one can see that the application of the aberration of light for measuring the speed of the ultra-relativistic “absolute" motion, is impossible.

2.5.9 The Doppler effect

Let an inertial frame S be parameterized with coordinates $x^\mu = (ct, \boldsymbol{x})$. Consider a source of light and an observer moving along worldlines $x_s^\mu(t) = (c\tau_s(t), \boldsymbol{x}_s(t))$ and $x_o^\mu(t) = (c\tau_o(t), \boldsymbol{x}_o(t))$ respectively, where τ_s and τ_o are proper times of the source of light and the observer respectively. Let the source of light emit a monochromatic electromagnetic wave at time t_1 that is received by the observer at time t_2. The events of emission and observation are interconnected by the null-cone equation

$$t_2 = t_1 + \frac{1}{c}|\boldsymbol{x}_o(t_2) - \boldsymbol{x}_s(t_1)| \,. \tag{2.214}$$

If the functions $\boldsymbol{x}_s(t)$ and $\boldsymbol{x}_o(t)$ are known, this equation describes the time t_1 as an implicit function of t_2 and vice versa. Taking a time derivative from both sides of equation (2.214) one obtains

$$\frac{dt_1}{dt_2} = \frac{1 - \boldsymbol{k}\cdot\boldsymbol{\beta}_o}{1 - \boldsymbol{k}\cdot\boldsymbol{\beta}_s} \,, \tag{2.215}$$

where the unit vector

$$\boldsymbol{k} = \frac{\boldsymbol{x}_o(t_2) - \boldsymbol{x}_s(t_1)}{|\boldsymbol{x}_o(t_2) - \boldsymbol{x}_s(t_1)|} \,, \tag{2.216}$$

points out in the direction of propagation of the electromagnetic wave, while $\boldsymbol{\beta}_s = \boldsymbol{v}_s(t_1)/c$ and $\boldsymbol{\beta}_o = \boldsymbol{v}_o(t_2)/c$ are velocities of the source of light and the observer, normalized to the fundamental speed c. Notice that motion of the observer and the source of light have not been restricted by a rectilinear uniform motion. Hence, velocities $\boldsymbol{v}_o$ and $\boldsymbol{v}_s$ are functions of time t and should be taken at the appropriate instants of time as indicated above.

The frequency, ν_s, of the emitted wave is related to its period, $\Delta\tau_s$, expressed in terms of the proper time of the source of light as $\nu_s = 1/\Delta\tau_s$. The observed frequency of the wave, ν_o, is related to its period, $\Delta\tau_o$, expressed in terms of the proper time of observer as $\nu_o = 1/\Delta\tau_o$. Assuming that both $\Delta\tau_s$ and $\Delta\tau_o$ are extremely small, one can interpret them as the infinitesimal increments, $d\tau_s$ and $d\tau_o$. Hence, the ratio of the two frequencies can be expressed as an ordinary derivative of the proper time τ_s of the source, taken with respect to the proper time τ_o of observer,

$$\frac{\nu_o}{\nu_s} = \frac{d\tau_s}{d\tau_o} \,. \tag{2.217}$$

The proper time τ is related to the coordinate time t in accordance with equation (2.162) which should be applied separately to the source of light and to the observer,

$$d\tau_s = dt_1 \sqrt{1-\beta_s^2}\,, \tag{2.218}$$

$$d\tau_o = dt_2 \sqrt{1-\beta_o^2}\,. \tag{2.219}$$

Substituting these expressions in equation (2.217) and recollecting that the time increment dt_2 is connected to dt_1 by the null cone equation (2.215), one obtains

$$\frac{\nu_o}{\nu_s} = \sqrt{\frac{1-\beta_s^2}{1-\beta_o^2}}\,\frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_o}{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_s}\,. \tag{2.220}$$

This equation relates the observed and emitted frequencies of the electromagnetic wave and expresses the *Doppler effect* in special relativity. It is named after Austrian physicist Christian Doppler who proposed it in 1842. Equation (2.220) is valid for arbitrary motion of the source of light and the observer, but neglects all gravitational effects which will be considered later in Chapter 7. Derivation of the Doppler effect given above, is based on the so-called time-transformation technique [Brumberg, 1972, 1991; Kopeikin and Ozernoy, 1999]. The frequency-transformation technique and the technique based on the Lorentz transformation of the four-vector of electromagnetic wave can be applied as well, resulting in the same equation [Kopeikin and Ozernoy, 1999].

The Doppler effect equation (2.220) depends on the coordinate velocities of observer and the source of light with respect to a chosen inertial reference frame S. It may make an impression that the Doppler effect depends on the choice of the reference frame but this impression is illusory. Indeed, one can introduce a null vector $k^\alpha = (1, \boldsymbol{k})$ and four-velocity of the observer $u_o^\alpha = \gamma_o(1, \boldsymbol{\beta}_o)$, and that of the source of light, $u_s^\alpha = \gamma_s(1, \boldsymbol{\beta}_s)$ where $\gamma_o = \left(1-\beta_o^2\right)^{-1/2}$ and $\gamma_s = \left(1-\beta_s^2\right)^{-1/2}$ are the corresponding Lorentz factors. In terms of these vectors the Doppler shift equation (2.220) can be written in apparently Lorentz-invariant form

$$\frac{\nu_o}{\nu_s} = \frac{k_\alpha u_o^\alpha}{k_\alpha u_s^\alpha}\,. \tag{2.221}$$

This form of the equation is valid in any inertial frame, and it makes no preference between them in accordance with the Einstein principle of relativity. It means that the magnitude of the Doppler effect depends only on the relative velocity of observer and the source of light, the conclusion that is not directly visible in equation (2.220).

2.6 Accelerated Frames

Inertial reference frames have preference in special relativity because the Minkowski spacetime is globally homogeneous and isotropic. It allows us to associate each inertial reference frame with a globally orthonormal basis (tetrad) $\boldsymbol{e}_\alpha$ which has one timelike, $\boldsymbol{e}_0$, and three spacelike, $\boldsymbol{e}_i$, unit vectors. The metric tensor, $g_{\mu\nu}$ in the inertial frame is a diagonal Minkowski metric $g_{\mu\nu} = \eta_{\mu\nu}$. The inertial basis $\boldsymbol{e}_\alpha$ corresponds to the Cartesian coordinates, $x^\alpha = (x^0, x^i) = (ct, x^i)$, covering the entire Minkowski spacetime. Coordinate time t and spatial coordinates $x^i = \boldsymbol{x}$ are identified respectively with the proper time and the proper length, measured by any observer, who is at rest with respect to the inertial frame. The static observers have their clocks synchronized. The Cartesian coordinates of any two inertial frames are connected with each other by the Lorentz boost transformation, which describes a uniform rectilinear motion of one frame with respect to another with a constant speed $\boldsymbol{V}$. It may also include additional constant rotation in space.

Newtonian mechanics admits non-inertial (that is accelerating and/or rotating) frames of reference which are more convenient in some cases to solve practical problems of motion of bodies in celestial mechanics. Is it possible to extrapolate the classic approach and to introduce non-inertial frames of reference to the theory of special relativity in a self-consistent way? The answer is affirmative, and this section demonstrates below how to build such frames in case of accelerated observers [22). This approach builds a bridge between the special and general theories of relativity.

In orbital mechanics, a perturbed trajectory of a particle gets assembled of points, each of which is collocated with an unperturbed instantaneous curve in a certain way. Typically, though not necessarily, the instantaneous curves are chosen to be osculating, i.e., tangent to the resulting perturbed orbit. It is natural to complete the unit tangent vector, $\boldsymbol{T}$, to the curve by two other orthogonal unit vectors– normal $\boldsymbol{N}$ and binormal $\boldsymbol{B} = \boldsymbol{T} \times \boldsymbol{N}$. The three vectors collectively form an orthonormal basis in three-dimensional space called the *Frenet-Serret frame* [Dubrovin *et al.*, 1984, §5] after the two French mathematicians who independently introduced and explored this frame: Jean Frédéric Frenet, in his thesis of 1847, and Joseph Alfred Serret in 1851.

A similar idea can be implemented to the construction of the locally-orthonormal tetrad which origin moves with an acceleration along a timelike worldline in the Minkowski spacetime. At each instant of time such tetrad coincides with an inertial reference frame moving in space with the same velocity as the accelerated tetrad. This concept was brought from differential geometry into the special relativity by Møller [1952] and Synge [1964], who had derived the analogue of the Frenet-Serret differential equations describing evolution of the tetrad moving along an arbitrary timelike curve in the Minkowski spacetime [23). It was discovered that the continu-

22) Rotating frames can be treated similarly. We refer the reader to the books [Møller, 1952, §84], [Landau and Lifshitz, 1975, §89], and [Zel′manov and Agakov, 1989, Section 10] for more detail.

23) Synge [1964] has also analyzed a more general case of the tetrad transported along an arbitrary space-

ous transition from one instantaneously comoving inertial frame to another must be accompanied with the infinitesimal Lorentz boost, to keep the timelike unit vector $\boldsymbol{e}_0$ of the tetrad tangent to the curve, and rotation of the tetrad's spatial vectors $\boldsymbol{e}_i$. If the spatial axes of the moving tetrad are kept parallel to the spatial axes of a global inertial reference frame, the law of transportation of the tetrad is described by the Fermi-Walker equation . In what follows, the spatial rotation of the local tetrad will be excluded from consideration.

Let us denote the tetrad of unit vectors of the accelerated reference frame S' as $\boldsymbol{e}_{\alpha'}$, and the Cartesian coordinates corresponding to the tetrad as $x^{\alpha'}$. The accelerated tetrad is to coincide at each instant of time with a comoving inertial reference frame so that one can interpret the coordinates $x^{\alpha'}$ as the coordinates of the locally inertial frame. Transformation from a global inertial frame S covered with the Cartesian coordinates x^α to the local inertial frame has to coincide at each instant of time with the Poincaré transformation (2.86) between two inertial reference frames given by equation

$$\mathbf{x} = \mathbf{z} + \boldsymbol{e}_{j'} x^{j'} \,, \tag{2.222}$$

or, in coordinate notations,

$$x^\alpha = z^\alpha + \Lambda^\alpha{}_{j'} x^{j'} \,, \tag{2.223}$$

where vector $\mathbf{z}$ points out from the origin of the global inertial frame S to the origin of the accelerated tetrad $\boldsymbol{e}_{\alpha'}$, and $\Lambda^\alpha{}_{\beta'}$ is the matrix of the Lorentz boost corresponding to the instantaneous value of four-velocity of the observer $\mathbf{u} = \boldsymbol{e}_{0'} = u^\alpha \boldsymbol{e}_\alpha$. However, now the unit vectors of the tetrad $\boldsymbol{e}_{\alpha'} = \boldsymbol{e}_{\alpha'}(\tau)$ are functions of the proper time τ of the observer, and vector of translation $\mathbf{z} = \mathbf{z}(\tau)$ describes instantaneous position of the observer on its worldline which is no longer a linear function of time as in case of a uniformly-moving inertial observer. The hypersurface of constant value of the proper time τ, taken on the observer's worldline, is a hypersurface of simultaneity of the comoving inertial observer. This hypersurface is instantaneously-orthogonal to the four-velocity of observer.

It is intuitively clear from equation (2.223) that the hypersurfaces of simultaneity should intersect in space at some distance from the tetrad's worldline as visualized in Figure 2.10. It means that the coordinates associated with the accelerated tetrad can not cover the entire Minkowski spacetime, but only a part of it. The domain of coverage is determined by the magnitude of the tetrad's acceleration. Thus, the observer who carries out the tetrad along an accelerated worldline, has to admit that the spacetime around him conforms to a Minkowski spacetime only locally. In fact, this is a starting point for going from the Minkowski world of special relativity to the realm of a curved spacetime of general relativity.

Coordinate transformation (2.223) between the global inertial coordinates x^α and the local inertial coordinates of an accelerated observer depends on the matrix of the Lorentz transformation which is a smooth function of observer's proper time,

like curve.

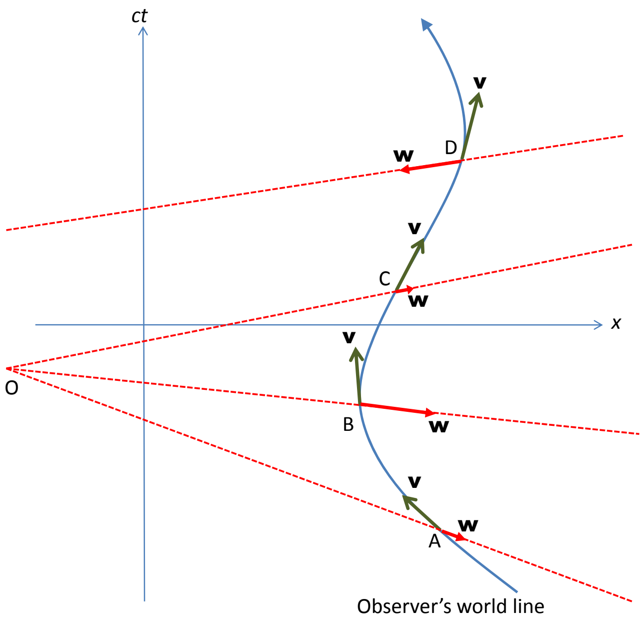

Figure 2.10 A Minkowski diagram for an observer moving along an arbitrary worldline with variable four-acceleration $\mathbf{w}$. Events A, B, C, D on the worldline are the origins of the local inertial frame comoving with the observer. Four-velocity of the observer is a time-like unit vector $\mathbf{v}$ being tangent to the worldline. Four-acceleration $\mathbf{w}$ is a space-like vector being orthogonal to four-velocity, $\mathbf{u} \cdot \mathbf{w} = 0$. Spacelike hypersurfaces of the comoving inertial frame are orthogonal to four-velocity, and are shown by dashed lines. Since the four-velocity changes its direction as the observer moves, the hypersurfaces are not parallel to each other and intersect at point O at the distance being inversely proportional to the magnitude of observer's proper acceleration $|\boldsymbol{A}|$. Hence, the local inertial coordinates of the observer fail to cover the entire Minkowski spacetime.

$\Lambda^{\alpha}{}_{\beta'}(\tau)$. Its temporal evolution is determined by the transportation law that take the matrix smoothly from one point of observer's worldline to another in a way that is to be consistent with the temporal change of the four-velocity of the observer that is defined by observer's four-acceleration. Derivation of the transportation law requires solution of the problem of motion of an observer moving with acceleration. For simplicity, one shall focus on the case of a uniform, rectilinear motion with a constant acceleration that will demonstrate that special relativity is robust enough to handle accelerated frames of reference.

2.6.1
Worldline of a uniformly-accelerated observer

Let the inertial frame S be static, and the accelerated frame S' be attached to an observer who is always at rest at the origin of S'. Keeping the three-dimensional acceleration $\boldsymbol{a} = d\boldsymbol{v}/dt = (d^2x/dt^2, d^2y/dt^2, d^2z/dt^2)$ of the observer constant in the inertial frame S, would linearly increase its speed with time, thus, leading to an unsurmountable obstacle – the ultimate speed c. Hence, the observer cannot be uniformly accelerated with respect to a static inertial frame and the only possibility is to admit that the proper acceleration of the observer, $\boldsymbol{A} = d\boldsymbol{v}'/dt = (d^2x'/dt^2, d^2y'/dt^2, d^2z'/dt^2)$, is kept constant in frame S' [Rindler, 1960a]. The proper acceleration $\boldsymbol{A}$ and its relation to four-acceleration $\boldsymbol{w}$ were defined in section 2.5.2.

In the case of one-dimensional motion, the proper acceleration $\boldsymbol{A}$ is related to the static-frame acceleration $\boldsymbol{a}$ by equation (2.190). Assuming that the motion is along x-axis, the acceleration components are: $\boldsymbol{a} = (dv/dt, 0, 0)$ and $\boldsymbol{A} = (A, 0, 0)$, where $\boldsymbol{v} = (v, 0, 0)$ is the variable speed of the observer in frame S, and A is constant. Two other spatial coordinates do not change as the observer moves: $y' = y$, $z' = z$. Making use of these notations, equation (2.190) can be written as follows

$$\frac{dv}{(1 - v^2/c^2)^{3/2}} = A dt \,. \tag{2.224}$$

It can be integrated, yielding the speed of the observer with respect to frame S as a function of time. Assuming for simplicity that the initial speed of the observer $v_0 = 0$ at the time $t = 0$, one obtains

$$v = \frac{At}{\sqrt{1 + \dfrac{A^2t^2}{c^2}}} \,. \tag{2.225}$$

Substituting for the speed, $v = dx/dt$, and integrating equation (2.225) one more time, one gets

$$x = \frac{c^2}{A}\sqrt{1 + \frac{A^2t^2}{c^2}} \,, \tag{2.226}$$

where one has chosen the constant of integration equals to zero, so that $x = x_0 = c^2/A$ at time $t = 0$.

Equation (2.226) represents a hyperbola in Minkowski spacetime,

$$x^2 - c^2t^2 = x_0^2\,, \qquad \left(x_0 \equiv \frac{c^2}{A}\right) \tag{2.227}$$

which occupies the right wedge of the null cone having the origin at point O, and intersects x-axis of the static frame S at point $x_0 = c^2/A$ at time $t = 0$, as shown in Figure 2.11. Equation (2.225) reveals that the observer will reach ultra-relativistic speed when time $t \to \pm\infty$. In this case one has velocity of the observer, $v \to \pm c$, and its coordinate $x \to \pm ct$, which represent two asymptotes of the hyperbola (2.227). In the slow-motion approximation, one has $At \ll c$, so that velocity of the observer is approximated by $v = At$, and its coordinate $x = x_0 + At^2/2$, which conforms to classical mechanics.

One notices that the two asymptotes of the hyperbola, $x = x_0 \pm ct$, do not depend on the particular value of the proper acceleration A. It means that a family of hyperbolae parameterized by the acceleration A, is filling up the right wedge enveloped by the null cone $x = \pm ct$ with its vertex at the event $t = 0, x = 0$. On the other hand, if one fixes the value of acceleration A, the straight lines (2.232) labeled with different values of the proper time τ, will make a family of rays emanating from point $t = 0, x = 0$ and filling up the right wedge of the null cone (see Figure 2.11).

The hyperbola can be parameterized with the proper time τ of the accelerated observer that relates to the coordinate time t of the static frame by equation

$$\frac{d\tau}{dt} = \sqrt{1 - \frac{v^2}{c^2}}\,, \tag{2.228}$$

where the speed, v, of the observer is given by equation (2.225). Integrating equation (2.228) with the initial condition, $\tau = 0$ at time $t = 0$, one obtains

$$c\tau = x_0 \ln\left[\frac{At}{c} + \sqrt{1 + \frac{A^2t^2}{c^2}}\right]\,. \tag{2.229}$$

This equation can be inverted, leading to the following, proper-time parametrization of the hyperbolic worldline $\mathbf{z} = (ct, x, y, z)$ of the uniformly-accelerated observer

$$ct = x_0 \sinh\left(\frac{A\tau}{c}\right)\,, \tag{2.230a}$$

$$x = x_0 \cosh\left(\frac{A\tau}{c}\right)\,, \tag{2.230b}$$

$$y = 0\,, \tag{2.230c}$$

$$z = 0\,. \tag{2.230d}$$

This parametrization of the worldline of the accelerated observer substituted in equation (2.225) for instantaneous velocity of the observer, yields

$$v = c\tanh\left(\frac{c\tau}{x_0}\right)\,, \tag{2.231}$$

Dividing equation (2.230b) by equation (2.230a), and accounting for equation (2.231), gives

$$ct = \frac{v}{c}\,x\,, \tag{2.232}$$

that is a straight line in the (ct, x) plane for each fixed value of the proper time τ. This line corresponds to a hypersurface of simultaneity of an inertial frame instantly comoving with the observer with the instantaneous velocity $v = c\tanh(c\tau/x_0)$.

2.6.2
A tetrad comoving with a uniformly-accelerated observer

The accelerated observer carries out a tetrad $\boldsymbol{e}_{\alpha'} = (\boldsymbol{e}_{0'}, \boldsymbol{e}_{1'}, \boldsymbol{e}_{2'}, \boldsymbol{e}_{3'})$, which origin traces the observer's worldline. Relationship between the local tetrad and the basis $\boldsymbol{e}_\alpha$ of the entire Minkowski space-time is given at each instant of time by the instantaneous Lorentz transformation

$$\boldsymbol{e}_{\alpha'} = \Lambda^{\beta}{}_{\alpha'}\boldsymbol{e}_\beta\,, \tag{2.233}$$

As the motion is along x-axis only, the components $\Lambda^{\beta}{}_{\alpha'}$ of the matrix of the instantaneous Lorentz boost are given by the inverse of equation (2.73) with the parameter of rapidity $\theta = \tanh^{-1}(v/c) = c\tau/x_0$,

$$\begin{bmatrix} \Lambda^0{}_{0'} & \Lambda^0{}_{1'} & \Lambda^0{}_{2'} & \Lambda^0{}_{3'} \\ \Lambda^1{}_{0'} & \Lambda^1{}_{1'} & \Lambda^1{}_{2'} & \Lambda^1{}_{3'} \\ \Lambda^2{}_{0'} & \Lambda^2{}_{1'} & \Lambda^2{}_{2'} & \Lambda^2{}_{3'} \\ \Lambda^3{}_{0'} & \Lambda^3{}_{1'} & \Lambda^3{}_{2'} & \Lambda^3{}_{3'} \end{bmatrix} = \begin{bmatrix} \cosh\left(\frac{A\tau}{c}\right) & \sinh\left(\frac{A\tau}{c}\right) & 0 & 0 \\ \sinh\left(\frac{A\tau}{c}\right) & \cosh\left(\frac{A\tau}{c}\right) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{2.234}$$

Making use of this parametrization the tetrad transformation (2.233) can be rewritten in more explicit form

$$\boldsymbol{e}_{0'}(\tau) = \cosh\left(\frac{A\tau}{c}\right)\boldsymbol{e}_0 + \sinh\left(\frac{A\tau}{c}\right)\boldsymbol{e}_1\,, \tag{2.235a}$$

$$\boldsymbol{e}_{1'}(\tau) = \sinh\left(\frac{A\tau}{c}\right)\boldsymbol{e}_0 + \cosh\left(\frac{A\tau}{c}\right)\boldsymbol{e}_1\,, \tag{2.235b}$$

$$\boldsymbol{e}_{2'}(\tau) = \boldsymbol{e}_2\,, \tag{2.235c}$$

$$\boldsymbol{e}_{3'}(\tau) = \boldsymbol{e}_3\,, \tag{2.235d}$$

where the unit vectors $\boldsymbol{e}'_2$ and $\boldsymbol{e}'_3$ of the observer's tetrad stay unaffected, and coincide with the appropriate vectors $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ of the basis in the static frame S.

Time evolution of the tetrad moving with constant acceleration is derived by differentiation of equations (2.235), and is given by

$$\frac{d\boldsymbol{e}_{0'}}{d\tau} = \frac{A}{c}\boldsymbol{e}_{1'}\,, \qquad \frac{d\boldsymbol{e}_{1'}}{d\tau} = \frac{A}{c}\boldsymbol{e}_{0'}\,, \qquad \frac{d\boldsymbol{e}_{2'}}{d\tau} = \frac{d\boldsymbol{e}_{3'}}{d\tau} = 0\,. \tag{2.236}$$

Its covariant relativistic generalization for arbitrary coordinates is given by

$$\frac{d\boldsymbol{e}_{\mu'}}{d\tau} = \left(w_{\mu'}u^{\nu'} - u_{\mu'}w^{\nu'}\right)\boldsymbol{e}_{\nu'}\,, \tag{2.237}$$

where $w^{\mu'}$ is four-acceleration, and $u^{\mu'}$ is four-velocity of the observer. Equation (2.237) is called the Fermi-Walker transport [Misner *et al.*, 1973]. It preserves the length of the transported vectors and leads to an interesting phenomenon called the Thomas precession [Thomas, 1926]. Mathematics associated with the Fermi-Walker transport was thoroughly discussed by Ni and Zimmermann [1978a] in a more general case of an accelerated tetrad which spatially-rotates at a rate $\boldsymbol{\Omega}$. Such an observer will find freely-moving test particle of mass m experiences not only the inertial force $-m\boldsymbol{A}$ but also the Coriolis force $-2m\boldsymbol{\Omega}\times\boldsymbol{v}$, with $\boldsymbol{v}$ being the spatial components of particle's velocity measured with respect to the comoving tetrad [24].

2.6.3
The Rindler coordinates

Equations (2.230) of the world line of a uniformly-accelerated observer combined with equations (2.235) defining the local tetrad are used to specify the local inertial coordinates $x^{\alpha'}$ of the observer. Indeed, employing these equations in transformation (2.223) between the global inertial coordinates $x^\alpha = (ct, x, y, z)$ of the Minkowski spacetime and the local inertial coordinates $x^{\alpha'} = \left(ct', x^{1'}, x^{2'}, x^{3'}\right) = (c\tau, \xi, \zeta, \eta)$ yields

$$ct = (\xi + x_0)\sinh\left(\frac{c\tau}{x_0}\right), \qquad (2.238a)$$

$$x = (\xi + x_0)\cosh\left(\frac{c\tau}{x_0}\right), \qquad (2.238b)$$

$$y = \zeta, \qquad (2.238c)$$

$$z = \eta, \qquad (2.238d)$$

where the local time $\tau \in (-\infty, \infty)$ coincides with the proper time of the observer on its worldline, and the local spatial coordinate $\xi \in (-x_0, \infty)$ with $x_0 = c^2/A$. Thus, the local inertial coordinates of a uniformly-accelerated observer do not cover the entire Minkowski spacetime but only its part due to the deformation of the coordinate grid caused by the acceleration of the worldline of the observer.

The local inertial coordinates $(c\tau, \xi, \zeta, \eta)$ in equations (2.238) were pioneered by Born [1909]. They have been also discussed by Møller [1943, 1952] and the most comprehensively by [Rindler, 1960b] so that they are known the most generally as the Rindler coordinates [D'Inverno, 1992; Misner *et al.*, 1973] shown graphically in Figure 2.11.

Let us mention that the half-line $ct = x$ of the null cone, plays the role of the observer horizon in the spacetime endowed with the metric (2.240) - the concept introduced by Rindler [1960b]. For example, the event $(t = 0, x = 0)$ in Figure 2.11 will never be seen by the accelerated observer, because the observer's worldline will never intersect with the worldline $ct = x$ of a photon (as well as a graviton, or any other massless particle) emitted at the event $(t = 0, x = 0)$. Hence, it is physically

24) See also discussion of equation (4.155) in Chapter 4.5.1.

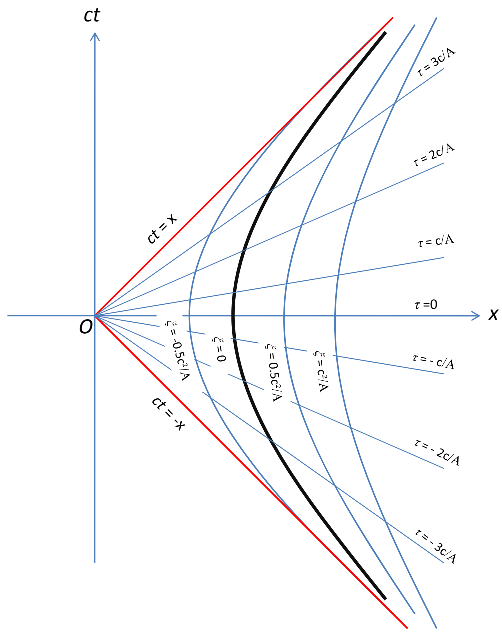

Figure 2.11 The local Rindler coordinates of an observer moving with a uniform acceleration A along x-axis. Worldline of the observer is shown as a heavy-black hyperbola. The grid of the local coordinates covers only the part of the Minkowski spacetime inside the right wedge of the null cone. The coordinate grid collapses at the distance $\xi = -c^2/A$ from the observer's worldline (at point O). This singularity is not the property of the Minkowski spacetime.

impossible to communicate any information from the point ($t = 0, x = 0$) to the accelerated observer.

The Rindler coordinates help to establish the mathematical equivalence between a homogeneous gravitational field and an inertial force caused by a constant acceleration of the local reference frame Einstein [1907]. This equivalence was promoted by Einstein to the rank of physical principle of equivalence which became a cornerstone of general theory of relativity. The principle of equivalence can be understood by calculating the metric tensor in the Rindler coordinates. The starting point is the metric tensor $g_{\mu\nu}$ in the static frame S covered with the Cartesian coordinates, x^μ, which is reduced in these coordinates to the Minkowski metric $g_{\mu\nu} = \eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$ in the entire Minkowski spacetime. Transformation (2.238) should be used in order to calculate the metric tensor $g_{\mu'\nu'}$ in the Rindler coordinates, $x^{\mu'} = (c\tau, \xi, \zeta, \eta)$. One has

$$g_{\mu'\nu'} = \frac{\partial x^\alpha}{\partial x^{\mu'}}\frac{\partial x^\beta}{\partial x^{\nu'}}\eta_{\alpha\beta} = -\frac{\partial x^0}{\partial x^{\mu'}}\frac{\partial x^0}{\partial x^{\nu'}} + \frac{\partial x^i}{\partial x^{\mu'}}\frac{\partial x^i}{\partial x^{\nu'}} , \tag{2.239}$$

where one assumes summation with respect to the repeated indices. Equation (2.239) yields

$$g_{0'0'} = -\left(1 + \frac{\xi}{x_0}\right)^2 , \qquad g_{1'1'} = g_{2'2'} = g_{3'3'} = 1 , \tag{2.240}$$

where $x_0 = c^2/A$, and all other components $g_{\mu'\nu'} = 0$ for indices $\mu' \neq \nu'$. Determinant g of the metric tensor in the Rindler coordinates

$$g = \det\left[g_{\mu'\nu'}\right] = -\left(1 + \frac{\xi}{x_0}\right)^2 , \tag{2.241}$$

gets equal to zero at the distance $\xi = -x_0$ from the origin of the Rindler coordinates, where the coordinate grid of these coordinates collapses. The metric tensor $g_{\mu'\nu'}$ is divergent at large distances and reduces to the Minkowski metric $\eta_{\mu'\nu'} = \mathrm{diag}(-1, 1, 1, 1)$ only at the origin of the Rindler coordinates. For small distances, $\xi \ll x_0$, the time component of the metric tensor behaves like

$$g_{0'0'} = -\left(1 + \frac{2\xi}{x_0}\right) , \tag{2.242}$$

that coincides up to a factor 2 with gravitational potential of a homogenous gravitational field, $U = -g\xi$, and $g = c^2/x_0 = A$. It is this equivalence that led Einstein to the idea to identify an arbitrary gravitational field with the metric tensor. The structure of the metric tensor of an accelerated observer is similar to that of the geocentric metric tensor in the theory of the astronomical reference frames, which is discussed in full details in Chapter 4.

It is instructive to establish a relationship between the Rindler coordinates of two different observers moving with different accelerations, A and $\bar{A}$, so that they are passing x-axis at different distances, $x_0 = c^2/A$ and $\bar{x}_0 = c^2/\bar{A}$. The Rindler coordinates of some event, measured by the first observer, are $x^{\alpha'} = (c\tau, \xi, \zeta, \eta)$ and those the same event measured by the second observer will be denoted as $\bar{x}^{\alpha'} = \left(c\bar{\tau}, \bar{\xi}, \bar{\zeta}, \bar{\eta}\right)$. It is clear that all relationships having been established above for the first observer will preserve their form for the second observer with a corresponding replacement of the coordinates $x^{\alpha'} \to \bar{x}^{\alpha'}$ and $x_0 \to \bar{x}_0$. In particular, coordinate transformation (2.238) referred to the second accelerated observer will read

$$ct = \left(\bar{\xi} + \bar{x}_0\right) \sinh\left(\frac{c\bar{\tau}}{\bar{x}_0}\right) , \tag{2.243a}$$

$$x = \left(\bar{\xi} + \bar{x}_0\right) \cosh\left(\frac{c\bar{\tau}}{\bar{x}_0}\right) , \tag{2.243b}$$

$$y = \bar{\zeta} , \tag{2.243c}$$

$$z = \bar{\eta} . \tag{2.243d}$$

Since the left side of equations (2.238) and (2.243) expresses the Minkowski coordinates of one and the same event, the right side of these equations should coincide. It result in the relationship between the Rindler coordinates of the two observers

$$\bar{\tau} = \frac{\bar{x}_0}{x_0} \tau , \qquad \bar{\xi} = \xi + (x_0 - \bar{x}_0) , \qquad \bar{\zeta} = \zeta , \qquad \bar{\eta} = \eta . \tag{2.244}$$

The relationship between the two Rindler coordinate grids is illustrated in Figure 2.12.

Since τ and $\bar{\tau}$ are the proper times of the observers which is physically measurable, equation (2.244) tells us that the physical time, τ, flows differently for differently-accelerating observers. It means that the observers are not able to maintain their clocks synchronized even if they were synchronized initially on some hypersurface, for example, on x-axis where $\bar{\tau} = \tau = 0$. Indeed, the second observer's worldline has the Rindler coordinates $\bar{\xi} = 0$, resulting in the value of $\bar{x}_0 = x_0 + \xi$. Equation (2.244) relating the proper times of two observers becomes

$$\bar{\tau} = \left(1 + \frac{\xi}{x_0}\right) \tau , \tag{2.245}$$

which shows that the clocks with $\xi > 0$ (acceleration $\bar{A} < A$) runs faster, and the clocks with $\xi < 0$ (acceleration $\bar{A} > A$) runs slower than the clock moving with acceleration A. This fact is used sometimes to explain the, so-called, "twin paradox" in special relativity as an effect caused by the acceleration of the traveling twin [Levi, 1967].

The Rindler coordinates are generalized in general theory of relativity to the Riemann normal coordinates and to the *Fermi normal coordinates* which are discussed in sections 3.3.8.3 and 3.8.2. Generalization of the Rindler coordinates to the case, when the "observer" is a massive, self-gravitating body is given in section 4.5. These coordinates have important practical applications in relativistic celestial mechanics of the solar system as stated in various resolutions of the International Astronomical Union (IAU) which are discussed in Chapter 9.

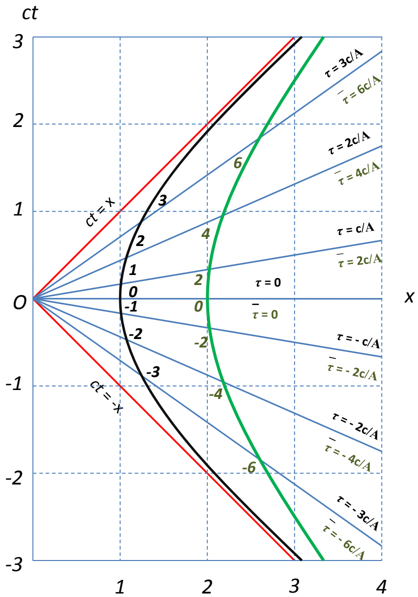

Figure 2.12 The local Rindler coordinates of two observers moving with a uniform acceleration A and $\bar{A}$ along x-axis. Worldlines of the observers are hyperbolas passing at distances $x_0 = c^2/A$ and $\bar{x}_0 = 2x_0$ from the origin of the Minkowski coordinates. Each point of spacetime has different values of the Rindler coordinates of the observers. Clocks of the two observers can not stay synchronized, and the lines of simultaneity have different values, τ and $\bar{\tau} = 2\tau$ in accordance with equation (2.244).

2.6.4 The radar coordinates

It is instructive to analyze the Einstein synchronization of clocks and the procedure of measuring distances conducted by a uniformly-accelerating observer by means of sending and receiving light (laser, radio) signals. Let the observer emits a light signal (photon) at the point E with coordinates $x^\alpha_E = (ct_1, x_1)$ in the static frame S toward another observer also moving along x-axis and keeping a constant Rindler coordinate ξ as measured by the first observer [25]. The light riches the second observer at point B, having coordinates $x^\alpha_B = (c\bar{t}, \bar{x})$ in frame S. It is immediately reflected back to the first observer, and arrives to him at point C with coordinates $x^\alpha_C = (ct_3, x_3)$. Figure 2.13 explains the relative positions of the points, and takes into account that light can propagate from point E either to an observer with positive Rindler coordinate $\xi > 0$ (reflection point B_+), or to an observer with negative Rindler coordinate $\xi < 0$ (reflection point B_-).

Let us consider, first, propagation of light from point E to point B_+ and, then, to point C_+. The proper time τ of the first observer at the instant of emission, t_1, and that of reception of the light signal, t_3, are calculated from equation (2.229) which yields

$$c\tau_1 = x_0 \ln\left[\frac{ct_1 + x_1}{x_0}\right] , \qquad c\tau_3 = x_0 \ln\left[\frac{ct_3 + x_3}{x_0}\right] , \tag{2.246}$$

where $x_0 = c^2/A$, and A is the proper acceleration of the observer. The observer's worldline is a spacetime hyperbola (2.227). Since the events E and C_+ are lying on this hyperbola, one has

$$x_1^2 - c^2t_1^2 = x_3^2 - c^2t_3^2 = x_0^2 . \tag{2.247}$$

On the other hand, the equations of propagation of light connecting the events E and C_+ with the event of reflection B_+ are null-cone equations

$$x_1 - ct_1 = \bar{x} - c\bar{t} , \qquad x_3 + ct_3 = \bar{x} + c\bar{t} . \tag{2.248}$$

Equations (2.247), (2.248) are employed for making algebraic transformation of equation (2.246) to a more symmetric form

$$c\tau_1 = -x_0 \ln\left[\frac{\bar{x} - c\bar{t}}{x_0}\right] , \qquad c\tau_3 = x_0 \ln\left[\frac{\bar{x} + c\bar{t}}{x_0}\right] . \tag{2.249}$$

The radar coordinate ϱ of point B_+ is defined as [Alba and Lusanna, 2007; Minguzzi, 2005] [26]

$$\varrho = \frac{c}{2}(\tau_3 - \tau_1) . \tag{2.250}$$

25) Because the light propagates between two observers along x-axis, the coordinates y and z remain constant and can be dropped out of the calculation.

26) The radar coordinates are also called the Märzke-Wheeler coordinates [Pauri and Vallisneri, 2000] or the emission coordinates [Coll and Pozo, 2006].

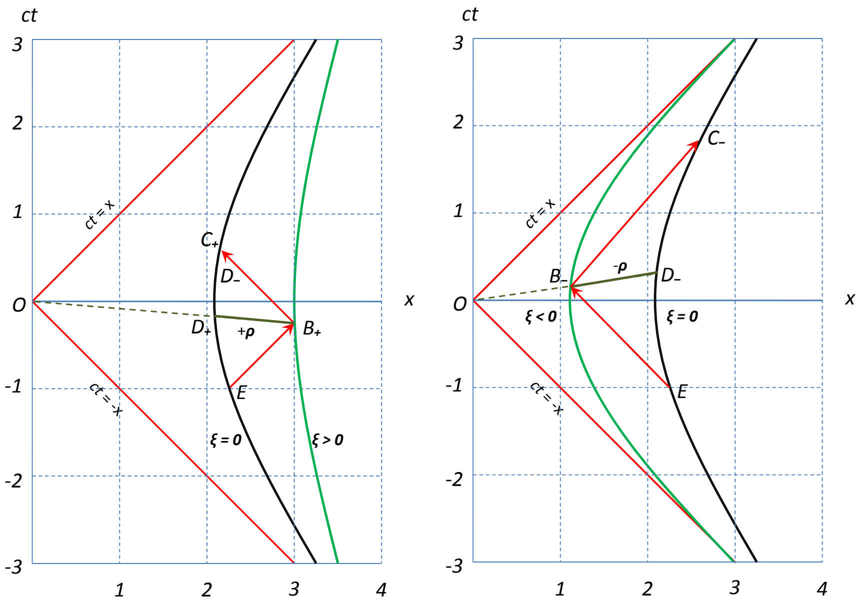

Figure 2.13 The radar coordinate ϱ of observer, moving with a uniform proper acceleration A along x-axis, is built with the help of light rays. The radar coordinate is counted as positive, $+\varrho$ (left diagram), in measuring the distance to an observer with the positive value of the Rindler coordinate, $\xi > 0$, and negative, $-\varrho$, otherwise (right diagram). Point D_+ (D_-) is simultaneous with the point of reflection B_+ (B_-) of the light ray according to Einstein's synchronization convention. The line passing through the points D_+ and B_+ on the left diagram (correspondingly D_- and B_- on the right diagram)is the line of the Rindler simultaneity as well.

It can be calculated with the help of equation (2.249) and the equation of the world-line of the second observer passing through the reflection point B_+:

$$\bar{x}^2 - c^2\bar{t}^2 = (x_0 + \xi)^2 \ . \tag{2.251}$$

The result is,

$$\varrho = x_0 \ln\left[1 + \frac{\xi}{x_0}\right] \ . \tag{2.252}$$

This equation relates the positive Rindler coordinate ξ to radar coordinate ϱ of the second observer. It is straightforward to see that for small distances both coordinates coincide, $\varrho = \xi + O(\xi^2)$. Since the Rindler coordinate ξ is the measure of the proper length ℓ measured by the comoving inertial observer, the radar coordinate ϱ provides an operational measurement of the proper length ℓ at small distances from the worldline of observer. However, the radar distance ϱ diverges logarithmically from the proper length ℓ for large values of the Rindler coordinate.

Equations of transformation from the global coordinates (ct, x) to the radar coordinates $(c\tau, \varrho)$ can be obtained directly from equations (2.238) and (2.252). One obtains,

$$ct = x_0 \exp\left(\frac{\varrho}{x_0}\right) \sinh\left(\frac{c\tau}{x_0}\right) \ , \tag{2.253a}$$

$$x = x_0 \exp\left(\frac{\varrho}{x_0}\right) \cosh\left(\frac{c\tau}{x_0}\right) \ , \tag{2.253b}$$

$$y = \zeta \ , \tag{2.253c}$$

$$z = \eta \ , \tag{2.253d}$$

Similar equations can be derived if one measures the radar coordinate of an observer with the negative value of the Rindler coordinate (path EB_-C_- in Figure 2.13). In this case, one defines the negative value of the radar coordinate by equation

$$\varrho = -\frac{c}{2}(\tau_3 - \tau_1) \ , \tag{2.254}$$

and replace equations (2.248) with

$$x_1 + ct_1 = \bar{x} + c\bar{t} \ , \qquad x_3 - ct_3 = \bar{x} - c\bar{t} \ . \tag{2.255}$$

Equations of transformation (2.253) will formally remain the same. Hence, one concludes that equations (2.253) are valid for any value of the coordinate ϱ from $-\infty$ to $+\infty$.

Calculation of the metric tensor in the radar coordinates by equation being similar to equation (2.239) yields the spacetime interval

$$ds^2 = \exp\left(\frac{2\varrho}{x_0}\right)\left(-c^2 d\tau^2 + d\varrho^2\right) + d\zeta^2 + d\eta^2 \ , \tag{2.256}$$

and determinant of the metric tensor

$$g = -\exp\left(\frac{4\varrho}{x_0}\right) . \tag{2.257}$$

These expressions are valid everywhere inside the right wedge of the null cone covered with the Rindler coordinates except of the null lines $ct = \pm x$. Though the radar coordinate ϱ ranges from $-\infty$ to $+\infty$, as contrasted to the Rindler coordinate ξ having a limited range of the negative values, it does not cover the entire Minkowski spacetime either. The reason is that only the events in the Rindler wedge, $-x < ct < x$, can admit two light beams passing through them to intersect the curve of the accelerated observer and defining both the process of emission of light and its reception. The observers with $ct > x$ lying outside of the Rindler horizon, $ct = x$, will receive light signal emitted by the accelerated observer but can not send it back to him because it would require the speed of the signal faster than the speed of light c. This limitation on the range of applicability of the radar coordinates in the Minkowski spacetime was noted by Bini *et al.* [2005] who also extended their range of applicability to the light signals propagating in arbitrary direction from the accelerated observer.

Einstein's synchronization convention is that a point D_+ taken on the worldline of the first observer at the proper time

$$\tau = \tau_1 + \frac{1}{2}(\tau_3 - \tau_1) = \frac{1}{2}(\tau_1 + \tau_3) , \tag{2.258}$$

is simultaneous with the event of the light's reflection B_+ (see the left diagram in Figure 2.13). Making use of the null-cone equations (2.249) and the definition of the hyperbolic motion (2.251), one can transform the above expression to

$$\tau = x_0 \ln\left[\frac{\bar{x} + c\bar{t}}{x_0 + \xi}\right] . \tag{2.259}$$

According to discussion following equation (2.243), the proper time $\bar{\tau}$ of the second observer at point B_+ is related to the inertial coordinates $(c\bar{t}, \bar{x})$ of this point by equation

$$c\bar{\tau} = (x_0 + \xi)\ln\left[\frac{\bar{x} + c\bar{t}}{x_0 + \xi}\right] , \tag{2.260}$$

where ξ is the Rindler coordinate of the observer at point B_+. Comparing equation (2.260) with expression (2.259) of the time τ at point D_+ one immediately derives

$$\bar{\tau} = \left(1 + \frac{\xi}{x_0}\right)\tau = \exp\left(\frac{\varrho}{x_0}\right)\tau , \tag{2.261}$$

which is, in fact, identical with equation (2.245). The same equation can be derived for the synchronization of clocks of the second observer with a negative Rindler coordinate as shown in the right diagram in Figure 2.13. Equation (2.261) reveals that the Einstein synchronization procedure places the point D_+ on the line of the

Rindler simultaneity (shown as a dashed line on the left diagram in Figure 2.13) of the second accelerated observer passing through point B_+. Similarly, the Einstein synchronization procedure places the point D_- of Einstein's simultaneity on the line of the Rindler simultaneity (shown as a dashed line on the right diagram in Figure 2.13) of the second accelerated observer passing through point B_-.

Equation (2.261) demonstrates in a different way that the accelerated observers being separated by some finite distance along x-axis, will gradually have their clocks desynchronized. In order to maintain the synchronization, the observers must constantly interchange light signals and apply the appropriate clock corrections to compensate the difference in their rates due to the unequal accelerations. One emphasizes that acceleration itself does not affect the proper time of any observer measured with an ideal clock. The desynchronization of clocks occur only because the two clocks have different accelerations causing the relative speed of the clocks vary as time goes on, thus, making the dilation of time effect (see section 2.5.5) dependent on time.

2.7 Relativistic Dynamics

So far, discussion of special theory of relativity has been focused on kinematics of the uniform motions of particles, propagation of light, and some aspects of physics of accelerated observers. This section deals with more essential part of special relativity – physical laws of dynamics that govern the ultra-relativistic motion of particles under the influence of forces exerted upon them. It requires relativistic generalization of the classic concepts of the linear momentum, force, energy, etc., that should agree with the group of the Lorentz transformations.

2.7.1 Linear momentum and energy

Special relativity demands that all physical quantities and laws be formulated in terms of conserved four-dimensional geometric objects - scalars, vectors, tensors. One of the basic quantities is the linear momentum of the particle. Its four-dimensional analogue must coincide with the Newtonian definition of the momentum, $\boldsymbol{p} = m\boldsymbol{v}$, in the slow-motion limit, where m is mass of the particle, and $\boldsymbol{v}$ is its three-dimensional velocity. For this reason, it is natural to define the four-momentum of particle, $p^\mu = (p^0, \boldsymbol{p})$, as the product of the four-velocity of the particle with its rest mass m, which is a constant referred to the rest frame of the particle,

$$p^\mu = mu^\mu = m\gamma_{\rm v}\,(c, \boldsymbol{v}) \ , \tag{2.262}$$

where $\gamma_{\rm v} = \left(1 - \beta_{\rm v}^2\right)^{-1/2}$ is the Lorentz factor, $\boldsymbol{\beta}_{\rm v} = \boldsymbol{v}/c$, and $\boldsymbol{v}$ is the three-dimensional velocity of the particle which may be a function of time depending on the worldline of the particle. Four-velocity is a vector being normalized as $u^\mu u_\mu = -c^2$. Hence, the magnitude of the four-momentum is constant,

$$p^\mu p_\mu = -\frac{\mathcal{E}^2}{c^2} + \boldsymbol{p}^2 = -m^2c^2 \ . \tag{2.263}$$

Four momentum is usually represented as a four-vector

$$p^\mu = (p^0, \boldsymbol{p}) = \left(\frac{\mathcal{E}}{c}, \boldsymbol{p}\right) \ , \tag{2.264}$$

whose temporal part, $p^0 = \mathcal{E}/c$, is identified with the relativistic energy of the particle,

$$\mathcal{E} = mc^2\gamma_{\rm v} = \frac{mc^2}{\sqrt{1 - \beta_{\rm v}^2}} \ , \tag{2.265}$$

and whose spatial part,

$$\boldsymbol{p} = m\gamma_{\rm v}\boldsymbol{v} = \frac{m\boldsymbol{v}}{\sqrt{1 - \beta_{\rm v}^2}} \ , \tag{2.266}$$

generalizes the Newtonian linear momentum, and coincides with it in the limit of $v/c \to 0$. Sometimes, the concept of relativistic mass

$$m_{\rm v} = m\gamma_{\rm v} = \frac{m}{\sqrt{1-\beta_{\rm v}^2}}\,, \tag{2.267}$$

is used in calculations. While the rest mass of the particle, m, is constant, the relativistic mass depends on particle's velocity and changes each time as the velocity changes. In terms of this mass, the spatial part of the four-momentum looks similar to the definition of the linear momentum in the Newtonian theory,

$$\boldsymbol{p} = m_{\rm v}\boldsymbol{v}\,, \tag{2.268}$$

while the relativistic energy of the particle becomes

$$\mathcal{E} = m_{\rm v}c^2\,. \tag{2.269}$$

that is one of the most famous Einstein's formulas expressing direct relationship between the relativistic mass and energy of the particle. Comparison of equations (2.268) and (2.269) yields, yet another relationship between the spatial components of four-momentum and energy,

$$\boldsymbol{p} = \frac{\mathcal{E}}{c}\boldsymbol{\beta}\,, \tag{2.270}$$

where $\boldsymbol{\beta} = \boldsymbol{v}/c$.

Sometimes, especially in particle physics, a kinetic energy K of a particle is required in calculations. The kinetic energy is defined as a difference between the total energy of the particle, $\mathcal{E}$, and its rest-mass energy, $\mathcal{E}_0 = mc^2$, that is

$$K = (\gamma_{\rm v} - 1)\,mc^2\,. \tag{2.271}$$

In ultra-relativistic regime, $\gamma_{\rm v} \gg 1$, and $K \simeq \mathcal{E}$. In the slow-motion approximation, when $v \ll c$, the kinetic energy $K \simeq (1/2)mv^2$, that matches with the expression for kinetic energy in the Newtonian physics.

Although originally Einstein regarded the relativistic mass $m_{\rm v}$ as one of the basic concepts of relativity, later he came to a conclusion that only the rest mass m has a fundamental physical meaning - see an excerpt from Einstein's unpublished letter to Barnett, cited by Adler [1987] and an excellent historic review by Okun [1989]. The rest mass is, indeed, a fundamental constant for a single, structureless particle like electron. However, it should be replaced with the effective rest mass in case of a composite particle (or a body) made up of many structureless particles. The reason is that one has to take into account the energy of interaction between the particles which leads to reduction of the effective mass - the effect known as the *mass defect*. Hence, the relativistic mass of the composite body makes sense especially in astronomy where one can measure only the effective masses of astronomical bodies. The issue remains subject to theoretical debates which at times become extremely heated [Khrapko, 2000; Okun, 2000].

Photons are relativistic particles with the rest mass equal to zero that always move with the fundamental speed c along straight lines on the null cone of the Minkowski spacetime. Four-momentum of photon is a null vector along the line of motion of the photon that satisfies relationship

$$p^\mu p_\mu = -\frac{\mathcal{E}^2}{c^2} + \boldsymbol{p}^2 = 0 \,. \tag{2.272}$$

This equation tells us that photon's energy and the absolute value of its three-momentum are equal up to the constant factor which is the fundamental speed c,

$$\mathcal{E} = |\boldsymbol{p}|c \,. \tag{2.273}$$

Photon's energy relates to its frequency by the Planck-Einstein equation, $\mathcal{E} = \hbar\omega$, where $\hbar = 1.05 \times 10^{27}$ erg·sec is Planck's constant h divided by 2π, $\omega = 2\pi\nu$ is the angular frequency of a monochromatic electromagnetic wave corresponding to the given photon, and ν is the electromagnetic frequency of the wave. Four-momentum of photon can not be defined as a time derivative of its worldline's coordinate with respect to the proper time because there is no "proper time" for photons. Taking equations (2.273) into account, the four-momentum of a photon can be written in terms of the four-vector

$$\mathcal{K}^\mu = \frac{\omega}{c} k^\mu \,, \tag{2.274}$$

of the electromagnetic wave as

$$p^\mu = \hbar \mathcal{K}^\mu \,, \tag{2.275}$$

where $k^\mu = (1, \boldsymbol{k})$ is a null vector which spatial components, $\boldsymbol{k}$, form a Euclidean unit vector pointing to the direction of propagation of the photon, $\boldsymbol{k} \cdot \boldsymbol{k} = 1$. Relationships (2.272)-(2.275) remain valid for any other massless particle moving with the fundamental speed c, for example, graviton - a hypothetical quantum of gravitational field.

2.7.2
Relativistic force and equations of motion

In the Newtonian mechanics equations of motion of a particle in an inertial frame of reference are expressed in terms of the Euclidean three-dimensional vectors and the absolute time. These equations are invariant with respect to the Galilean transformations. Special relativity operates on four-dimensional manifold of the Minkowski spacetime. Hence, the particle's equation of motion must be expressed in terms of four-vectors and obey the principle of the Lorentz covariance when transforming from one inertial frame to another with the help of the Lorentz transformation.

Let us consider motion of a single particle in an inertial reference frame S defined by a vector basis $\boldsymbol{e}_\alpha$ and covered with the Cartesian coordinates $x^\alpha = (x^0, x^i) = (ct, \mathbf{x})$.

The particle's worldline is given by the coordinates, $\mathbf{x} = x^\alpha \mathbf{e}_\alpha$, parameterized by the proper time τ taken on the worldline of the particle: $x^\alpha = x^\alpha(\tau)$. The Lorentz invariant equations of motion of a single particle are formulated in terms of the particle's linear momentum, $\mathbf{p} = p^\alpha \mathbf{e}_\alpha$, where $p^\alpha = mu^\alpha = mdx^\alpha/d\tau$, and the relativistic four-force, $\mathbf{F} = F^\alpha \boldsymbol{e}_\alpha$, exerted on the particle by an external environment and changing direction and magnitude of particle's linear momentum,

$$\frac{dp^\mu}{d\tau} = F^\mu \ . \tag{2.276}$$

Historically, the concept of the four-dimensional relativistic force, F^μ, was introduced in electrodynamics of moving charges by Poincaré [1906] by extending the concept of the three-dimensional, electromagnetic force $\boldsymbol{f}_L$ derived by H. Lorentz in 1892.

If the rest mass m of particle is a constant of motion, the Lorentz-covariant equation of motion (2.276) can be also formulated as follows

$$mw^\mu = F^\mu \ , \tag{2.277}$$

where $w^\mu = du^\mu/d\tau$ is four-acceleration of the particle. Equation (2.276) imposes a strong limitation on the mathematical nature of four-force. Indeed, multiplying both sides of equation (2.276) by four-momentum p^μ, and taking into account that p^μ is orthogonal to $dp^\mu/d\tau$ [27], one concludes that the four-force is orthogonal to the four-momentum of the particle,

$$p_\mu F^\mu = 0 \ , \tag{2.278}$$

This equation must be upheld for any relativistic force, F^μ. When four-momentum p^μ is proportional to four-velocity u^μ, like in case of a point-like particle, equation (2.278) is reduced to the condition of orthogonality of the four-force and four-velocity

$$u_\mu F^\mu = 0 \ . \tag{2.279}$$

It immediately tells us that the time component F^0 of the four-force exerted on the particle is uniquely expressed in terms of its spatial components, $(F^i) = \boldsymbol{F}$, as follows

$$F^0 = \boldsymbol{\beta}_{\rm v} \cdot \boldsymbol{F} \ , \tag{2.280}$$

where $\boldsymbol{\beta} = \boldsymbol{v}/c$, and $\boldsymbol{v} = d\boldsymbol{x}/dt$ is three-dimensional velocity of the particle with respect to frame S. Hence, only three, out of four components of four-force are algebraically independent. Another conclusion, which immediately follows from the condition of orthogonality (2.279) is that the four-force must be velocity-dependent, otherwise it will not obey the special relativity principle. Indeed, the assumption that

27) Remember that $p_\mu p^\mu = -m^2c^2$ according to equation (2.263).

the four-force depends on coordinates of the particle only, $F^\mu = F^\mu(x^\alpha)$, can be valid only in one particular inertial frame. Components of four-force in any other frame are obtained by the Lorentz transformation which is velocity-dependent. Moreover, the time component of four-force must vanish in the comoving frame of the particle as follows from (2.279). However, the force that depends only on coordinates of the particle can not comply with this requirement.

It is important to emphasize that not all three-dimensional forces of classic mechanics are associated with four-dimensional vectors. The most famous exception is the gravitational force (1.10) of the Newtonian law of the universal gravitation. The principle of equivalence, that is discussed in section 3.1, tells us that the gravitational force exerted on a freely-moving particle can be eliminated in the particle's comoving frame of reference that is in a free fall in the gravitational field. A four-vector, which components are nil in one frame must be identically equal to zero in any other reference frame. However, it apparently contradicts to existence of the universal gravitational attraction between massive bodies - planets to the Sun, Moon to the Earth, etc. - on the scale of the solar system. It indicates that the force of gravity should not be attributed to a four-vector but to a more complicated geometric object which can be eliminated locally at each point of the spacetime manifold but does not vanish on a global scale [28)]. In order to admit the generalization to a four-dimensional vector, a three-dimensional force must be localizable, that is it must not disappear in the local frame of reference comoving with particle. The electromagnetic and most other forces of physics hold this property.

In practical calculations, relativistic equations of motion of particles are often referred to the coordinate time t of the laboratory frame S rather than to the proper time τ of the particle which is not directly measurable quantity. The equations of motion, given in terms of the coordinate time, are obtained from (2.276) after replacing the differential of the proper time with that of the coordinate time, $dt = \gamma_{\rm v} d\tau$, and explicitly factorizing the four-force with the Lorentz factor,

$$F^\mu = \left(F^0, F^i\right) \equiv \gamma_{\rm v}\left(f^0, f^i\right) , \tag{2.281}$$

where $f^0 = \boldsymbol{\beta}_{\rm v} \cdot \boldsymbol{f}$, and $\gamma_{\rm v} = (1 - \beta_{\rm v}^2)^{-1/2}$ is the Lorentz factor associated with the velocity of motion of the particle, $\boldsymbol{v}$, with respect to the frame S. We shall call the relativistic three-dimensional vector, $\boldsymbol{f}$, the *Minkowski force* [29)]. Equation (2.276) can be projected, then, to the time and spatial dimensions, yielding

$$\frac{dK}{dt} = \boldsymbol{v} \cdot \boldsymbol{f} , \tag{2.282}$$

$$\frac{d\boldsymbol{p}}{dt} = \boldsymbol{f} , \tag{2.283}$$

28) The Newtonian force of gravity is generalized to the, so-called, affine connection on a four-dimensional spacetime manifold of general relativity. The affine connection is neither a vector nor a tensor as explained in section 3.4.

29) There is no consensus on what exactly is called the Minkowski force. Some textbooks attribute this name to the four-force, F^μ, others - to its spatial components $\boldsymbol{F}$ only.

where $K = \mathcal{E} - mc^2$ is the kinetic energy of the particle, $\boldsymbol{p} = m\gamma_{\rm v}\boldsymbol{v}$ is its three-dimensional linear momentum. The projected equations look similar to the equation of motion of a particle in classic mechanics. However, this similarity should not be misinterpreted as the physical meaning of quantities entering both sides of the equation of motion are different from their Newtonian analogues. One should remark that the kinetic-energy equation (2.282) is not independent from equation of motion (2.283). Indeed, differentiating the momentum $\boldsymbol{p}$ with respect to time and taking the Euclidean dot-product with particle's velocity, one obtains

$$\boldsymbol{v} \cdot \frac{d\boldsymbol{p}}{dt} = \frac{dK}{dt} , \tag{2.284}$$

which proves that equation (2.282) can be derived from equation (2.284).

Introducing the three-acceleration $\boldsymbol{a} = d\boldsymbol{v}/dt$, and taking a time derivative in the left side of equation (2.283), one can recast it to

$$m\gamma_{\rm v}\boldsymbol{a} + m\gamma_{\rm v}^3 (\boldsymbol{\beta}_{\rm v} \cdot \boldsymbol{a})\boldsymbol{\beta}_{\rm v} = \boldsymbol{f} . \tag{2.285}$$

Calculating a dot-product of this equation with $\boldsymbol{\beta}_{\rm v}$ yields

$$m\gamma_{\rm v}^3 (\boldsymbol{\beta}_{\rm v} \cdot \boldsymbol{a}) = \boldsymbol{\beta}_{\rm v} \cdot \boldsymbol{f} , \tag{2.286}$$

that allows us to convert equation (2.285) to a simple expression

$$m\gamma_{\rm v}\boldsymbol{a} = \boldsymbol{f} - (\boldsymbol{\beta}_{\rm v} \cdot \boldsymbol{f})\boldsymbol{\beta}_{\rm v} . \tag{2.287}$$

Equation (2.287) differs from its Newtonian counterpart by the presence of the Lorentz factor $\gamma_{\rm v}$ in the left side, and apparent relativistic correction, $(\boldsymbol{\beta}_{\rm v} \cdot \boldsymbol{f})\boldsymbol{\beta}_{\rm v}$, to the force $\boldsymbol{f}$ in its right side[30]. It makes evident that in special relativity one cannot define the inertial mass of the accelerating particle simply as the ratio of the three-dimensional force, $\boldsymbol{f}$, to the three-dimensional acceleration $\boldsymbol{a}$ as this ratio depends on the direction of the particle's velocity with respect to the force. Instead, the rest mass m of the particle must be used as an analogue of the inertial mass.

In the slow-motion limit, when $\beta \ll 1$, equation (2.287) can be approximated as

$$m\boldsymbol{a} = \left(1 - \frac{3}{2}\beta_{\rm v}^2\right)\boldsymbol{f} + \boldsymbol{\beta}_{\rm v} \times (\boldsymbol{f} \times \boldsymbol{\beta}_{\rm v}) + O\left(\beta^4\right) . \tag{2.288}$$

This approximation was important at the very beginning of experimental exploration of special relativity. However, modern high-energy accelerators of elementary particles reach a Lorentz factor $\gamma_{\rm v} \simeq 7500$ and speed-up the particles at about $\beta_{\rm v} = 0.999999991$ [Evans and Bryant, 2008] so that one can not use the post-Newtonian expansion (2.288) for computing. Instead, the exact equation (2.287) serves as a fundamental basis of particle's accelerator engineering. The very fact

30) One should keep in mind that the force $\boldsymbol{f}$ itself, is defined within the framework of relativistic theory and contains relativistic corrections in its own definition as well. The most obvious example is the Lorentz force in electrodynamics that besides the electric force contains the magnetic force which vanishes in the limit of vanishing $\beta_{\rm v}$.

that particle's accelerators - most importantly, super-colliders in Fermilab and CERN [Dremin, 2009], work in a fascinating agreement with equation (2.287) validates the special relativity with unprecedented precision, leaving no doubt in its solid physical foundation.

Equation (2.287) tells us that in relativity the three-acceleration of the particle $\boldsymbol{a}$ is, generally, not parallel to the applied three-force $\boldsymbol{f}$. Specifically, if the force is orthogonal to particle's velocity, one has $\boldsymbol{f}\cdot\boldsymbol{\beta}_{\rm v} = 0$, and equation (2.287) becomes

$$m\boldsymbol{a} = \frac{\boldsymbol{f}}{\gamma_{\rm v}} \,. \tag{2.289}$$

In case of the force being parallel to the velocity, $\boldsymbol{f}\cdot\boldsymbol{\beta}_{\rm v} = f\beta_{\rm v}$, and equation (2.287) yields

$$m\boldsymbol{a} = \frac{\boldsymbol{f}}{\gamma_{\rm v}^3} \,. \tag{2.290}$$

These equations reveal that it is more effective to accelerate the particles moving with ultra-relativistic speed along a circle than those moving along a straight line.

2.7.3
The relativistic transformation of the Minkowski force

In Newtonian mechanics, force is remains the same in all inertial reference frames connected to each other by the Galilean transformation. In special relativity the force does not remain the same, and changes its magnitude and direction in accordance with the Lorentz transformation. Let us consider an inertial reference frame S, where a particle has instantaneous velocity $\boldsymbol{v}$ and experiences the Minkowski three-force $\boldsymbol{f} = (f^i)$. Let another inertial frame S' move with respect to S with constant velocity $\boldsymbol{V}$. The particle moves in the frame S' with the instantaneous velocity $\boldsymbol{v}'$, and experiences force $\boldsymbol{f}' = (f^{i'})$. Four-vector of the force has components $F^{\mu'} = \gamma_{\rm v'}(f^{0'}, f^{i'})$ in frame S', and $F^{\mu} = \gamma_{\rm v}(f^0, f^i)$ in frame S, where the Lorentz factors $\gamma_{\rm v'} = (1-\boldsymbol{v}'^2/c^2)^{-1/2}$, and $\gamma_{\rm v} = (1-\boldsymbol{v}^2/c^2)^{-1/2}$. The four-force is transformed as four-vector

$$F^{\nu'} = \Lambda^{\nu'}{}_{\mu}F^{\mu} \,, \tag{2.291}$$

where the matrix of the Lorentz transformation $\Lambda^{\nu'}{}_{\mu}$ is given by equations (2.82). Taking into account definition (2.281), one finds out that the Lorentz transformation of spatial components of the Minkowski three-force reads

$$\gamma_{\rm v'}f^{i'} = \gamma_{\rm v}\left(\Lambda^{i'}{}_0 f^0 + \Lambda^{i'}{}_j f^j\right) \,, \tag{2.292}$$

where $f^0 = \boldsymbol{f}\cdot\boldsymbol{\beta}_{\rm v}$, and $\boldsymbol{\beta}_{\rm v} = \boldsymbol{v}/c$. Lorentz factor $\gamma_{\rm v'}$ is related to two other Lorentz factors, $\gamma = (1-V^2/c^2)^{-1/2}$ and $\gamma_{\rm v} = (1-v^2/c^2)^{-1/2}$, by equation (2.179) that yields

$$\gamma_{\rm v'} = \gamma\gamma_{\rm v}(1-\boldsymbol{\beta}_{\rm v}\cdot\boldsymbol{\beta}) \,, \qquad \gamma_{\rm v} = \gamma\gamma_{\rm v'}(1+\boldsymbol{\beta}_{\rm v'}\cdot\boldsymbol{\beta}) \,, \tag{2.293}$$

where $\boldsymbol{\beta} = \boldsymbol{V}/c$. Substituting the Lorentz matrix and the first relationship (2.293) in equation (2.292) bring about

$$\boldsymbol{f}' = \frac{\sqrt{1-\beta^2}}{1-\boldsymbol{\beta}_{\rm v}\cdot\boldsymbol{\beta}}\left[\boldsymbol{f} - \gamma(\boldsymbol{f}\cdot\boldsymbol{\beta}_{\rm v})\boldsymbol{\beta} + (\gamma-1)\frac{(\boldsymbol{f}\cdot\boldsymbol{\beta})\boldsymbol{\beta}}{\beta^2}\right] . \tag{2.294}$$

This equation represents the transformation of Minkowski three-force from the inertial frame S to S'. It is rather straightforward to prove that the inverse transformation reads

$$\boldsymbol{f} = \frac{\sqrt{1-\beta^2}}{1+\boldsymbol{\beta}_{\rm v'}\cdot\boldsymbol{\beta}}\left[\boldsymbol{f}' + \gamma(\boldsymbol{f}'\cdot\boldsymbol{\beta}_{\rm v'})\boldsymbol{\beta} + (\gamma-1)\frac{(\boldsymbol{f}'\cdot\boldsymbol{\beta})\boldsymbol{\beta}}{\beta^2}\right] , \tag{2.295}$$

where $\boldsymbol{\beta}_{\rm v'} = \boldsymbol{v}'/c$. Notice that the inverse transformation of the force can be achieved by making formal replacements, $\boldsymbol{\beta}_{\rm v} \to \boldsymbol{\beta}_{\rm v'}$ and $\boldsymbol{\beta} \to -\boldsymbol{\beta}$, in the direct transformation (2.294).

Notice that in the case when the inertial frame S' is the rest frame of the particle, $\boldsymbol{v}' = 0$, and $\boldsymbol{v} = \boldsymbol{V}$, that is $\boldsymbol{\beta}_{\rm v} = \boldsymbol{\beta}$. The force-transformation equations (2.294) and (2.295) become

$$\boldsymbol{f}' = \gamma\boldsymbol{f} - (\gamma-1)\frac{(\boldsymbol{f}\cdot\boldsymbol{\beta})\boldsymbol{\beta}}{\beta^2} , \tag{2.296}$$

$$\boldsymbol{f} = \frac{1}{\gamma}\boldsymbol{f}' + \left(1-\frac{1}{\gamma}\right)\frac{(\boldsymbol{f}'\cdot\boldsymbol{\beta})\boldsymbol{\beta}}{\beta^2} , \tag{2.297}$$

Relativistic equation of the force transformation (2.296) can be also derived from the transformation equation (2.188) for acceleration. Indeed, in the rest frame of the particle, equation of motion (2.287) is simple, and reduces to

$$m\boldsymbol{A} = \boldsymbol{f}' , \tag{2.298}$$

where $\boldsymbol{A} \equiv \boldsymbol{a}'$ denotes the proper acceleration of the particle in its rest frame. The acceleration $\boldsymbol{A}$ is transformed to the laboratory frame in accordance with equation (2.188), where acceleration $\boldsymbol{a}$ must be replaced with equation (2.287). It yields equation (2.296) as expected.

Be mindful of that equations (2.296) and (2.297) can not be applied to massless particles moving with the fundamental speed c like photons, gravitons, etc., because the speed c is invariant and remains the same in any inertial reference frame making existence of a rest frame for these particles impossible. Formally speaking, equations (2.296) and (2.297) are not applicable for the massless particles because both the Lorentz factors $\gamma_{\rm v}$ and $\gamma_{\rm v'}$ are divergent for $v \to c$ or $v' \to c$.

2.7.4 The Lorentz force and transformation of electromagnetic field

The most famous example of the relativistic three-dimensional force is given by the Lorentz force acting on a point charge,

$$\boldsymbol{f}_{\scriptscriptstyle L} = q\left(\boldsymbol{E} + \boldsymbol{\beta}_{\rm v}\times\boldsymbol{B}\right) , \tag{2.299}$$

where q is the electric charge [31], $\boldsymbol{E} = (E_x, E_y, E_z)$ and $\boldsymbol{B} = (B_x, B_y, B_z)$ are the electric and magnetic fields respectively, $\boldsymbol{\beta}_v = \boldsymbol{v}/c$ with $\boldsymbol{v}$ being velocity of the charge. Taking this expression into account, and introducing a four-velocity $u^\mu = \gamma_v(c, \boldsymbol{v})$ of the charge, expression (2.299) for the electromagnetic force after being substituted in equation (2.281), can be recast to the following Lorentz-invariant form

$$F^\mu = \frac{q}{c} F^\mu{}_\nu u^\nu \,, \tag{2.300}$$

where $F^\mu{}_\nu$ is the Faraday tensor of electromagnetic field made of the electric, $\boldsymbol{E}$, and magnetic, $\boldsymbol{B}$, vector fields,

$$F^\mu{}_\nu = \begin{bmatrix} F^0{}_0 & F^0{}_1 & F^0{}_2 & F^0{}_3 \\ F^1{}_0 & F^1{}_1 & F^1{}_2 & F^1{}_3 \\ F^2{}_0 & F^2{}_1 & F^2{}_2 & F^2{}_3 \\ F^3{}_0 & F^3{}_1 & F^3{}_2 & F^3{}_3 \end{bmatrix} = \begin{bmatrix} 0 & E_x & E_y & E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{bmatrix} . \tag{2.301}$$

Tensor $F^{\mu\nu} = F^\mu{}_\beta \eta^{\beta\nu}$ is anti-symmetric, $F^{\mu\nu} = -F^{\nu\mu}$, and has the following components

$$F^{\mu\nu} = \begin{bmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{bmatrix} . \tag{2.302}$$

Equations of motion (2.276) of a charged particle with the Lorentz force exerted upon it, acquire the following form

$$\frac{dp^\mu}{d\tau} = \frac{q}{c} F^\mu{}_\nu u^\nu \,. \tag{2.303}$$

The anti-symmetry of the Faraday tensor ensues that $F^\mu{}_\nu u_\mu u^\nu = F^{\mu\nu} u^\mu u^\nu \equiv 0$ as expected.

The Faraday tensor is also instrumental in formulation of the apparently Lorentz-invariant form of the Maxwell equations (2.16) of electromagnetic field. More specifically, the Maxwell equations written in terms of the Faraday tensor are

$$\partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} + \partial_\gamma F_{\alpha\beta} = 0 \,, \tag{2.304}$$

$$\partial_\beta F^{\alpha\beta} = \frac{4\pi}{c} j_e^\alpha \,, \tag{2.305}$$

where

$$j_e^\alpha = (j_e^0, j_e^i) = (c\rho_e, \boldsymbol{j}_e) \,, \tag{2.306}$$

is a four-vector of the electric current in case of a continuous distribution of charge, or

$$j_e^\alpha = q u^\alpha \,, \tag{2.307}$$

31) The electric charge is invariant under the Lorentz transformations.

in case of a single, point-like charge q [32]. Equation (2.304) is equivalent to the Maxwell equations (2.16b) and (2.16c) that tells us that the Faraday tensor is made of the partial derivatives of the electromagnetic four-potential, $A^\mu = (A^0, A^i) = (c\varphi, \boldsymbol{A})$,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \,, \tag{2.308}$$

where φ is an electric scalar potential, and $\boldsymbol{A}$ is a magnetic vector-potential. Equation (2.305) is a tensorial form of two other field equations (2.16a) and (2.16d).

It is customary to speak about the electric, $\boldsymbol{E}$, and magnetic, $\boldsymbol{B}$, fields as vectors as they appear in this form in the expression (2.299) for the Lorentz force. However, $\boldsymbol{E}$ and $\boldsymbol{B}$, are components of the Faraday tensor of the second rank and, hence, are not vectors. Therefore, though the Lorentz force is transformed as a vector under the Lorentz transformations, the electric and magnetic fields are transformed as components of the second-rank tensor,

$$F^{\mu'\nu'} = \Lambda^{\mu'}{}_{\alpha}\Lambda^{\nu'}{}_{\beta}F^{\alpha\beta} \,, \tag{2.309}$$

from an inertial frame S to S' moving with respect to S with a constant velocity $\boldsymbol{V}$. Making use of definition of the Faraday tensor (2.302) and the matrix (2.82) of the Lorentz transformation, the transformation of the electromagnetic field can be written

$$\boldsymbol{E}' = \gamma\,(\boldsymbol{E} + \boldsymbol{\beta}\times\boldsymbol{B}) - (\gamma - 1)\frac{(\boldsymbol{E}\cdot\boldsymbol{\beta})\boldsymbol{\beta}}{\beta^2} \,, \tag{2.310}$$

$$\boldsymbol{B}' = \gamma\,(\boldsymbol{B} - \boldsymbol{\beta}\times\boldsymbol{E}) - (\gamma - 1)\frac{(\boldsymbol{B}\cdot\boldsymbol{\beta})\boldsymbol{\beta}}{\beta^2} \,, \tag{2.311}$$

It shows that $\boldsymbol{E}$ and $\boldsymbol{B}$ are not independent. A purely electric or magnetic field in one inertial frame will appear as a mixture of electric and magnetic field in another frame.

2.7.5 The aberration of the Minkowski force

It is instructive to express the force-transformation equation (2.294) in a different form by introducing the velocity of motion of the particle in the inertial frame S' explicitly in this equation. It is achieved with the equation of the velocity transformation (2.176) that yields

$$\boldsymbol{f}\cdot\boldsymbol{\beta}_{\rm v} = \frac{\boldsymbol{f}\cdot\boldsymbol{\beta}_{\rm v'} + \gamma\boldsymbol{f}\cdot\boldsymbol{\beta} + \beta^{-2}(\gamma-1)(\boldsymbol{\beta}_{\rm v'}\cdot\boldsymbol{\beta})(\boldsymbol{f}\cdot\boldsymbol{\beta})}{\gamma(1+\boldsymbol{\beta}_{\rm v'}\cdot\boldsymbol{\beta})} \,. \tag{2.312}$$

Substituting this expression in equation (2.294), making use of the relationships (2.293) and reducing similar terms give us another expression for the force acting

32) In case of a system of charged particles, the four-current is a linear superposition of the currents of individual particles.

on the particle in the frame S' [Ohanian, 2001, Section 6.2]

$$f' = \gamma(1 + \beta_{v'} \cdot \beta) f - \gamma(f \cdot \beta_{v'})\beta - \frac{\gamma - 1}{\beta^2}(f \cdot \beta)\beta \,. \tag{2.313}$$

One reduces this equation to yet another form by making use of the double cross product of two vectors. Specifically, equation (2.313) can be brought to the following form

$$f' = f + \frac{\gamma^2}{\gamma + 1}\beta \times (f \times \beta) + \gamma\beta_{v'} \times (f \times \beta) \,, \tag{2.314}$$

which can be termed as the *aberration-of-force* equation. This equation clearly shows that the force in the new frame S' becomes velocity-dependent even if it was not such in the primary frame S. It is also clear that the velocity-dependent (aberrational) terms change both the magnitude and direction of the force. Furthermore, the force f' depends not only on the relativity velocity V of the frames but also on the velocity v' of the particle in the new reference frame. This circumstance is rarely emphasized.

In the slow-motion approximation, when parameters $\beta \ll 1$ and $\beta_{v'} \ll 1$, the difference between forces f' and f in two different inertial frames, is proportional to $\beta^2 = V^2/c^2$ and $\beta\beta_{v'} = Vv'/c^2$ after expansion of equation (2.314) in the Taylor series with respect to the small parameters. It makes the effect of relativistic aberration of force in the slow-motion approximation be always quadratic with respect to velocity and eliminates the so-called *Laplace effect* in the orbital motion of binary stars and planets in the solar system [Damour, 1987]. Laplace correctly assumed the gravity does not propagate instantaneously but with a finite speed [Laplace, 1805]. This assumption led him to the conclusion that the Newtonian force of gravity acting on the celestial bodies should be not purely radial but have a correction of the order V/c_g, where V is the relative speed between the bodies and c_g is the speed of gravity. Speaking modern language, the Laplace modification of the gravity force was an attempt to take into account the fact that the propagating gravity force connects positions of the moving bodies with retardation along the null-like directions made of the characteristics of the hyperbolic gravity field equation that should replace the elliptic-type Poisson equation for the gravitational potential. Extrapolating the Laplace' s idea, one says that the force of attraction must be a null vector, which Lorentz transformation is to look similar to equation (2.212) for light aberration containing a linear, with respect to V/c, term.

The Laplace approach to a finite speed of gravity inevitably leads to the gravitational torque exerted on the planet and leading to dissipation of orbital energy of the planet making its orbit unstable. The effect of the dissipation goes to zero linearly as $c_g \to \infty$ and, thence, can be used to limit the speed of gravity from observation of the planetary motions. As no effect was observed, Laplace concluded that the speed of gravitational interaction is at least 7×10^6 faster than the speed of light. Laplace's argument was critically reconsidered by Lorentz who noticed that the relativistic force is orthogonal to a timelike vector of four-velocity and, hence, must be always a spacelike vector. It led Lorentz to the conclusion that the aberration of force acting between two moving particles do exist but it must be of the quadratic order of V^2/c^2, not of the linear order V/c, as was argued by Laplace. This can be clearly seen from equation (2.296) after its expansion in the Taylor series with respect to V/c. Similar conclusion about the magnitude of the aberration of gravity force for a system of slowly-moving massive bodies is valid in general theory of relativity as discussed by Carlip [2000]. Lorentz's point of view on the aberration of force is currently accepted by all scientists.

Van Flandern [1998, 1999] had brought a peer attention to the experimental issue of measuring the aberration-of-gravity effect by emphasizing that the stability of the planetary orbits can be interpreted in many possible ways and among them, the infinite value of the speed of gravity is still a possible alternative. Van Flandern' arguments were subject to criticism on the basis that the linear aberrational effect of the order V/c does not exist due to a number of theoretical reasons [Carlip, 2000; Marsh and Nissim-Sabat, 1999]. However, all these reasons are applicable mostly to the particles moving slowly while the aberration-of-force formula (2.314) does not discard, in fact, the possibility to have the aberration-of-force effect of the order of V/c for particles that move with ultra-relativistic speed $\boldsymbol{v}'$ approaching the speed of light c. Indeed, in case of an ultra-relativistic particle, its velocity $\boldsymbol{v}' = \boldsymbol{k}'c$, where the unit vector $\boldsymbol{k}'$ is directed along an (almost) straight path of the particle. One can still assume the relative velocity $\boldsymbol{V}$ between the two frames, S and S', small. However, for the ultra-relativistic particle the last term in equation (2.314) is "amplified" making the difference between forces, $\boldsymbol{f}'$ and $\boldsymbol{f}$, of the first order in $\beta = V/c$,

$$\boldsymbol{f}' = \boldsymbol{f} + \frac{1}{c}\boldsymbol{k}' \times (\boldsymbol{f} \times \boldsymbol{V}) + O(c^{-2}) \,. \tag{2.315}$$

This theoretical conclusion remains valid in general theory of relativity as will be discussed in section 7.9. It can be used to measure the numerical value of the fundamental speed c entering the equations of general relativity - the *speed of gravity*. One have to notice that at the time being there is some disagreement among theorists about the concept of the speed of gravity. For example, Will [1993] identifies it with the speed of gravitational waves by *excluding* gravitomagnetic time-dependent gravitational phenomena having a non-wave character. Definition accepted in this book is less restrictive - the speed of gravity is the fundamental speed of Einstein's general theory of relativity that governs *all* time-dependent gravitational phenomena *including* the propagation of gravitational waves [Kopeikin, 2004; Kopeikin and Fomalont, 2006]. This is similar to Maxwell's electrodynamics where the "speed of light" c governs *all* time-dependent electromagnetic laws, not only the propagation of electromagnetic waves. In fact, Maxwell measured the fundamental speed c in the experiments with quasi-stationary magnetic fields. He discovered that c is equal to the speed of light that had been already known at that time independently from astronomical measurements. On the ground of the coincidence of the two speeds, Maxwell came to the conclusion that light is an electromagnetic wave – a great theoretical achievement.

2.7.6 The center-of-momentum frame

Let us consider equations of motion for a composite system made up of N particles. The particles can interact with each other via internal forces making the system self-bounded in space. One can call such a system as a composite body. It is also possible to have a system consisting of non-interacting particles like in physics of nuclear reactions in particle' accelerators. In such system particles move independently before and after the reaction and interact only during a short time interval when the reaction takes place.

Let the system consists of N particles where each particle has a four-momentum

$$p_a^\mu = \left(\frac{\mathcal{E}_a}{c}, \boldsymbol{p}_a\right) , \qquad (a = 1, 2, ..., N) . \tag{2.316}$$

One can formally define the total four-momentum of the system by equation

$$\mathcal{P}^\mu = \left(\frac{\mathcal{E}}{c}, \boldsymbol{\mathcal{P}}\right) , \tag{2.317}$$

where the left side is an algebraic sum of the four-momenta of individual particles,

$$\mathcal{P}^\mu = \sum_{a=1}^{N} p_a^\mu = p_1^\mu + p_2^\mu + ... + p_N^\mu , \tag{2.318}$$

measured with respect to an inertial reference frame S. This definition assumes that the relativistic energy and three-momentum of the ensemble are additive quantities [Rindler, 1960b]

$$\mathcal{E} = \sum_{a=1}^{N} \mathcal{E}_a = \mathcal{E}_1 + \mathcal{E}_2 + ... + \mathcal{E}_N , \tag{2.319}$$

$$\boldsymbol{\mathcal{P}} = \sum_{a=1}^{N} \boldsymbol{p}_a = \boldsymbol{p}_1 + \boldsymbol{p}_2 + ... + \boldsymbol{p}_N . \tag{2.320}$$

One draws the attention of the reader that summation in equations (2.318)-(2.320) is performed on a hyperplane of simultaneity of a particular reference frame S. If one takes another inertial frame of reference S', moving with respect to frame S, its hyperplane of simultaneity does not coincide with that of frame S. In case of uniformly moving particles the result of the summation will not depend on the particular choice of the hyperplane, making the total four-momentum $\mathcal{P}^\mu$ of the system of particles to be a four-vector defined unambiguously. However, if the particles in the system move with acceleration, their four-momenta p^μ, are functions of time. In this case, the total four-momentum $\mathcal{P}^\mu$ of the ensemble of particles depends on the choice of the hyperplane of simultaneity, and different observers may not agree on the overall value of $\mathcal{P}^\mu$ which will get dependent on the four-acceleration of the particles and their mutual displacements. This situation is typical in the relativistic dynamics of gravitating bodies, where the definition of the four-momentum requires more detailed investigation taking into account the mutual gravitational attraction of the bodies. We discuss it in more detail in section 6.1.3.

It is convenient to introduce a concept of the relativistic center-of-momentum frame (CM frame) of the system of particles. The center-of-momentum frame is defined as a frame in which the spatial components of the total momentum are zero,

$$\boldsymbol{\mathcal{P}} = \boldsymbol{\mathcal{P}}_{CM} = 0\,. \tag{2.321}$$

It is always possible to chose such a frame, though it may be not inertial. To this end, let us consider motion of each individual particle in the system that is affected by a four-force, F^μ_a $(a = 1, 2, ..., N)$. The force can be algebraically separated in two parts: internal $F^\mu_{\mathrm{int}\,a}$, and external, $F^\mu_{\mathrm{ext}\,a}$,

$$F^\mu_a = F^\mu_{\mathrm{int}\,a} + F^\mu_{\mathrm{ext}\,a}\,, \tag{2.322}$$

so that equation of motion of the a-th particle is

$$\frac{dp^\mu_a}{d\tau_a} = F^\mu_{\mathrm{int}\,a} + F^\mu_{\mathrm{ext}\,a}\,, \tag{2.323}$$

where τ_a is the proper time on the worldline of the particle. Because one considers motion of the system of particles, it is reasonable to use the coordinate time t instead of the proper time of each particle which differ from each other as the particles move with different velocities. In terms of the coordinate time, equation (2.323) becomes

$$\frac{dp^\mu_a}{dt} = f^\mu_{\mathrm{int}\,a} + f^\mu_{\mathrm{ext}\,a}\,, \tag{2.324}$$

where the normalized force $f^\mu = \gamma_a^{-1} F^\mu$ in accordance with definition (2.281). Equation of motion of the four-momentum of the system are obtained after taking a time derivative from expression (2.318) and applying equation (2.324),

$$\frac{dP^\mu}{dt} = f^\mu_{\mathrm{int}} + f^\mu_{\mathrm{ext}}\,, \tag{2.325}$$

where the net forces

$$f^\mu_{\mathrm{int}} = \sum_{a=1}^{N} f^\mu_{\mathrm{int}\,a}\,, \qquad\qquad f^\mu_{\mathrm{ext}} = \sum_{a=1}^{N} f^\mu_{\mathrm{ext}\,a}\,. \tag{2.326}$$

In special relativity the third Newton law is valid in case of contact forces [Rindler, 1960b]. If the forces act at distance (electromagnetic, gravitational) there is an issue of simultaneity of time and the third Newton law must be proven in each particular case. Nevertheless, the most fundamental physical principles dictate that an isolated system can not accelerate itself [33] and this effectively eliminates the net internal

33) Rocket motion is an exception but one does not consider it here. See [Ohanian, 2001, Section 5.6].

force, $f^\mu_{\rm int} = 0$. If, in addition to that, the total external force $f^\mu_{\rm ext}$ exerted on the system of particles is zero, the total four-momentum of the system is conserved,

$$\frac{dP^\mu}{dt} = 0\,, \tag{2.327}$$

and the total energy $\mathcal{E}$ and the four-momentum $\boldsymbol{\mathcal{P}}$ are constants of motion,

$$\mathcal{E} = \mathcal{E}_{CM}\,, \qquad \boldsymbol{\mathcal{P}} = \boldsymbol{\mathcal{P}}_{CM}\,. \tag{2.328}$$

By choosing the constant of motion $\boldsymbol{\mathcal{P}}_{CM} = 0$, one will have reach the goal of construction of the center-of-momentum frame.

Equations (2.327), (2.328) are used in particle's physics to calculate the reactions of disintegration or collision of the particles [Tsamparlis, 2010]. Specifically, if there is no interaction between the particles before and after a reaction, then, the initial value of the total four-momentum, $\mathcal{P}^\mu_{\rm before} = \sum_{b=1}^M p^\mu_b$, must be equal to its final value, $\mathcal{P}^\mu_{\rm after} = \sum_{a=1}^N p^\mu_a$,

$$\sum_{b=1}^M p^\mu_b = \sum_{a=1}^N p^\mu_a\,, \tag{2.329}$$

where N and M are the total number of particles before and after the reaction, and in general case, $M \neq N$. If the identity and number of particles before and after the reaction is preserved, the reaction is called *elastic*; otherwise it is called *inelastic*. In case of the elastic reaction both energy and three-momentum are conserved. In case of inelastic reaction, only the three-momentum is conserved, and the energy dissipates.

In the absence of the external forces, the law of conservation of the linear momentum (2.327) tells us that the center-of-momentum frame moves with respect to the inertial frame S with a constant three-velocity

$$\boldsymbol{v}_{CM} = \frac{\boldsymbol{\mathcal{P}}}{\mathcal{M}}\,, \tag{2.330}$$

where

$$\mathcal{M} = \frac{\mathcal{E}}{c^2}\,, \tag{2.331}$$

is the total rest mass of the particles that is constant, $\mathcal{E} = \mathcal{E}_{CM}$, in the case of an isolated system. One defines the four-velocity of the center-of-momentum frame as

$$u^\mu_{CM} \equiv \gamma_{CM}\left(c, \boldsymbol{v}_{CM}\right)\,, \tag{2.332}$$

where $\gamma_{CM} = \left(1 - v^2_{CM}/c^2\right)^{-1/2}$. Then, the total four-momentum can be written down as follows

$$\mathcal{P}^\mu = \mathcal{M}\left(c, \boldsymbol{v}_{CM}\right)\,, \tag{2.333}$$

or, equivalently,

$$\mathcal{P}^\mu = m_{CM} u^\mu_{CM} \ , \tag{2.334}$$

where the effective rest mass of the system of particles

$$m_{CM} \equiv \frac{\mathcal{M}}{\gamma_{CM}} \ , \tag{2.335}$$

relates to the effective rest energy in the standard manner,

$$\mathcal{E}_{CM} = m_{CM} c^2 \ . \tag{2.336}$$

In case of a closed system of N particles, velocities $\boldsymbol{v}_a = d\boldsymbol{x}_a/dt$ $(a = 1, 2, ..., N)$ of the individual particles are not so much different from the velocity $\boldsymbol{v}_{CM}$ of the center-of-momentum frame. It allows us to expand the effective energy in the power series with respect to the ratio of v'_a/c, where $\boldsymbol{v}'_a$ is the relative velocity of a-th particle with respect to the center-of-momentum frame. Indeed, in the slow-motion approximation the relative velocity, given by equation (2.174), can be approximated as $\boldsymbol{v}'_a = \boldsymbol{v}_a - \boldsymbol{v}_{CM}$, and the Taylor expansion of equation (2.336) with respect to v'_a/c yields

$$\begin{aligned} \mathcal{E}_{CM} &= \sum_{a=1}^{N} \left[\mathcal{E}_a + \left(\frac{\gamma_a}{\gamma_{CM}} - 1 \right) \mathcal{E}_a \right] \\ &= \sum_{a=1}^{N} E_a + \frac{1}{2} \sum_{a=1}^{N} m_a \boldsymbol{v}'^2_a + O\left(v'^4_a / c^2 \right) \ . \end{aligned} \tag{2.337}$$

This indicates that the effective rest energy of the composite body is not simply the sum of the rest energies of the individual particles but includes an additional term that is the overall kinetic energy of the particles in the center-of-momentum frame. The same reasoning, when applied to the effective rest mass, tells us that

$$m_{CM} = \sum_{a=1}^{N} m_a + \frac{1}{2c^2} \sum_{a=1}^{N} m_a \boldsymbol{v}'^2_a + O\left(v'^4_a / c^4 \right) \ , \tag{2.338}$$

which means that the effective mass of a composite body is larger than the sum of the rest masses of individual particles because of their motion inside the body. One should emphasize that in this (kinematic) consideration the potential energy effects caused by the internal interactions between the particles due to electromagnetic, gravitational, or any other forces, have been omitted. Taking into account the potential energy of those internal interactions will make the effective rest mass of the system of particles *smaller* than the sum of rest masses of particles. This effect is known in physics as the *defect of mass* . We discuss it in application to the relativistic dynamics of self-gravitating bodies in section 6.1.3.

In case, when the system of particles is not self-closed and interacts with an external environment, the overall four-force may be not equal to zero. Then, the spatial components of equation (2.325) describes the accelerated motion of the center-of-momentum frame of the system of particles with respect to the inertial frame,

$$\frac{d\mathcal{P}}{dt} = \boldsymbol{f}_{\rm ext} \ . \tag{2.339}$$

One can still introduce the center-of-momentum frame by making a coordinate transformation to a frame moving with some acceleration which introduces the inertial force to the right side of equation (2.339) canceling the external force of interaction of the system with the environment. Discussion of this complicated question is deferred to section 6.1.

2.7.7
The center-of-mass frame

The center-of-momentum frame is not the only frame associated with a system of particles. Another useful frame is identified with the center of mass of the system and is called the *center-of-mass frame*. In special relativity the center of mass is defined by a three-dimensional vector

$$\boldsymbol{x}_{CM} = \frac{1}{\mathcal{M}} \sum_{a=1}^{N} m_{\text{v}\,a} \boldsymbol{x}_a \ , \tag{2.340}$$

where $m_{\text{v}\,a}$ is the relativistic mass, and $\boldsymbol{x}_a = \boldsymbol{x}_a(t)$ is the spatial position of a-th particle taken on a hyperplane of simultaneity of an inertial frame S. The center-of-mass frame is defined by the condition

$$\boldsymbol{x}_{CM} = 0 \ . \tag{2.341}$$

If one can neglect possible time variations of the total mass $\mathcal{M}$ and accelerations of the particles, differentiation of equation (2.340) with respect to time will yield

$$\frac{d\boldsymbol{x}_{CM}}{dt} = \frac{\mathcal{P}}{\mathcal{M}} \ , \tag{2.342}$$

which tells us that the center of mass moves uniformly with respect to the frame in which the total momentum, $\mathcal{P}$, of the system of particles is constant. Moreover, by comparing equations (2.342) and (2.330), one sees that the velocity $d\boldsymbol{x}_{CM}/dt$ of the center of mass is equal to velocity, $\boldsymbol{v}_{CM}$, of the center-of-momentum frame. Hence, the center of mass is at rest in the center-of-momentum frame. This proves that the center-of-mass frame can be identified with the center-of-momentum frame after a constant displacement between the origins of the two frames is set to zero. If the accelerations and the time variation of the total mass $\mathcal{M}$ can not be neglected, the construction of the center-of-mass frame becomes more involved. It is discussed in section 6.1.3 of this book.

Equation (2.340) admits relativistic four-dimensional generalization in the Minkowski spacetime after introducing definition of the *angular momentum* of the system of particles,

$$L^{\mu\nu} = \sum_{a=1}^{N} \left(x_a^\mu p_a^\nu - x_a^\nu p_a^\mu \right) , \tag{2.343}$$

where $p_a^\mu = m_a u_a^\mu$ is a four-momentum of a-th particle with m_a being the rest mass of the particle, and $u_a^\mu = dx_a^\mu / d\tau_a$ is its four-velocity. Due to the action-reaction property of the internal forces, the net internal torque on the system must be equal to zero [34)]

$$\sum_{a=1}^{N} \left(x_a^\mu f_{\text{int}\,a}^\nu - x_a^\nu f_{\text{int}\,a}^\mu \right) = 0 . \tag{2.344}$$

Assuming that the external forces are absent and the particles move without acceleration, one can obtain the law of conservation of the angular momentum

$$\frac{dL^{\mu\nu}}{dt} = 0 , \tag{2.345}$$

where the differentiation is with respect to the coordinate time t of the inertial frame S under consideration. Relativistic center of mass appears in the components L^{0i} of the anti-symmetric tensor $L^{\mu\nu}$. Indeed, after making use of definitions of the total mass, $\mathcal{M}$, and the total momentum, $\mathcal{P} = (\mathcal{P}^i)$, of the system of particles, one obtains

$$\begin{aligned} L^{0i} &= \sum_{a=1}^{N} \left(x_a^0 p_a^i - x_a^i p_a^0 \right) = ct \sum_{a=1}^{N} p_a^i - c \sum_{a=1}^{N} m_{\text{v}\,a} x_a^i \\ &= c \left(\mathcal{P}^i t - \mathcal{M} x_{CM}^i \right) , \end{aligned} \tag{2.346}$$

which is equivalent to the first integral of the equation of motion of the center of mass (2.342),

$$\boldsymbol{x}_{CM} = \boldsymbol{v}_{CM} t + \boldsymbol{x}_0 , \tag{2.347}$$

where velocity of the center of mass, $\boldsymbol{v}_{CM} = \mathcal{P}/\mathcal{M}$, the constant displacement $\boldsymbol{x}_0 = \boldsymbol{L}/(\mathcal{M}c)$, and vector $\boldsymbol{L} = (L^{0i})$.

Accounting for equation (2.346), one can give a special-relativistic definition of the center-of mass frame as follows,

$$L^{\mu\nu} \mathcal{P}_\nu = 0 \tag{2.348}$$

where $\mathcal{P}^\nu$ is the total four-momentum of the system. Indeed, in the center-of-momentum frame, where $\mathcal{P} = 0$, definition (2.348) is reduced to the previous formula (2.346) with $L^{0i} = 0$, yielding $x_{CM}^i = 0$. Equation (2.348) is known as the Dixon-Tulczyjew supplementary condition [Dixon, 1979; Tulczyjew, 1959].

34) Strictly speaking, the vanishing property of the net internal torque must be proven.

2.8 Energy-Momentum Tensor

2.8.1 Non-interacting particles

In the Newtonian mechanics any continuous distribution of matter - a gas, a liquid, a solid - possess a mass density, and energy density, flux density, and stresses. Relativistic mechanics unite these quantities to a single object, called the *energy-momentum tensor* $T_{\alpha\beta}$. To see how it emerges in the mechanics of continuum let us first consider a simple case of a medium consisting of non-interacting particles (dust) moving with respect to an inertial reference frame $x^\alpha = (x^0, x^i)$. Each particle has its own four-velocity $u^\alpha = (u^0, u^i) = c\gamma(1, \beta^i)$, where the Lorentz factor $\gamma = (1 - \boldsymbol{\beta}^2)^{-1/2}$, and $\beta^i = v^i/c$ is a dimensionless velocity of the particle being normalized to the invariant speed c. Each particle is at rest with respect to its own comoving frame, $\xi^\alpha = (\xi^0, \xi^i) = (c\tau, \xi^i)$, where τ is the proper time on the worldline of the particle, and the particle's four-velocity $u^\alpha = (c, 0)$. Let us consider the proper mass density ϵ of the medium which is defined in the comoving frame as the ratio of energy ΔE_0 of the particles contained in three-dimensional element of the proper volume, $\Delta V_0 = d\xi^1 d\xi^2 d\xi^3$, to this volume,

$$\epsilon = \frac{\Delta E_0}{\Delta V_0} \ . \tag{2.349}$$

We define now the energy density

$$\varsigma = \frac{\Delta E}{\Delta V} \ , \tag{2.350}$$

measured by an observer being at rest in the inertial frame x^α. Our goal is to find out the transformation property of the energy density.

Three-dimensional volume is not invariant with respect to the Lorentz transformation, $\Delta V \neq \Delta V_0$. Real invariant in the Minkowski spacetime is a four-dimensional volume [35)]

$$\Delta\Omega = dx^0 dx^1 dx^2 dx^3 = d\xi^0 d\xi^1 d\xi^2 d\xi^3 \ . \tag{2.351}$$

One can easily establish the law of transformation of the three-volume after noticing that $dx^0 = (dx^0/d\tau)d\tau = u^0 d\xi^0/c$ along the worldline of the particle, where $u^0 = c\gamma$. Making use of this relationship in equation (2.351), one obtains that three-volume, $\Delta V = dx^1 dx^2 dx^3$, calculated in the inertial coordinates, x^α, relates to three-volume, ΔV_0, in the local frame, by equation

$$\Delta V_0 = \gamma \Delta V \ . \tag{2.352}$$

35) Transformation law of the four-volume from one inertial frame to another, must include the Jacobian of the Lorentz transformation, $J = \det\left[\partial\xi^\alpha/\partial x^\beta\right]$. Direct calculation of the Jacobian by making use of the matrix (2.82) of the Lorentz transformation reveals that $J = 1$. For this reason, the Jacobian is not shown in equation (2.351). In a more general case of curvilinear coordinates, the Jacobian, $J = \sqrt{-g}$, where $g = \det[g_{\alpha\beta}] < 0$ is the determinant of the metric tensor $g_{\alpha\beta}$.

On the other hand, the energy is transformed according to equation (2.265), that is

$$\Delta E_0 = \gamma^{-1}\Delta E \,. \tag{2.353}$$

Equations (2.352) and (2.353), together constitute the law of transformation of the energy density

$$\varsigma = \gamma^2 \epsilon \,. \tag{2.354}$$

that tells us that the energy density of matter is neither a scalar nor a vector but a component of a tensor of a second rank. It is identified with a time-time component, T^{00}, of a tensor of energy-momentum (see Figure **??**).

The energy density ϵ relates to the mass density ρ according to Einstein's mass-energy formula $\epsilon = \rho c^2$. However, the mass density in the Newtonian mechanics obeys the equation of continuity (1.32) which must be extrapolated to relativistic mechanics. This prompt us to identify the energy flux with the components T^{0i} of the energy-momentum tensor. The energy-momentum tensor must be symmetric,

$$T^{\alpha\beta} = T^{\beta\alpha} \,, \tag{2.355}$$

in order to comply with the law of conservation of the intrinsic angular momentum of the medium under consideration. This is related to the absence of intrinsic, uncompensated torques that, if existed, might bring the medium to self-rotation in violation of the third law of Newton [Landau and Lifshitz, 1975; Misner *et al.*, 1973]. Due to the symmetry of the energy-momentum tensor, the energy flux, T^{0i}, must be equated to the momentum density, T^{i0}, of the medium. Again, since the momentum density obeys in the Newtonian physics the macroscopic equations of motion (1.55), the components T^{ij} of the energy-momentum tensor should be identified with the linear superposition of the momentum flux and internal stresses in the medium. The energy-momentum tensor satisfies in the inertial coordinates, $x^\alpha = (x^0, \boldsymbol{x}) = (ct, \boldsymbol{x})$, the law of conservation

$$\partial_\beta T^{\alpha\beta} = 0 \,, \tag{2.356}$$

or more explicitly,

$$\frac{1}{c}\frac{\partial T^{00}}{\partial t} + \frac{\partial T^{0j}}{\partial x^j} = 0 \,, \tag{2.357}$$

$$\frac{1}{c}\frac{\partial T^{i0}}{\partial t} + \frac{\partial T^{ij}}{\partial x^j} = 0 \,. \tag{2.358}$$

Covariant expression for the energy-momentum tensor of a continuous medium consisting of non-interacting (dust) particles is given by

$$T^{\alpha\beta} = \rho u^\alpha u^\beta \,, \tag{2.359}$$

where ρ is the mass density (a scalar), and u^α is the medium's four-velocity. It is postulated that the density ρ and the three-velocity $v^i = u^i/u^0$ satisfy the Newtonian

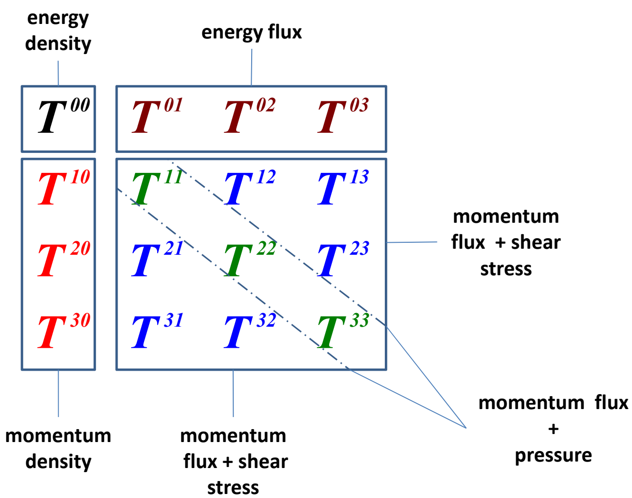

Figure 2.14 Physical meaning of the components of the stress-energy tensor $T^{\alpha\beta}$. Be aware that due to the symmetry of the tensor, $T^{\alpha\beta} = T^{\beta\alpha}$, the momentum density, T^{i0}, is equal to the energy flux T^{0i}. Off-diagonal space-space components represent a linear combination of anisotropic momentum flux and stress. Diagonal space-space components describe isotopic pressure and momentum flux.

equation of continuity (1.32) exactly. The energy-momentum tensor of dust has no stresses or pressure. Its trace is negative

$$T^{\alpha}{}_{\alpha} = \eta_{\alpha\beta} T^{\alpha\beta} = -\rho c^2 < 0 \,, \tag{2.360}$$

because the rest mass density is always positive.

One can easily check that tensor (2.359) satisfies the law of conservation (2.356). Indeed, applying this law to equation (2.359), one obtains

$$\partial_\beta \left(\rho u^\alpha u^\beta\right) = u^\alpha \partial_\beta (\rho u^\beta) + \rho u^\beta \partial_\beta u^\alpha \,. \tag{2.361}$$

Here the second term in the right side is proportional to a four-acceleration of the particle $w^\alpha = u^\beta \partial_\beta u^\alpha$ as introduced in equation (2.165). Since the particles do not interact the four-acceleration $w^\beta = 0$. The first term in the right side of equation (2.361) can be simplified after expressing four-velocity u^i in terms of three velocity, $u^i = u^0 v^i$, and using the Newtonian equation of continuity ((1.32)). It yields

$$u^\alpha \partial_\beta (\rho u^\beta) = u^\alpha u^\beta \partial_\beta u^0 = u^\alpha w^0 = 0 \,, \tag{2.362}$$

because all components of the four-acceleration, w^α, are zero for non-interacting particles.

Particular interest represents a tensor of energy-momentum of an isolated particle with mass m moving along trajectory $\boldsymbol{x}_p(t)$ with velocity $\boldsymbol{v}_p = d\boldsymbol{x}_p/dt$. In special relativity, this tensor is given by expression [Landau and Lifshitz, 1975; Misner *et al.*, 1973]

$$T^{\alpha\beta} = m\gamma^{-1} u^\alpha u^\beta \delta^{(3)} \left[\boldsymbol{x} - \boldsymbol{x}_p(t)\right] \,, \tag{2.363}$$

where $\gamma = \left(1 - v_p^2/c^2\right)^{-1/2}$, $u^\alpha = \gamma(c, \boldsymbol{v}_p)$ is the particle's four-velocity with $\boldsymbol{v}_p = \boldsymbol{v}_p(t) = d\boldsymbol{x}_p/dt$, and $\delta^{(3)}(\boldsymbol{x})$ is three-dimensional Dirac's delta function that is defined by the condition that integration of a smooth function $f(t, \boldsymbol{x})$ over the entire space yields

$$\int_{R^3} \delta^{(3)} \left[\boldsymbol{x} - \boldsymbol{x}_p(t)\right] f(t, \boldsymbol{x}) d^3x = f\left[t, \boldsymbol{x}_p(t)\right] \,. \tag{2.364}$$

If one uses four-dimensional Dirac's delta-function

$$\delta^{(4)} \left[x - x_p(\tau)\right] = \delta\left[t - t_p(\tau)\right] \delta^{(3)} \left[\boldsymbol{x} - \boldsymbol{x}_p(\tau)\right] \,, \tag{2.365}$$

where $t_p(\tau)$ is the coordinate time and τ is the proper time on the worldline of the particle, tensor (2.363) can be written in the form

$$T^{\alpha\beta} = m \int u^\alpha(\tau) u^\beta(\tau) \delta^{(4)} \left[x - x_p(\tau)\right] d\tau \,, \tag{2.366}$$

where the integral is taken along the particle's worldline. In case of a system of N particles with masses m_i ($i = 1, 2, ..., N$), the energy-momentum tensor of the system is

$$T^{\alpha\beta} = \sum_{i=1}^{N} m_i \int u^\alpha(\tau_i) u^\beta(\tau_i) \delta^{(4)} \left[x - x_p(\tau_i)\right] d\tau_i \,, \tag{2.367}$$

that is a linear superposition of integrals taken along the worldlines of the particles with τ_i being the proper time of i-th particle. Tensor (2.367) will be used in section 7.2 in order to find out a gravitational field of a moving massive body (star, planet, etc.) by solving Einstein's field equations.

2.8.2
Perfect fluid

One can move on and consider more complicated example of the energy-momentum tensor of a continuous medium that cannot sustain a tangential or shearing force. Such medium is known as a perfect (or isentropic, or ideal) fluid which entropy does not change as the fluid moves. This is because the absence of shear stresses does not allow elements of the fluid to exchange heat. Perfect fluid can be completely characterized by its rest frame energy density, ϵ, and isotropic pressure, p. Energy-momentum tensor of a perfect fluid is given in any inertial coordinates by

$$c^2 T^{\alpha\beta} = (\epsilon + p)u^\alpha u^\beta + c^2 p\eta^{\alpha\beta} , \tag{2.368}$$

where $\eta^{\alpha\beta}$ is the reciprocal Minkowski metric tensor, and p is an isotropic pressure related to ϵ by an equation of state $p = p(\epsilon)$, which is postulated separately. Perfect fluids are often used in general relativity to model idealized distributions of matter, such as in the interior of a star or in cosmology. This model of matter is also a good approximation for solving various problems of relativistic celestial mechanics in the solar system [Brumberg, 1972, 1991; Will, 1993]. If fluid is at rest, the energy-momentum tensor (2.368) is reduced to a diagonal form

$$T^{\alpha\beta} = \mathrm{diag}[\epsilon, p, p, p] , \tag{2.369}$$

which elucidates that the perfect fluid has no tangential stresses.

Tensor (2.368) is often presented not in terms of the rest energy density, ϵ, but in terms of the rest mass energy, ρc^2, and the energy of compression, Π, per unit mass. The compression energy, Π, is caused by pressure p, and vanishes in a pressureless fluid. The overall rest energy density is a linear superposition of the rest mass energy of fluid's particles and the compression energy per unit mass,

$$\epsilon = \rho(c^2 + \Pi) . \tag{2.370}$$

After substitution this equation to definition (2.368) the energy-momentum tensor acquires a new form [Fock, 1964]

$$T^{\alpha\beta} = \rho\left(1 + \frac{\Pi}{c^2}\right)u^\alpha u^\beta + \frac{p}{c^2}\left(u^\alpha u^\beta + c^2\eta^{\alpha\beta}\right) . \tag{2.371}$$

Pressure is related to the mass density by *equation of state*

$$p = p(\rho) , \tag{2.372}$$

which is a smooth, fully invertible function. In order to find out an equation relating the compression energy Π to ρ and p, let us apply the law of conservation (2.356) to tensor (2.371), and use the equation of continuity

$$\partial_\beta(\rho u^\beta) = u^\beta \partial_\beta \rho + \rho \partial_\beta u^\beta = 0 \,. \tag{2.373}$$

One obtains

$$\rho u^\alpha u^\beta \partial_b \Pi + \rho(c^2 + \Pi) w^\alpha + \left(u^\alpha u^\beta + c^2 \eta^{\alpha\beta}\right) \partial_\beta p + p\left(w^\alpha + u^\alpha \partial_\beta u^\beta\right) = 0 \,, \tag{2.374}$$

where $w^\alpha = u^\beta \partial_\beta u^\alpha$ is fluid's four-acceleration, which does not vanish since the non-zero pressure p means that fluid's particle are self-interacting. One can simplify equation (2.374) by projecting it on the direction of (that is contracting it with) four-velocity u^α. Taking into account that u_α is orthogonal to both w^α and $u^\alpha u^\beta + c^2\eta^{\alpha\beta}$, the projection yields

$$\rho u^\beta \partial_b \Pi + p \partial_\beta u^\beta = 0 \,. \tag{2.375}$$

Four-velocity divergence, $\partial_\beta u^\beta$, in this equation can be replaced by making use of equation of continuity (2.373) that is

$$\partial_\beta u^\beta = -\frac{1}{\rho} u^\beta \partial_\beta \rho \,. \tag{2.376}$$

Finally, noticing that the derivative $u^\beta \partial_b$ is just a derivative $d/d\tau$ with respect to a proper time *tau* taken along the worldline of fluid's particle, equation (2.375) is reduced to a relation between differentials

$$d\Pi - \frac{p}{\rho^2} d\rho = 0 \,, \tag{2.377}$$

which is nothing else but the second law of thermodynamics applied to a compressible perfect fluid [36]. Integration of this equation yields

$$\Pi = -\frac{p}{\rho} + \int_0^p \frac{dp}{\rho} \,, \tag{2.378}$$

where the integration constant was fixed by the condition $\Pi = 0$ for $p = 0$.

2.8.3
Non-perfect fluid and solids

Real matter composing planets and stars is not made of a perfect fluid. There are strong dissipation processes, heat exchange, tangential stresses, and other non-equilibrium processes going on in the interior of celestial bodies. Tensor of energy-momentum of such non-ideal matter has more complicated form and is characterized

36) In the rest frame of fluid the unit mass of fluid element is a constant of motion, and the increment of density, $d\rho$, is related to the increment of volume, dV, by a simple relationship $d\rho = -\rho^2 dV$. It brings equation (2.377) to a canonical thermodynamic form $d\Pi + pdV = 0$ for isentropic medium.

by several other parameters. Without going into details of derivation, which can be found in [Fock, 1964; Landau and Lifshit's, 1959; Weinberg, 1972] the most general form of the energy-momentum tensor is given by

$$T^{\alpha\beta} = \rho\left(1 + \frac{\Pi}{c^2}\right)u^\alpha u^\beta + \sigma^{\alpha\beta}\;, \tag{2.379}$$

where ρ is the rest mass density, and Π is the compression energy per unit mass, and $\sigma^{\alpha\beta}$ is the symmetric stress tensor that is orthogonal to the four-velocity

$$\sigma^{\alpha\beta}u_\beta = 0\;. \tag{2.380}$$

Specific form of $\sigma^{\alpha\beta}$ depends on the particular type of the medium, and can be fairly complicated. We shall use tensor (2.379) for derivation post-Newtonian equations of motion of celestial bodies in section 6. It is remarkable that one can pursue the derivation without specification of $\sigma^{\alpha\beta}$ that makes the post-Newtonian equations of motion valid for any system of astronomical bodies.

2.8.4
Electromagnetic field

Electromagnetic field is a vector field (tensor field of rank one) described classically by a vector potential A^α. In quantum electrodynamics the field description is replaced with spin-1 particles, called photons. Electromagnetic energy-momentum tensor is build out of the Faraday tensor of electromagnetic field, $F^{\alpha\beta}$, that was introduced in section 2.7.4. One has

$$4\pi T^{\alpha\beta} = F^{\alpha\mu}F^{\beta}_{\mu} - \frac{1}{4}\eta^{\alpha\beta}F_{\mu\nu}F^{\mu\nu}\;. \tag{2.381}$$

Employing equation (2.301) allows us to write down the components of $T^{\alpha\beta}$ explicitly in terms of electric, $\boldsymbol{E} = (E^i)$, and magnetic, $\boldsymbol{B} = (B^i)$, fields:

$$T^{00} = \frac{1}{8\pi}\left(\boldsymbol{E}^2 + \boldsymbol{B}^2\right)\;, \tag{2.382a}$$

$$T^{0i} = T^{i0} = \frac{1}{4\pi}\left(\boldsymbol{E}\times\boldsymbol{B}\right)^i\;, \tag{2.382b}$$

$$T^{ij} = \frac{1}{4\pi}\left[-\left(E^iE^j + B^iB^j\right) + \frac{1}{2}\left(\boldsymbol{E}^2 + \boldsymbol{B}^2\right)\delta^{ij}\right]\;, \tag{2.382c}$$

The component T^{00} is the energy density $\epsilon = (\boldsymbol{E}^2 + \boldsymbol{B}^2)/8\pi$ of electromagnetic field, and T^{i0} is the *Pointing flux*. The spatial components T^{ij} of the energy-momentum tensor of electromagnetic field give (with opposite sign) the *Maxwell stress tensor* [Jackson, 1998]. If one chooses the inertial coordinates in which the electric and magnetic fields are parallel and directed, let say, along x axis, the Maxwell tensor is reduced to a diagonal form with a tension (negative pressure) of the electromagnetic field, $(\boldsymbol{E}^2 + \boldsymbol{B}^2)/8\pi$, along the field lines and a (positive) pressure, $(\boldsymbol{E}^2 + \boldsymbol{B}^2)/8\pi$, perpendicular to the field lines

$$T^{\alpha\beta} = \mathrm{diag}[\epsilon, -\epsilon, \epsilon, \epsilon]\;. \tag{2.383}$$

If yet another inertial frame is chosen, where the electric and magnetic fields are equal and perpendicular to each other, and axis x is orthogonal to their plane, the energy-momentum tensor is not diagonal. It has components

$$T^{00} = T^{01} = T^{10} = T^{11} = \epsilon \,, \tag{2.384}$$

with all other components being nil. Tensor of energy-momentum of electromagnetic field is traceless

$$T^{\alpha}{}_{\alpha} = 0 \,, \tag{2.385}$$

that is a characteristic feature of any massless field moving with the speed of light.

The energy-momentum tensor of electromagnetic field is conserved, if and only if, there are no electric charges in the space occupied by the electromagnetic field. In case, when there are electric charges the divergence of the tensor becomes equal to the Lorentz force exerted on the charges. Indeed. taking the divergence from tensor (2.381) yields

$$\partial_\beta T^{\alpha\beta} = \frac{1}{4\pi}\left[\partial_\beta F^{\alpha\mu} F^{\beta}{}_{\mu} + F^{\alpha\mu}\partial_\beta F^{\beta}{}_{\mu} - \frac{1}{2}\eta^{\alpha\beta}\partial_\beta F_{\mu\nu}F^{\mu\nu}\right] . \tag{2.386}$$

After lowering the free index with the Minkowski metric tensor, and making use of anti-symmetry of the Faraday tensor, the above equation can be recast to the following form

$$\partial^\beta T_{\alpha\beta} = -\frac{1}{4\pi}\left[F_{\alpha\mu}\partial_\beta F^{\mu\beta} - \frac{1}{2}F^{\mu\nu}\left(\partial_\beta F_{\mu\nu} + \partial_\mu F_{\nu\beta} + \partial_\nu F_{\beta\mu}\right)\right] . \tag{2.387}$$

However, the term enclosed to the round parentheses, vanishes identically due to the Maxwell equation (2.304). The remaining term can be expressed in terms of the electric four-current j_e^α by using the Maxwell equation (2.305). We conclude that

$$\partial_\beta T^{\alpha\beta} = -\frac{1}{c}F^{\alpha}{}_{\beta} j_e^{\beta} \,, \tag{2.388}$$

coincides with the Lorentz force (2.300).

Though the energy-momentum tensor of electromagnetic field does not conserve in the presence of electric charge, it does not mean that there is a violation of the law of conservation. The charges are particles having masses so that the tensor of energy-momentum of the field+charge system must be described by a linear superposition consisting of the tensors of energy-momentum of the field and the particles. It leads to the equations of motion of charges (2.19).

2.8.5 Scalar field

A scalar field is a tensor field of rank zero. In classical physics, there are many examples of scalar fields such as the Newtonian gravitational potential or the electric potential in electrostatics. A temperature, humidity or pressure are also scalar fields. In quantum field theory, a scalar field is associated with spin-0 particles. The most famous examples includes the Higgs boson as well as the charged pion mediating the strong nuclear interaction. In the Standard Model of elementary particles, a scalar Higgs field is used to give the leptons and massive vector bosons their mass, via a combination of the Yukawa interaction and the spontaneous symmetry breaking. This mechanism is known as the Higgs mechanism [Higgs, 1964]. The Higgs boson is the subject of searches at particle accelerators like Tevatron in Fermilab and the Large Hadron Collider (LHC) near Geneva, Switzerland. Scalar fields are supposed to cause the exponentially-accelerated expansion of the universe known as inflation [Guth, 1981], helping to solve the *horizon problem* and giving a phenomenological reason for the non-vanishing *cosmological constant* of cosmology.

Scalar fields can represent the gravitational interaction mixed up with a tensor field of rank two which is the metric tensor in Einstein's general relativity. Such scalar-tensor theories are alternatives to general theory of relativity where the only field that mediates gravity is the metric tensor. Examples of the scalar-tensor theories are the Jordan-Fiertz theory [Fierz, 1956; Jordan, 1949, 1959] and the Brans-Dicke theory [Brans and Dicke, 1961] as well as various multi-dimensional generalizations of the Kaluza-Klein theory [Overduin and Wesson, 1997; Wesson, 2000]. Had a long-range scalar mediator of gravity been present in nature, it might lead to a violation of the principle of equivalence between gravitational force and the force of inertia. It would change the Einstein-Infeld-Hoffmann (EIH) equations of motion [Einstein *et al.*, 1938] of massive bodies in the solar system [37)]. Detection of these hypothetical violations is one of the primary goals of relativistic celestial mechanics in the solar system and in binary systems with compact astrophysical objects like neutron stars and/or black holes. Taking into account the importance of scalar fields for modern gravitational physics and for facilitating discussion in section (4.2.1 of this book, a brief description of the energy-momentum tensor of a scalar field is given below.

Exact form of the energy-momentum of scalar field depends on a particular theory. The most simple case is a classical scalar field ϕ with potential $V(\phi) > 0$ in the Minkowski spacetime (no gravity). The energy-momentum tensor of such a field is given by the expression [Mukhanov, 2005, page 21]

$$T_{\alpha\beta} = \partial_\alpha\phi\partial_\beta\phi - \frac{1}{2}\eta_{\alpha\beta}\partial^\mu\phi\partial_\mu\phi + \eta_{\alpha\beta}V(\phi)\ . \tag{2.389}$$

Taking divergence of the energy-momentum tensor, leads to the scalar field equations

$$\partial_\beta T^{\alpha\beta} = \Box\,\phi + \frac{\partial V}{\partial\phi} = 0\ , \tag{2.390}$$

where the second-order differential operator of partial derivatives

$$\Box \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu = -\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \triangle\ , \tag{2.391}$$

37) EIH force of gravity depends on two parameters, β and γ, characterizing the coupling of the scalar field with matter and gravity as shown in equation (6.82).

is known as the D'Alambert operator describing propagation of the field in spacetime with the fundamental speed c [38].

It is remarkable that if $\partial^\mu\phi\partial_\mu\phi < 0$, then the energy-momentum tensor for a scalar field can be reformulated in the form (2.368) of a perfect fluid by defining

$$\epsilon = -\frac{1}{2}\partial^\mu\phi\partial_\mu\phi + V(\phi)\,, \tag{2.392a}$$

$$p = -\frac{1}{2}\partial^\mu\phi\partial_\mu\phi - V(\phi)\,, \tag{2.392b}$$

$$u^\alpha = \frac{\partial^\alpha\phi}{\sqrt{-\partial^\mu\phi\partial_\mu\phi}}\,. \tag{2.392c}$$

In particular, assuming that the field is spatially homogeneous ($\partial_i\phi = 0$), one has [39]

$$\epsilon = \frac{1}{2}\dot{\phi}^2 + V(\phi)\,, \qquad p = \frac{1}{2}\dot{\phi}^2 - V(\phi)\,, \tag{2.393}$$

where the time derivative $\dot{\phi} = d\phi/dt$. Equation (2.393) implies the weak energy dominance condition, $\epsilon + p \geq 0$.

The energy-momentum tensor of a scalar field in the presence of gravity was introduced by Brans and Dicke [1961] and has the following form

$$T_{\alpha\beta} = \frac{\theta(\phi)}{\phi}\left(\partial_\alpha\phi\partial_\beta - \frac{1}{2}\eta_{\alpha\beta}\partial^\mu\phi\partial_\mu\phi\right) + \partial_{\alpha\beta}\phi - \eta_{\alpha\beta}\,\Box\,\phi\,, \tag{2.394}$$

where $\partial_{\alpha\beta} \equiv \partial_\alpha\partial_\beta$, and the constant $\theta(\phi)$ is a dimensionless coupling function which is to be determined from experiment. The boundary conditions are chosen such that as $\theta(\phi) \to \infty$ the scalar field $\phi \to$ const. thus making $T^{\alpha\beta}$ effectively nil. Brans-Dicke tensor of energy-momentum for a scalar field can be further generalized to include a potential $V(\phi)$ and the coupling of the scalar field with gravity. We discuss it in section 4.2.1.

Santiago and Silbergleit [2000] argued that terms with the second derivatives on the right side of equation (2.394) should not be included to the energy-momentum tensor of the scalar field. These terms originate in the Brans-Dicke theory from variation of the gravitational part of the action, ϕR, after an integration by parts. Hence they form a part of the dynamical description of gravity, and not of the scalar field alone. They occur because the dynamics of gravity and that of the purely scalar excitations are entangled in the physical (Jordan-Fierz) frame, as a result of the non-minimal coupling between gravity and the scalar field. Santiago and Silbergleit [2000] introduced a new affine connection in terms of which the dynamical terms of the gravitational field can be explicitly separated from those associated with the scalar field. When doing this, the energy-momentum tensor of the scalar field is simplified and is given by equation (compare with equation (4.7))

$$T_{\alpha\beta} = \frac{\theta(\phi) + 3/2}{\phi}\left(\partial_\alpha\phi\partial_\beta - \frac{1}{2}\eta_{\alpha\beta}\partial^\mu\phi\partial_\mu\phi\right)\,, \tag{2.395}$$

38) One reminds that the operator $\triangle$ is the Laplace differential operator introduced in equation (1.39).
39) One reminds that the Minkowski metric has signature $\eta_{\alpha\beta} = \text{diag}(-1, 1, 1, 1)$.

This energy-momentum tensor is well-defined as long as the energy density is positive. The above-mentioned argument about the energy-momentum tensor of a scalar field in the presence of gravity is closely associated with the discussion of the meaning of the Jordan-Fierz and Einstein frames in formulation of the field equations for gravitational field. Tensor (2.394) is given in the Jordan-Fierz frame while tensor (2.395) belongs to the Einstein frame.

3
General Relativity

3.1 The Principle of Equivalence

General theory of relativity stemmed from Einstein's attempt to incorporate the Newtonian theory of gravity to the framework of special theory of relativity. The guiding idea that led Einstein, was his remarkable observation of the identity between the force of gravity exerted on a particle in a homogeneous gravitational field and the force of inertia existing in each uniformly accelerated reference frame. This identity is now known as the *principle of equivalence* that can be also formulated in terms of the equivalence between the *inertial* and *gravitational* masses of a particle.

3.1.1 The inertial and gravitational masses

One reminds that in a classical mechanics a point-like particle of mass m obeys Newton's second law of motion

$$m\boldsymbol{a} = \boldsymbol{F} \,, \tag{3.1}$$

where $\boldsymbol{a} = d^2\boldsymbol{x}/dt^2$ is particle's acceleration, and $\boldsymbol{F}$ is the external force exerted on the particle. In the case of free fall in a gravity field, the force is defined by Newton's law of universal gravity

$$\boldsymbol{F} \equiv m\boldsymbol{g} \,, \tag{3.2}$$

where $\boldsymbol{g}$ is the gravitational force per unit mass that is a gradient of the Newtonian gravitational potential U taken at the position of the particle [1)]

$$\boldsymbol{g} = \boldsymbol{\nabla} U \,. \tag{3.3}$$

The coefficient m in equation (3.1) describes the inertial properties of the particle, that is how the particle responds to the force applied. On the other hand, the coefficient m in equation (3.2) determines the overall magnitude of the gravitational force exerted on the particle by the external bodies.

The two laws, (3.1) and (3.2), are fundamentally different in the Newtonian mechanics. Therefore, rigorously speaking, one should have formulated them more carefully as

$$m_{\rm i}\boldsymbol{a} = \boldsymbol{F} \,, \tag{3.4}$$

and

$$\boldsymbol{F} \equiv m_{\rm g}\boldsymbol{g} \,, \tag{3.5}$$

where $m_{\rm i}$ and $m_{\rm g}$ denote now the *inertial* and *gravitational* masses respectively. Hence, instead of one mass m one has two masses for one and the same particle.

1) See sections 1.2.1 and 1.2.5.

Combining two equations results in the acceleration acquired by a particle in gravity field

$$\boldsymbol{a} = \frac{m_g}{m_i}\boldsymbol{g} \,. \tag{3.6}$$

One is now faced with the question if the ratio m_g/m_i is exactly equal to unity, or, equivalently, whether the *inertial mass*, m_i, and the *gravitational mass*, m_g, have the same numerical value. Had one concerned about only one particle this question would have a positive answer. Indeed, by conveniently renormalizing the gravity constant G (and, accordingly, the value of $|\boldsymbol{g}|$) one would be always able to set the ratio m_g/m_i equal to unity for this particle. The answer would remain positive if the world consisted of many particles but all the particles in the universe were identical to one another - for all of them the ratio m_i/m_g would be the same. However, the world does not consist of identical particles. Hence, if one considers, in the same gravity field $\boldsymbol{g}$, two particles made of different substances [2), one can not keep the ratio $m_g/m_i = 1$ for both particles by renormalizing the universal gravity constant G. The other aspect of this question is associated with special relativity. According to this theory, mass is proportional to energy. Therefore, one should expect that two particles made of the same material but having different values of their internal energy (say, different degrees of magnetization or different temperatures) are to have different ratio of m_g/m_i. The equality of the inertial and gravitational mass can also depend, in principle, on location of the particle in space and time.

Current opinion, supported by numerous experiments [Gundlach *et al.*, 2009], is that the inertial and gravitational masses of all particles in the universe are equal. The reader, however, should understand that this point of view is an extrapolation to infinite accuracy of observations having a finite precision - future experiments may reveal the violation of the equality $m_g/m_i = 1$ [Braginsky, 1994; Damour, 2009a].

3.1.2
The weak equivalence principle

The assertion that the ratio m_g/m_i is the same for all point-like particles is called by Dicke [1965] the *weak equivalence principle* (WEP), also known as *universality of free fall*. Free fall implies that the particle is subject to the homogeneous gravity force and the tidal force is negligibly small. Formulation of WEP also neglects the particle's gravitational potential and excludes self-gravity effects. Another important aspect of WEP is that the experimental setup (including observer) must not influence the test particle's motion [3). Taking into account all the above considerations, WEP is formulated as follows:

In a homogeneous gravitational field, the acceleration of a freely-falling, structureless, test particle is independent of the particle's properties - its mass, com-

2) That is, the particles have different chemical compositions.
3) Observer has its own gravitational field which can affect the state of free motion (or rest) of test particle. This phenomenon must be carefully taken into account in the most precise experiments with test masses [Thorne and Winstein, 1999]

position, or thermodynamical state.

As a result, the worldline of a freely-falling test particle in a given gravitational field depends only on the particle's initial position and velocity. Thus, all test particles in this gravitational field will undergo the same acceleration, independent of their properties.

Provided all particles fall in the same manner in an external gravity field, a freely-falling observer should find the other freely-falling particles, in observer's immediate proximity, moving at uniform velocities relative to him. Thus, from the viewpoint of a freely-falling observer, the mechanics of particles in free fall is indistinguishable from their mechanics in the absence of gravity. To see this, consider a test particle moving along worldline $\boldsymbol{r} = \boldsymbol{r}(t)$ with nonrelativistic velocity in a constant homogenous gravitational field $\boldsymbol{g} = \boldsymbol{g}_0$. Assuming the WEP is valid, the Newtonian equation of motion (3.6) of the particle reads

$$\ddot{\boldsymbol{r}} = \boldsymbol{g}_0 \,, \tag{3.7}$$

where overdot denotes a time derivative. Switching to freely-falling coordinates, $\boldsymbol{w}$, connected with $\boldsymbol{r}$ via

$$\boldsymbol{r}(t) = \boldsymbol{w}(t) + \frac{1}{2}\boldsymbol{g}_0 t^2 \,, \tag{3.8}$$

transform equation (3.7) into

$$\ddot{\boldsymbol{w}} = 0 \,, \tag{3.9}$$

which means that in the absence of inhomogeneity of the gravitational field caused by tidal forces, no gravity field can be detected in the freely-falling frame of reference. Notice that if the WEP were not hold, the ratio $m_{\rm g}/m_{\rm i} \neq 1$, and equation (3.9) would not be satisfied for *all* freely-falling particles. This is the theoretical basis for all experiments that are trying to find out the violation of WEP [Will, 1993].

This development enables us to cast the WEP into another form:

In a homogenous gravitational field, the laws of mechanics in a freely-falling reference frame are the same as in the inertial reference frame in the absence of gravity.

For thirty six years, the experiment by Braginsky and Panov [1972] with collaborators had remained the most accurate test of the WEP. According to their measurement, the relative difference between the inertial and gravitational masses did not exceed 10^{-12}. This result was superseded only recently by Schlamminger *et al.* [2008] who managed to decrease this difference down to 3×10^{-13}.

3.1.3
The Einstein equivalence principle

WEP is formulated basically for mechanical motions of test particles having rest mass. Einstein proposed that it would be natural to extend this principle from mechanics to electrodynamics and to any other type of non-gravitational fundamental interactions. He justified the proposal by making use of a new type of experiment in physics called him *Gedankenexperiment* [4)] that suggested that in a close vicinity of a freely-falling observer *all* non-gravitational laws of physics are indistinguishable from the same laws formulated in the inertial reference frames. In modern terms, the *Einstein equivalence principle* (EEP) reads:

> *In a given gravitational field, the outcome of any local, non-gravitational experiment is independent of the freely-falling experimental apparatus' velocity, of where and when in the gravitational field the experiment is performed and of experimental technique applied.*

The word *local* in the formulation of the EEP means that, with a necessary degree of precision, the external gravity field can be assumed static and homogeneous. From observing the worldline of a single particle dropped in an elevator, one cannot infer whether the elevator is accelerating or is subject to gravity. However, measuring the relative motion of two particles placed initially at a sufficiently large spatial separation, one can discriminate between gravity and inertia by checking whether the trajectories of the particles are either parallel or are converging towards (diverging from) each other. Likewise, the EEP assumes that none of the observed particles is heavy enough to alter the background gravity field with its own gravitational potential.

In the EEP, the independence from velocity, position, and time can be referred to as local Poincaré invariance. Beside being an extension of Einstein's relativity principle employed in special relativity, the local Poincaré invariance also demands constancy of the fundamental physical parameters showing up in the laws of physics. Such parameters include, for example, the invariant speed c, the fine-structure constant, electron-to-proton mass ratio, etc. However, they do not include the universal gravitational constant G, because the EEP applies to non-gravitational experiments only. For a comprehensive discussion of tests of the EEP, see the article by Haugen and Lämmerzahl [2001] and references therein.

3.1.4
The strong equivalence principle

Even more general than the EEP is the strong equivalence principle (SEP), which extends the idea of equivalence of the inertial and gravitational masses of test particles further on to self-gravitating objects like planets, stars, etc., that is to those which intrinsic gravitational field is strong enough to influence the background gravitational field of external masses [Dicke, 1965]. This principle requires, among other things, that the gravitational constant G be the same all the time and everywhere in the universe. The SEP is formulated as follows:

> *The results of any local experiment,* gravitational or not*, in a freely-falling frame of reference are independent of where and when in the universe it is conducted.*

4) *Gedankenexperiment* is a German word meaning a "thought experiment".

The word *local* now pertains to the effects of *external* gravity field of other bodies alone, implying that, with a sufficient precision, this field remains static and homogeneous in the spacetime domain where the experiment is conducted. On the other hand, the effects of *internal* gravitational field of the "experimental apparatus" (which can be a particle, a planet, or even a star) are not ignored and must be incorporated to the definitions of the "apparatus" inertial and gravitational masses. There is a certain mathematical difficulty in separation of the internal and external gravitational fields for a self-gravitating body. The fact of the matter is that in general theory of relativity gravitational field does not obey the linear superposition principle. For this reason, it is rather unlikely, as shown for the first time by Fichtengolz [1950], to formulate the concept of the background gravitational field of the external masses that is fully independent of the internal gravitational field of the body under consideration. As of today, theoretical research in this direction has a real progress but still is a way too far from a final resolution of the problem.

The SEP can be also formulated as equivalence between the inertial and gravitational masses of the self-gravitating body participating in the local gravitational experiment. The equivalence assumes that the gravitational field contributes to the both masses of the body on the same footing. Again, the problem is that the intrinsic gravitational field of the body under consideration interacts non-linearly with the background gravitational field of the external masses. The concept of mass becomes quite complicated and to some extent ambiguous. Moreover, the above formulation of the SEP refers to locally-inertial frames. In special relativity, one postulates the existence of inertial frames due to homogeneity and isotropy of the Minkowski spacetime. In the presence of a self-gravitating body, however, one cannot demand that the spacetime remains homogenous and isotropic as gravitational field of the body affects the metric properties of the spacetime according to general theory of relativity. One can perhaps imagine that the body is originally imbedded to the Minkowski spacetime, whereafter its gravity field is “switched on" adiabatically by slowly increasing the body’s gravitational potential from zero to its "nominal" value without changing the background Minkowski spacetime. This thought experiment, though, looks quite uncertain. Particularly, it remains unclear to what degree the body in question can be kept isolated from other massive bodies in the universe which may affect the body under consideration in a number of various ways. For this reason, one would say that in the presence of gravity one has no a straightforward physical procedure for constructing a local inertial frame associated with a self-gravitating body though, in principle, such a procedure can be constructed, at least in the post-Newtonian approximation of general relativity (and scalar-tensor theory of gravity) as demonstrated in Chapter 4 below.

Taken these comments into consideration, a better form of the SEP would be:

All the laws of nature formulated in an external homogeneous and static gravitational field are the same as in a uniformly accelerated reference frame. These laws are independent of where and when in the gravitational field the experiment is performed, and are independent of the falling experimental apparatus’ velocity.

The SEP is not a necessary building block of Einstein's general relativity, in that the Einstein gravity field equations can be derived without referring to this principle [5]. The question is rather opposite, whether or not, the SEP follows from Einstein's field equations. Theoretically, fulfilment of the SEP for extended and self-gravitating bodies within general relativity has been confirmed, at least, up to the second post-Newtonian approximation [6] by Breuer and Rudolph [1982] and Kopeikin [1985]. Experimentally, its validity was confirmed, with a good precision, through timing of binary pulsars [Taylor, 1994] and, to lesser extent, by laser ranging to the Moon [Müller *et al.*, 2008b]. So far, it remains unknown if the SEP is fulfilled in general relativity *exactly* that is in *all* post-Newtonian approximations.

In section 6.3, one will touch upon the SEP also in the context of one of the alternative theories of gravity, the so-called scalar-tensor theory where the SEP gets violated.

3.1.5 The Mach principle

Despite the deceptively simple formulation of the equivalence principle, the road to this principle was not easy. When embarking on the quest for a relativistic theory of gravitation, Einstein was largely inspired by the philosophical views of Ernst Mach. Many a time Einstein mentioned Mach's theory of inertia as one of the inspirations for general relativity. Mach was cited also by later authors - Weinberg, Rindler, Bondi, and others, - who came to different, sometimes opposite conclusions on whether the general relativity theory goes along with the ideas of Mach.

The issue will forever remain subject to various exegeses, because Mach left to us not a theory of inertia in a rigorous mathematical sense, but a set of qualitative considerations, with no detailed calculations involved. Here comes a representative excerpt from Mach [1883]:

> "[The] investigator must feel the need of ... knowledge of the immediate connections, say, of the masses of the universe. There will hover before him as an ideal insight into the principles of the whole matter, from which accelerated and inertial motions will result in the same way."

Pais [2005] complains about Mach's style:

> "Reading his discourse is not unlike reading the Holy Scriptures. The text is lucid but one senses, perhaps correctly, perhaps wrongly, a deeper meaning behind the words".

Therefore, what really matters is how Einstein and other scholars of gravity interpreted Mach's philosophy.

5) For that, the EEP would be sufficient.
6) Post-Newtonian approximations represent an expansion of the solution of the Einstein equations in powers of a small parameter $\epsilon = v/c$, where v is a characteristic velocity of matter, and c is the invariant speed (see section 4.2 for more detail).

One theme by Mach is straightforward: he reiterated Descartes' idea that only relative motion is observable, but refused to follow Descartes in interpreting rest and motion as two different states of a body. In this, Mach took the side of Leibniz, rejecting the possibility of absolute space.

The second theme is the special way in which Mach rejected the absolute space. According to Newton, the existence of the inertial force in accelerated reference frame confirms that acceleration takes place against the background of absolute space or æther. According to Mach, the background of absolute space is fictitious and should be substituted with the direct interaction of a body under consideration with *all* matter in the universe or of the *"fixed stars"*, as Bishop Berkeley put it in his "De Motu" essay of 1721, where he also dwelled upon the problem of inertia. This way, Mach asserted that inertia is a phenomenon that relates the motion of a body to the average motion of all other bodies. In Mach's own words,

> "When ... we say that a body preserves unchanged its direction and velocity in space, our assertion is nothing more or less than an abbreviated reference to the entire universe" [Mach, 1883].

In modern terms, the property of a reference frame to be inertial depends upon whether or not it is in uniform motion relative to the distribution of distant quasars realized as a fundamental reference system (see section 9.4). Although the above quotation from Mach carries certain parallels with Berkeley, there also exists a difference in their views on inertia. Addressing the famous Newton's experiment with a rotating pail of water, Berkeley introduced the fixed stars only as reference bodies, to emphasize that motion must be relative. When explaining the reason for the rotating water surface to become concave, Berkeley accounted for the Newtonian concept of the centrifugal force by considering the interaction of water with the pail's wall but not with the stars [Newburgh, 2007; Suchting, 1967]. Mach, however, believed that the source of curvature of the water's surface lies in the influence from the fixed stars. Thus, the key point of Mach's approach was that it is the fixed stars, which distantly interact with a body, and they are the ultimate source of its inertia. This conjecture of the dynamic origin of inertia is what Einstein [1918] later termed *Mach's principle*.

Developing the relativity theory, Einstein borrowed from Mach's ideas only a necessary minimum - that inertia originates as a kind of interaction between bodies. In this interpretation, Mach's principle perfectly fits into Einstein's edifice. In particular, Einstein believed that the relativistic precession of the orbit of a test particle caused by rotation of a central body, currently known as the Lense-Thirring effect [Ciufolini and Wheeler, 1995], was a manifestation of the so-interpreted Mach's principle. Einstein's point of view was later questioned by Rindler [1994], who came to the conclusion that Mach's principle is antagonistic to the effect of the Lense-Thirring precession. On the other hand, a careful investigation by Bondi and Samuel [1997] demonstrated that Mach's principle is multivalued, and Rindler [1994] had attributed to Mach a strong particular assertion, the one that overall rigid rotations and translations of a system are unobservable. Bondi and Samuel [1997] confirmed that a milder version of the principle - that local inertial frames are affected by the cosmic motion and distribution of matter - is in a perfect agreement with Einstein's understanding of the Lense-Thirring effect.

Prior to Rindler [1994], Einstein's viewpoint on Mach's principle was challenged also by Weinberg [1972] who interpreted the negative results of an experiment that had been proposed by Cocconi and Salpeter [1958, 1960] and carried out by Hughes *et al.* [1960] and Drever [1961] as a failure of Mach's principle to conform to the equivalence principle. Since the inertial mass of a body may, according to Mach's principle, be affected by the global distribution of matter, Cocconi and Salpeter [1958, 1960] enquired if slight asymmetries in the distribution of matter at large would cause slight deviations from, at least, some of the known laws of mechanics and gravitation which are commonly assumed to be exact. They proposed an experiment based on the Zeeman splitting of energy levels in a nucleus of excited atom, that would check whether concentration of matter near the center of the Milky Way could generate asymmetries in inertia. These measurements were performed by a number of experimentalists starting from Hughes *et al.* [1960] and Drever [1961], that resulted in a negative answer [Chupp *et al.*, 1989; Lamoreaux *et al.*, 1986; Prestage *et al.*, 1985] to find any effect causing the critics of Mach's principle by Weinberg.

To acquit Mach's principle, let us recall Cocconi and Salpeter's line of reasoning. As a point of departure, they rightly mentioned that Mach's principle alone does not specify the nature of the effect that matter at large has on the inertia of a body. Considering the contribution of an external gravitational mass M to the inertia of a test particle placed from M at a vector distance $\boldsymbol{r}$ away, they proposed that this contribution is proportional to both M and to a negative power of $|\boldsymbol{r}|$. As a next step, they hypothesized that the contribution also depends on the angle θ between the directions of $\boldsymbol{r}$ and the acceleration of the test particle. Cocconi and Salpeter [1958] remarked that, if the contribution to inertia is θ-independent, then inertia would remain isotropic for any distribution of matter though the numerical value of the inertial mass of a test particle still could depend upon its position. This makes it clear that the Zeeman-splitting-based experiments were, in fact, tests not of Mach's principle *per se*, but of Mach's principle *amended* with the hypothesis of θ-dependence of the inertial mass. This hypothesis is neither a part of Einstein's relativity foundation nor a part of the original Mach's principle, but is an independent assertion. Thus, refutation of this supplementary hypothesis does not disprove Mach's principle.

Nevertheless, the first hypothesis by Cocconi and Salpeter [1958, 1960] - that the contribution to the inertia of a test particle is a function of the distance $|\boldsymbol{r}|$ of the particle from the external mass M - is still remain untouched. If this hypothesis is repudiated by experiment, then perhaps the entire Mach's principle will have to be questioned. The topic evidently needs further experimental research. One possible theoretical framework to handle this problem was offered by Sciama [1953, 1964a,b], who developed a theory based on the assumption that the inertial mass m of a particle is related to the gravitating mass M of the visible (causally-connected) patch of the universe through $mc^2 = G\int dM/r$ where the integral to be taken over the whole visible universe.

Another approach inspired by Mach's principle was suggested by Brans and Dicke [1961] and Dicke [1962b] who provided an early version of a scalar-tensor theory, a treatment in which the gravitational interaction is mediated by a scalar field as well as the metric tensor of general relativity. For a condensed explanation of the Brans-Dicke theory see, for example, Weinberg [1972, Chapter 7] and section 4.2 of the present book. For a more detailed review of other aspects of Mach's principle and its versions, the reader is referred to the paper by Bondi and Samuel [1997] and to the collection of articles in the book by Barbour and Pfister [1995].

3.2
The Principle of Covariance

Any physical theory must be *covariant* in the sense that its content must be independent of a coordinate choice [7]. The covariance demands the form of physical laws be unchanged under arbitrary differentiable coordinate transformations. In its own turn it implies that the mathematical objects entering the physical laws ought to transform, under coordinate transformations, as tensors or spinors of the same rank [8], and so should do the differential operators acting on the tensors that is their derivatives. The latter necessitates modification of the derivative operator, to make sure that irrespectively of the choice of the coordinates it always maps a tensor to another tensor. This section discusses the mathematical procedures associated with the covariance principle.

3.2.1
Lorentz covariance in special relativity

The spacetime in special relativity is a four-dimensional manifold $\mathcal{M}$ endowed with the Minkowski metric. In inertial coordinates, x^α, the Minkowski metric is a diagonal matrix $\eta_{\alpha\beta} = \mathrm{diag}(-1, 1, 1, 1)$. Equations of special relativity obey the Lorentz group of symmetry which leaves the special relativistic laws form-invariant under Lorentz transformation between the inertial coordinates. This form-invariance constitutes the Lorentz covariance principle in special relativity. In particular, the Minkowski metric does not change its form, when transformed from one inertial coordinates x^α to another $x^{\alpha'} = \Lambda^{\alpha'}{}_\beta x^\beta$, where $\Lambda^{\alpha'}{}_\beta$ is the matrix of the Lorentz transformation having been explained in section 2.3.5. In any inertial coordinates the Minkowski metric

$$\eta_{\alpha\beta} = \Lambda^{\mu'}{}_\alpha \Lambda^{\nu'}{}_\beta \eta_{\mu'\nu'} \ , \tag{3.10}$$

where the repeated indices denote summation from 0 to 3.

Because the Lorentz transformation is linear, the operator of a partial derivative transforms as a covector, that is

$$\frac{\partial}{\partial x^\alpha} = \Lambda^{\mu'}{}_\alpha \frac{\partial}{\partial x^{\mu'}} \ . \tag{3.11}$$

The same rule remains valid for higher-order partial derivatives. It automatically makes all physical laws of special relativity such as Maxwell equations, second Newton's law, etc., Lorentz invariant (see sections 2.1.5 and 2.7).

Lorentz covariance is the most fundamental principle of modern physics. Nevertheless, the word-combination "Lorentz violation" is mentioned quite often nowadays [Kostelecky, 2008; Kostelecký and Mewes, 2007]. What exactly does it mean?

7) The term covariant should not be confused with the related concept of a covariant vector from section 2.4.

8) In what follows, discussion of spinors is omitted.

The phrase “Lorentz violation" refers actually to *incomplete* theories which are described by the Lagrangians containing some terms which can be interpreted under certain circumstances as violating the Lorentz invariance. The Lorentz covariance is tightly connected with the fundamental nature of the invariant speed c [9)] which ensures that the spacetime in special relativity has neither preferred directions, nor absolute reference frame, nor any other additional structure besides the Minkowski metric. However, one should keep in mind that the Lorentz symmetry of special relativity is only a low-energy limit of the laws of physics, which may be not kept and involve new phenomena at some high-energy level like the fundamental Planck scale [10)] where the effects of yet unknown quantum fields and gravity become extremely strong. At that scale one may expect deviations from the known fundamental laws (electrodynamics, chromodynamics, etc.) that may be observed through the tiny violation of the CPT symmetry [Greenberg, O. W., 2002] tightly connected with the Lorentz invariance of the Minkowski spacetime. Specifically those presumable deviations are called as a spontaneous violation of the Lorentz covariance. It is clear, however, that such “violation" is nothing else but a matter of ignorance of the true mathematical structure of the Lagrangian of the theory which should preserve the Lorentz invariance at a new, more fundamental level.

Lorentz symmetry violation of a fundamental law is to be governed by an energy-dependent parameter of the Lagrangian which tends to zero at the low-energy limit where the violation vanishes. The parameter is linked to a privileged direction, $\boldsymbol{V}$, in spacetime that indicates to the existence of a preferred frame in vacuum. If, at high energy scale, one attempts to ignore the additional vector field, $\boldsymbol{V}$, the fundamental law under consideration will not preserve its form under the Lorentz transformation. On this occasion, Wald [1984] comments that the seemingly non-tensorial nature of the law stems from a failure to explicitly incorporate the extra geometrical object, $\boldsymbol{V}$, into the equation. When the preferred direction $\boldsymbol{V}$ is incorporated into the fundamental law explicitly, the law recovers its symmetry with respect to the Lorentz transformation at the cost of parametrization of the fundamental law by the "vacuum-rest-frame" field $\boldsymbol{V}$. Detection of the presence of such preferred frame effects, caused by the existence of the hypothetical æther-like field $\boldsymbol{V}$, at ultra-high energies can be tested in cosmic ray experiments [Pierre Auger Collaboration, 2007] where the energy threshold of the most energetic cosmic rays ($\simeq 10^{20}$ eV) is much higher than that of the Large Hadron Collider ($\simeq 1.4 \times 10^{13}$ eV).

3.2.2
Lorentz covariance in arbitrary coordinates

According to the special relativity principle, the two inertial frames are equivalent, in that any fundamental law of physics formulated in the two frames has an identical

9) One reminds that the invariant speed c is numerically equal to the speed of any massless fundamental field. Electromagnetic and gravitational interactions propagate with the invariant speed c.

10) The Planck scale corresponds to energy 1.22×10^{28} electron-volt (eV) that is equivalent to mass 2.18×10^{-8} kg.

form. However, the inertial frames are limited to spacetime coordinate systems related to each other by uniform relative motions only [11]. Einstein recognized that the special principle of relativity should be also applied to accelerated relative motions, and he used the newly developed tool of tensor calculus to extend the global Lorentz covariance (applying only to inertial frames) to the more general local Lorentz covariance which applies to all frames. This eventually led Einstein to producing general theory of relativity.

The essential idea is that coordinates do not exist a priori in nature, but are only a supplementary tool used in describing nature. Hence, they should not play any role in the formulation of fundamental physical laws. The first step, in switching from the Lorentz covariance in global inertial coordinates to arbitrary ones, is to accept that transformation from the inertial coordinates, $x^{\alpha'}$, to arbitrary coordinates, $x^\beta = x^\beta(x^{\alpha'})$, brings the Minkowski metric, $\eta_{\mu'\nu'} = \mathrm{diag}(-1, 1, 1, 1)$, to about a more general form [12]

$$g_{\alpha\beta}(x) = \frac{\partial x^{\mu'}}{\partial x^\alpha}\frac{\partial x^{\nu'}}{\partial x^\beta}\eta_{\mu'\nu'} , \tag{3.12}$$

where, as usual, the repeated indices assume the summation, and the metric $g_{\alpha\beta} = g_{\alpha\beta}(x)$ is a differentiable tensor field of rank 2, with each component being a function of the new coordinates x^α. Of course, transformation (3.12) does not change the signature of the metric. Any object entering a fundamental law must transform from one coordinate chart to another in accordance with its tensorial nature like a scalar, vector, or covariant/contravariant tensor of a higher rank (see section 2.4).

However, since the fundamental laws of physics are given in the form of the differential equations, a problem arises with the transformation laws of the derivatives of the tensor fields. Indeed, let us consider an arbitrary vector field $\boldsymbol{V}$ defined on spacetime manifold $\mathcal{M}$ that is locally parameterized with coordinates x^α. Any other chart $x^{\beta'} = x^{\beta'}(x^\alpha)$ parameterizing the same patch of the spacetime is equally usable, insofar as the coordinate transformation is non-singular and differentiable. While the components V^α of $\boldsymbol{V}$ transform as a contravariant tensor of the first rank

$$V^{\alpha'}(x') = \frac{\partial x^{\alpha'}}{\partial x^\beta}V^\beta(x) , \tag{3.13}$$

their partial derivatives, generally, do *not* behave as components of a tensor,

$$\begin{aligned}\frac{\partial V^{\alpha'}}{\partial x^{\gamma'}} &= \frac{\partial x^{\alpha'}}{\partial x^\beta}\frac{\partial V^\beta}{\partial x^{\gamma'}} + \frac{\partial^2 x^{\alpha'}}{\partial x^{\gamma'}\partial x^\beta}V^\beta \\ &= \frac{\partial x^{\alpha'}}{\partial x^\beta}\frac{\partial x^\lambda}{\partial x^{\gamma'}}\frac{\partial V^\beta}{\partial x^\lambda} + \frac{\partial^2 x^{\alpha'}}{\partial x^\lambda\partial x^\beta}\frac{\partial x^\lambda}{\partial x^{\gamma'}}V^\beta .\end{aligned} \tag{3.14}$$

The tensor law is “spoiled" by the presence of a term with a second derivative in equation (3.14). Only when both coordinate charts x^α and $x^{\alpha'}$ are rectilinear, the “bad" term vanishes and the transformation becomes tensorial.

11) As one has mentioned in section 2.3.5, transformation between any two inertial frames can also involve in the most general case a constant spatial rotation and translation [Richtmyer, 1982].

12) The coordinate transformation, $x = x(x')$, is differentiable, and admits an inverse transformation, $x' = x'(x)$, at least, locally.

This non-tensorial nature of partial derivatives can be surmounted by changing the partial derivative operator to the new, covariant derivative operator that transforms as a tensor. In inertial frames with rectilinear coordinates, the so-modified covariant derivative coincides with the regular partial derivative. Because the entire Minkowski spacetime can be always covered by a single chart, representing an inertial frame, this modification of the derivative is intended simply to make up for the "twisting" of curvilinear coordinates that fit better for description of physics in non-inertial (accelerating, rotating, etc.) frames as explained in section 2.6. However, the mathematical value of the covariant derivative goes beyond that as will become clear later in discussing the general relativity theory.

3.2.2.1 *Covariant derivative and the Christoffel symbols in special relativity*

Formula (3.14) is equivalent to the following expression for the vector differential

$$dV^{\alpha'} = \frac{\partial x^{\alpha'}}{\partial x^\beta} dV^\beta + \frac{\partial^2 x^{\alpha'}}{\partial x^\lambda \partial x^\beta} V^\beta dx^\lambda \; . \tag{3.15}$$

This differential transforms in a non-tensorial (non-vectorial) manner, except when both coordinate grids are rectilinear as the term with a second derivative vanishes in such case. It is possible to replace the ordinary differential with another linear operator of differentiation which transforms as a tensor. For this purpose, let us introduce a new object, $\Gamma^\alpha_{\mu\nu}$, which is symmetric with respect to the sub-indices, $\Gamma^\alpha_{\mu\nu} = \Gamma^\alpha_{\nu\mu}$, and transforms according to a non-tensorial rule [13)]

$$\Gamma^\alpha_{\mu\nu} = \frac{\partial x^\alpha}{\partial x^{\beta'}} \frac{\partial x^{\lambda'}}{\partial x^\mu} \frac{\partial x^{\rho'}}{\partial x^\nu} \Gamma^{\beta'}_{\lambda'\rho'} + \frac{\partial x^\alpha}{\partial x^{\beta'}} \frac{\partial^2 x^{\beta'}}{\partial x^\mu \partial x^\nu} \; . \tag{3.16}$$

It is now easy to prove that a linear combination

$$DV^{\alpha'} \equiv dV^{\alpha'} + \Gamma^{\alpha'}_{\mu'\nu'} V^{\mu'} dx^{\nu'} \; , \tag{3.17}$$

behaves as a vector. Indeed, taking into account equations (3.13), (3.15), and (3.16), one obtains

$$dV^{\alpha'} + \Gamma^{\alpha'}_{\mu'\nu'} V^{\mu'} dx^{\nu'} = \frac{\partial x^{\alpha'}}{\partial x^\beta} \left(dV^\beta + \Gamma^\beta_{\mu\nu} V^\mu dx^\nu \right) \; , \tag{3.18}$$

that is

$$DV^{\alpha'} = \frac{\partial x^{\alpha'}}{\partial x^\beta} DV^\beta \; . \tag{3.19}$$

The differential operator D introduced in equation (3.17) is called the *absolute* differential, and the object $\Gamma^\alpha_{\mu\nu}$ is known as the *Christoffel* symbol.

The *covariant derivative* of a vector field is defined by the following[14)] equation

$$\nabla_\beta V^\alpha = \frac{\partial V^\alpha}{\partial x^\beta} + \Gamma^\alpha_{\mu\beta} V^\mu \; . \tag{3.20}$$

13) The rule is intentionally chosen to be non-tensorial because it has to compensate for the non-tensorial part of the transformation of the regular partial derivative as expressed by equation 3.14.

14) One reminds that a regular partial derivative is denoted with $\partial_a V^\beta = \partial V^\beta / \partial x^\alpha$. Covariant derivative is denoted with $\nabla_\alpha V^\beta$.

It is straightforward to prove that the covariant derivative of a vector field is transformed as a tensor of the second rank

$$\nabla_{\alpha'} V^{\beta'} = \frac{\partial x^{\beta'}}{\partial x^{\nu}} \frac{\partial x^{\mu}}{\partial x^{\alpha'}} \left(\nabla_{\mu} V^{\nu}\right) . \tag{3.21}$$

By definition, the covariant derivative of a scalar field ϕ coincides with its regular partial derivative

$$\nabla_{\alpha}\phi = \partial_{\alpha}\phi = \phi_{,\alpha} = \frac{\partial \phi}{\partial x^{\alpha}} . \tag{3.22}$$

The covariant derivative must also satisfy the chain rule for the product of two tensors. In particular, taking a covariant derivative from the scalar $V^{\alpha} W_{\alpha}$, and making use of the chain rule, one can check by inspection that the covariant derivative of a covector must be defined as

$$\nabla_{\alpha} W_{\beta} = \frac{\partial W_{\beta}}{\partial x^{\alpha}} - \Gamma^{\mu}_{\alpha\beta} W_{\mu} , \tag{3.23}$$

where the term with the Christoffel symbols enters with the sign minus as opposed to the definition (3.20) of a covariant derivative from a vector. The covariant derivative from the covector is transformed as a covariant tensor of type $(0, 2)$,

$$\nabla_{\alpha'} W_{\beta'} = \frac{\partial x^{\mu}}{\partial x^{\alpha'}} \frac{\partial x^{\nu}}{\partial x^{\beta'}} \left(\nabla_{\mu} W_{\nu}\right) . \tag{3.24}$$

The machinery of covariant differentiation easily extends from equations (3.20) and (3.23) to tensor fields of all ranks. For example, for tensors of the second rank of various types, the covariant derivatives have the following form

$$\nabla_{\alpha} S^{\mu\nu} = \partial_{\alpha} S^{\mu\nu} + \Gamma^{\mu}_{\alpha\beta} S^{\beta\nu} + \Gamma^{\nu}_{\alpha\beta} S^{\mu\beta} , \tag{3.25}$$

$$\nabla_{\alpha} T^{\mu}_{\nu} = \partial_{\alpha} T^{\mu}_{\nu} + \Gamma^{\mu}_{\alpha\beta} T^{\beta}_{\nu} - \Gamma^{\beta}_{\alpha\nu} T^{\mu}_{\beta} , \tag{3.26}$$

$$\nabla_{\alpha} U_{\mu\nu} = \partial_{\alpha} U_{\mu\nu} - \Gamma^{\beta}_{\alpha\mu} U_{\beta\nu} - \Gamma^{\beta}_{\alpha\nu} U_{\mu\beta} . \tag{3.27}$$

The reader can easily verify that these equations yield tensors of the third rank.

3.2.2.2 *Relationship between the Christoffel symbols and the metric tensor*

Once the components of the Christoffel symbols have their values set in one chart, their values in any other coordinate chart are given by (3.16). Let us start from the inertial coordinates $x^{\alpha'}$ in which the components of the Christoffel symbol assume zero values, $\Gamma^{\beta'}_{\lambda'\rho'} = 0$, and transform it to the curvilinear [15] coordinates x^{α}. Then, equation (3.16) tells us that in the new coordinates

$$\Gamma^{\alpha}_{\mu\nu} = \frac{\partial x^{\alpha}}{\partial x^{\beta'}} \frac{\partial^2 x^{\beta'}}{\partial x^{\mu} \partial x^{\nu}} , \tag{3.28}$$

and it is not equal to zero. It is remarkable that the Christoffel symbols $\Gamma^{\alpha}_{\mu\nu}$ can be expressed through the components of the metric and their derivatives. To this end,

15) The new coordinates, x^{α}, are curvilinear, if they are not linear functions of the old coordinates $x^{\alpha'}$.

let us rewrite equation (3.28) as

$$\frac{\partial^2 x^{\beta'}}{\partial x^\mu \partial x^\nu} = \frac{\partial x^{\beta'}}{\partial x^\alpha}\Gamma^\alpha_{\mu\nu} , \tag{3.29}$$

and compare it to the result of differentiation of equation (3.12)

$$\frac{\partial g_{\alpha\beta}}{\partial x^\lambda} = \eta_{\mu'\nu'}\frac{\partial^2 x^{\mu'}}{\partial x^\alpha \partial x^\lambda}\frac{\partial x^{\nu'}}{\partial x^\beta} + \eta_{\mu'\nu'}\frac{\partial x^{\mu'}}{\partial x^\alpha}\frac{\partial^2 x^{\nu'}}{\partial x^\beta \partial x^\lambda} . \tag{3.30}$$

Substituting equation (3.29) to the right side of equation (3.30) allows us to re-write it in the following form

$$\frac{\partial g_{\alpha\beta}}{\partial x^\lambda} = g_{\gamma\alpha}\Gamma^\gamma_{\beta\lambda} + g_{\gamma\beta}\Gamma^\gamma_{\alpha\lambda} , \tag{3.31}$$

where one has used equation (3.12) for the metric tensor in curvilinear coordinates. Making use of two similar equalities - with indices α and λ transposed, and with β and λ transposed, one can resolve equation (3.31) with respect to the Christoffel symbols. It yields

$$g_{\lambda\alpha}\Gamma^\lambda_{\beta\gamma} = \frac{1}{2}\left(\frac{\partial g_{\alpha\beta}}{\partial x^\gamma} + \frac{\partial g_{\gamma\alpha}}{\partial x^\beta} - \frac{\partial g_{\gamma\beta}}{\partial x^\alpha}\right) \tag{3.32}$$

According to definition the contravariant metric tensor $g^{\alpha\beta}$ is the inverse with respect to its covariant components

$$g_{\alpha\nu}g^{\nu\beta} = \delta^\beta_\alpha . \tag{3.33}$$

Applying this equality in equation (3.31), one finally obtains

$$\Gamma^\alpha_{\beta\gamma} = \frac{1}{2}g^{\alpha\lambda}\left(\frac{\partial g_{\lambda\beta}}{\partial x^\gamma} + \frac{\partial g_{\lambda\gamma}}{\partial x^\beta} - \frac{\partial g_{\beta\gamma}}{\partial x^\lambda}\right) , \tag{3.34}$$

a milestone result relating the Christoffel symbols to the metric and its first derivatives. Importantly, equation (3.34) makes the Christoffel symbols vanish when $g_{\alpha\beta}$ is the (constant) Minkowski metric. This agrees with the starting point of the calculation - setting the Christoffel symbols zero in the inertial coordinates.

At this point, it is worthwhile to mention that the Christoffel symbols represent a particular example of a new geometric object, the *affine connection*, that will be introduced in section 3.4 on a more systematic basis. The affine connection is independent of the metric tensor in the most general case of an affine manifolds which are specific to some alternative theories of gravity. Spacetime manifold of general relativity is more restrictive and does not allow existence of the affine connection being independent of the metric tensor. This limitation on the affine connection is also employed in the parameterized post-Newtonian (PPN) formalism [Will, 1993] which is commonly used for experimental testing of general relativity.

3.2.2.3 *Covariant derivative of the metric tensor*

Taking covariant derivatives from the metric tensor $g_{\alpha\beta}$ with the help of equation (3.27) and using (3.31) one easily proves that the covariant derivative from the metric tensor is identically zero

$$\nabla_\gamma g_{\alpha\beta} \equiv 0 . \tag{3.35}$$

Differentiation of equation (3.33) with accounting for the chain rule and equation (3.35) also yields

$$\nabla_\gamma g^{\alpha\beta} \equiv 0 . \tag{3.36}$$

An important consequence of these two equations is that the metric tensor is constant with respect to covariant differentiation.

3.2.3
From Lorentz to general covariance

The introduction of the covariant derivative to special theory of relativity allows us to formulate it in more general way. The global Lorentz covariance is preserved but it is now applicable in non-inertial frames covered with arbitrary coordinates. Any fundamental law of special relativity becomes covariant in the sense that it transforms in accordance with the laws of tensor calculus in arbitrary coordinates introduced in the Minkowski spacetime.

This formulation significantly extends the domain of applicability of special theory of relativity. For example, it allows us to study physical processes in accelerated and rotating frames of reference, which are not inertial. Nevertheless, the extended Lorentz covariance is still limited to the case of the global Minkowski spacetime. The mathemtical objects that appear in the arbitrary coordinates - the metric tensor, $g_{\alpha\beta}$, and the Christoffel symbols, $\Gamma^\alpha_{\beta\gamma}$ - does not bear a new geometric significance because they are obtained directly from the Minkowski metric by the coordinate transformation. It is always possible to make the Christoffel symbols, $\Gamma^\alpha_{\beta\gamma}$, vanish, and the metric tensor $g_{\alpha\beta}$ to become a Minkowski metric $\eta_{\alpha\beta} = \mathrm{diag}(-1, 1, 1, 1)$, in the entire Minkowski spacetime by transforming the curvilinear coordinates back to those employed by the inertial observers.

It is conceivable however to make a next step ahead. Minkowski spacetime is not the most general type of manifolds as it can be covered by a single coordinate chart corresponding to a global inertial reference frame. Most of the manifolds studied in mathematics, do not admit such a luxury - two or more coordinate charts are, in general, required to cover them. It makes impossible to nullify the Christoffel symbols at each point of such manifolds by making a global coordinate transform - it can be done only in a limited domain of the manifold $\mathcal{M}$ covered by the local coordinates. The Christoffel symbol, introduced in section 3.2.2.1 as a tool to make partial derivatives covariant in arbitrary coordinates, becomes a non-trivial geometric object on the manifold $\mathcal{M}$, called *affine connection*. The affine connection is not entirely associated with the choice of coordinates. It determines the affine structure of the manifold $\mathcal{M}$ that is the law of parallel transport of tensors from one point of the manifold $\mathcal{M}$ to another. The law of transformation (3.16) of the affine connection from one coordinate chart to another remains the same on arbitrary manifold but its applicability becomes local and can not be extended to entire manifold.

Einstein realized that the formalism of differential geometry on manifolds endowed with the affine connection, is a key to formulation of the most general type of covariance in nature including gravity. According to Einstein, the principle of general covariance consists of two parts:

A. Fundamental laws of physics must not depend on the choice of coordinates.
B. Fundamental laws of physics must not depend on a prior geometry of spacetime.

Part A of the principle tells us that a fundamental law must be formulated in the language of tensor calculus that relies upon general formalism of the covariant derivative. Part B of the principle tells us that the fundamental law is local and its tensorial form is not affected by the choice of the structure of the spacetime manifold $\mathcal{M}$ which is defined by the distribution of matter and other physical fields.

The principle of general covariance still remains the center of a discussion that ignited almost immediately after the birth of Einstein's theory of general relativity. The question, that goes back to Kretschmann [1917], is whether general covariance is merely a mathematical requirement which is physically vacuous, or is a manifestation of a non-trivial physical circumstance distinguishing Einstein's theory of gravity from others. In a brief form, the counter-argument furnished by Kretschmann [1917] against the fundamental role of general covariance can be put like this:

> *Any* theory of space and time can be given a generally covariant formulation as long as we introduce a sufficient number of auxiliary fields transforming in a manner necessary to serve one's demands; so the principle of general covariance contains no specific information that could be used to build a new theory.

Kretschmann's objection would make sense if, and only if, one did not go beyond the framework of special relativity. Indeed, the part A of the principle of general covariance is a pure mathematical statement which can be applied to formulate *any* physical law, even if it has no fundamental value, in a covariant form [Kadomtsev *et al.*, 1972; Tyapkin, 1972]. On the other hand, the part B of the general covariance principle is a cornerstone in the building of the Einstein theory of gravity - the general relativity. In special relativity the covariance was achieved through a formal rearrangement of the already known equations under condition that prior geometry is fixed and remains Minkowski. General theory of relativity has no prior geometry, it must be determined from the gravitational field equations formulated by Einstein in 1915 [16].

16) In approximately the same time, D. Hilbert derived Einstein's equations from a variational principle.

3.2.4
Two approaches to gravitation in general relativity

Einstein's theory of general relativity identifies gravitation and geometry of spacetime manifold $\mathcal{M}$, which is not globally Minkowski. This leads to geometrization of the fundamental laws of nature, which become naturally coupled with the gravitational field through the affine connection (the Christoffel symbols) that defines the covariant derivative on the manifold $\mathcal{M}$. The manifold $\mathcal{M}$ has a curvature [17] that is not constant, and may be topologically nontrivial [18]. Thus, there is no chance to cover the spacetime manifold $\mathcal{M}$ with one non-singular chart. However, one expects the manifold $\mathcal{M}$ to be locally, in some vicinity of each point P, equivalent to the Minkowski spacetime in the sense that there is a local inertial frame at P in which the metric tensor reduces to the Minkowski metric, $g_{\alpha\beta}(P) = \eta_{\alpha\beta} = \mathrm{diag}(-1, 1, 1, 1)$, and the Christoffel symbols vanishes, $\Gamma^{\alpha}_{\beta\gamma}(P) = 0$. Physical fields and matter are described by various geometric objects - scalar, vector, tensor, etc. The local Minkowski spacetime is tangent to the manifold $\mathcal{M}$ at the point P, and algebraic operations on the geometric objects are performed in the tangent spacetime. Comparison of similar geometric objects being pinned down to two different tangent spacetimes is performed with the law of parallel transportation of the object from one tangent space to another. This involves the concept of the covariant derivative and the affine connection. General relativity *postulates* that the covariant derivative of the metric tensor is zero, which is a mathematical expression of the physical principle of equivalence that has been discussed in section 3.1. This postulate defines unambiguously the Christoffel symbols through equation (3.34). The metric tensor itself is determined by solving the Einstein gravity field equations discussed below in section 3.9. This brief discussion of the Einstein's theory of gravity is fully consistent with the general principle of covariance and does not rely upon any assumption about prior geometry.

Many researchers were not completely satisfied with the geometrization of physics as proposed by Einstein. Indeed, the geometrization makes it difficult to put general relativity on the same footing as quantum mechanics and/or other particle physics theories which require some prior background geometry. A number of researchers suggested an alternative approach to gravity in which the gravitational field is considered as a classic analogue of the quantum field theory of tensor particles having spin 2 and admitting a non-linear self-interaction. This tensor field is defined on the flat Minkowski background of the special relativity like a vector electromagnetic field in Maxwell's theory or a spinor fields in particle's theory were defined. Original approaches to find out the equations of the spin-2 field were limited by a linear approximation with respect to the amplitude of the field [Kraichnan, 1955; Ogiyevetsky and Barinov, 1965; Weinberg, 1965]. It is not surprising that they disagreed with general

17) Curvature is a new geometric object on the manifold. It corresponds to the tidal force in the Newtonian theory of gravity (see section 3.7 for further details).

18) For more accurate definition of manifold, formulated in the language of topologies see, for example, the book by de Felice and Clarke [1990].

relativity. It was Deser [1970] who had taken into account the non-linear nature of the tensor field and discovered that the final equations for this field exactly coincides with the Einstein equations. Grishchuk *et al.* [1984] and Popova and Petrov [1988] went further, demonstrating that *any* background spacetime can be employed as a starting arena for the tensor-field theory of gravity, and that a non-linear dynamical theory developed thereupon turns out to be mathematically equivalent to Einstein's general relativity. These works strongly support the idea that the background geometry, though rather useful for some mathematical applications [Grishchuk, 2009], is physically unobservable, and can not be an integral part of the theory of general relativity [19]. Thus, the field-theoretical study of general relativity confirmed the physical importance of the principle of general covariance in a new, fascinating way.

This book relies upon geometrical approach to general theory of relativity. It requires a rather detailed discussion of differential geometry on manifolds and geometric structures associated with it. These concepts will be introduced in a few subsequent chapters. One starts from the definition of differential manifold and tangent space.

19) Scalar-tensor theories of gravity have no prior geometry either [Will, 1993]. Scalar-tensor theory was worked out by Brans and Dicke [1961]; Fierz [1956]; Jordan [1949, 1959] and further extended to take into account the scalar field multiplet by Damour and Esposito-Farese [1992] (see section 4.2.1).

3.3 A Differentiable Manifold

One has already discussed the concept of manifold $\mathcal{M}$ in relation to the Minkowski spacetime in section 2.2.1. It is well-known that it can be entirely covered by a single coordinate chart forming an inertial reference frame in the whole spacetime. The spacetime of general relativity and any other viable theory of gravity is more complicated and is not reduced to the Minkowski spacetime globally. It means that the mathematical description of topological structure of the manifold $\mathcal{M}$ should be further developed to take into account its local properties in different locations.

3.3.1 Topology of manifold

In differential geometry and topology, a manifold is a continuous set of points that on a small enough scale resembles the Euclidean space R^n of a dimension n. The dimension of the manifold is, then, n and it is a local invariant. For example, spacetime in general relativity is a four-dimensional manifold $\mathcal{M}$, $n = 4$. Manifold looks locally like a Euclidean space in a purely topological sense which means that each manifold can be locally reduced to the Euclidean space by doing homeomorphic transformation of a sufficiently small region of the manifold that is, roughly speaking, its continuous stretching and bending. Definition of manifold does not include any particular and consistent choice of such concepts as distance, angles, parallelism of vectors, etc. In particular, topologically similar manifolds can have an infinite number of ways to specify distances and angles. In order to discuss these and other properties of a manifold, one needs to endow it with additional structures and to consider differentiable and Riemannian manifolds discussed below.

An *intrinsic* definition for differentiable manifolds in terms of local coordinate charts was introduced by Weyl in 1913 [Weyl, 1964]. On the other hand, Whitney [1936] and Nash [1956] have developed an *extrinsic* definition of manifold from the point of view of its embedding to a Euclidean space of higher dimensions. They showed that the intrinsic and extrinsic definitions are equivalent [20]. This basic result shows that manifolds may be treated intrinsically or extrinsically, as one wishes. One, however, prefers the intrinsic definition because it neither relies upon particular details of the procedure of embedding of the manifold to the Euclidean space of higher dimensions nor on assumptions about the number of these dimensions. Extrinsic point of view is valuable in some mathematical and physical applications like in string theory or quantum gravity but there are no yet experimental evidences that the four-dimensional spacetime of general relativity is really embedded to a space of higher dimensions.

20) The pioneering role of other scientists in developing the concept of a differential manifold is discussed by Shields [1988].

3.3.2
Local charts and atlas

The structure of a differentiable manifold is encoded by a collection of local coordinate charts that form an atlas of the manifold. The composition of one chart with the inverse of another chart is a differentiable function called a transition map that is equivalent to a coordinate transformation from an open subset of Euclidean space to the manifold and then back to another open subset of Euclidean space. In case of the spacetime manifold $\mathcal{M}$ the transition map is given by four (transition) functions $x^\alpha = x^\alpha(y^\beta)$ $(\alpha, \beta = 0, 1, 2, 3)$ where x^α and y^α are local coordinates of a point P on the manifold $\mathcal{M}$ corresponding to two charts.

One assumes the coordinates are smooth functions so a determinant of the matrix of transformation $\Lambda^\alpha{}_\beta = \partial x^\alpha / \partial y^b$ is not singular at the domain of intersection of the two local charts. We also assume the determinant of the transition map is positively defined in all intersections of the charts of the manifold's atlas that excludes non-orientable manifolds like the Möbius strip or the Klein bottle. We shall further assume that the transition functions are analytic that is they can be expanded in the infinite Taylor series around any point of the spacetime manifold $\mathcal{M}$.

3.3.3
Functions

Function f on a differentiable manifold $\mathcal{M}$ maps each point of an open set of the manifold to one-dimensional Euclidean space R^1. In local coordinates the function is represented as $f(x^\alpha)$ where x^α are the coordinates of a point on the manifold $\mathcal{M}$. If all partial derivatives of the function with respect to the local coordinates exist up to the order k, the function is called C^k-differentiable or simply C^k-function. We shall basically work with functions which are differentiable an infinite number of times, that is C^∞-functions. We shall also admit that these functions are analytic, that is they can be expanded in an infinite Taylor series in the vicinity of each point.

3.3.4
Tangent vectors

Vector space has been introduced axiomatically in section 2.3 in relation to the Minkowski spacetime. Spacetime of general relativity has more complicated structure making the concept of vector meaningful only locally at each point of the manifold $\mathcal{M}$. This localization of vector can be performed with the help of the operator of directional derivative introducing a tangent vector to the manifold $\mathcal{M}$.

To formalize this concept, let us consider a curve C_τ on a manifold $\mathcal{M}$ that maps points of an open interval of one-dimensional Euclidean space R^1 to the manifold $\mathcal{M}$. In local coordinates, x^α, the curve C_τ is given in the form of four functions, $x^\alpha(\tau)$, with τ being a continuous parameter along the curve. At this stage the metrical nature of the parameter τ is irrelevant and, in fact, can not be identified as the metric tensor has not yet been introduced. Fact of the matter is that the definition of a parameterized curve on manifold does not depend on whether it is endowed with the metric or not.

Tangent (contravariant) vector $\boldsymbol{V}$ at a point P of a manifold $\mathcal{M}$ is defined as an operator of differentiation $d/d\tau$ along the curve C_τ passing through the point P. It maps any function f at the point P of the manifold $\mathcal{M}$ to a number $df/d\tau$ from one-dimensional Euclidean space R^1. In the local coordinates the mapping of function f to R^1 becomes a function $g(\tau) = f[x^\alpha(\tau)]$. Its differentiation along the curve yields

$$\frac{dg}{d\tau} = \frac{\partial f}{\partial x^\alpha}\frac{dx^\alpha}{d\tau} , \tag{3.37}$$

As this equation must be valid for any function f, one can omit it from the equation which becomes

$$\frac{d}{d\tau} = \frac{\partial}{\partial x^\alpha}\frac{dx^\alpha}{d\tau} . \tag{3.38}$$

Four numbers $dx^\alpha/d\tau$ define components V^α of the vector $\boldsymbol{V}$ being tangent to the curve C_τ, that is $V^\alpha = dx^\alpha/d\tau$. The components depend on the particular parametrization of the curve because if another parameter $\lambda \neq \tau$ is chosen along it, the components of the vector will change their magnitude

$$W^\alpha = \frac{dx^\alpha}{d\lambda} = \frac{dx^\alpha}{d\tau}\frac{d\tau}{d\lambda} = V^\alpha \frac{d\tau}{d\lambda} , \tag{3.39}$$

where one has assumed that the parameter $\tau = \tau(\lambda)$ is a smooth function, and $d\tau/d\lambda \neq 1$ on the curve. We shall accept that two curves, C_τ and C_λ, are different even if they are passing through the same points on the manifold but with different values of their parameters. This is because they define two formally different mappings of the points of the manifold $\mathcal{M}$ to R^1. Thus, one postulates that each curve on the manifold defines only one tangent vector at each point. In a coordinate-free language the tangent vector $\boldsymbol{V}$ to a curve C_τ is identified with the operator $d/d\tau$

$$\boldsymbol{V} \equiv \frac{d}{d\tau} , \tag{3.40}$$

that acts on scalar differentiable function f and converts them to numbers,

$$\boldsymbol{V}[f] \equiv \frac{df}{d\tau} . \tag{3.41}$$

Notice that the square brackets in the notation $\boldsymbol{V}[f]$ mean that function f is considered as an argument of the operator of differentiation $\boldsymbol{V}$. In particular, the value of a scalar function f on a basis vector $\boldsymbol{e}_\alpha$ is a directional derivative of function f along this basis vector, $\boldsymbol{e}_\alpha[f]$. In a coordinate basis $\boldsymbol{e}_\alpha = \partial/\partial x^\alpha$, and hence

$$\boldsymbol{e}_\alpha[f] = \partial_\alpha f = \frac{\partial f}{\partial x^\alpha} . \tag{3.42}$$

This allows us to recast equation (3.41) to more transparent form

$$\boldsymbol{V}[f] = V^\alpha \partial_\alpha f . \tag{3.43}$$

Tangent vectors are not linear functions on scalars in the sense that for the product of two smooth functions, f and g, tangent vector $\boldsymbol{V}$ must obey the Leibnitz rule:

$$\boldsymbol{V}[gf] = g\boldsymbol{V}[f] + f\boldsymbol{V}[g] \,. \tag{3.44}$$

It turns out that while each curve defines a tangent vector uniquely, each vector at a given point of manifold is tangent to an infinite number of curves. Indeed, two different curves, C_τ and C_λ, passing through a point P of the manifold produce one and the same vector at this point, if the parameters, τ and λ are related in such a way that at the point P the derivative $(d\tau/d\lambda)_{\rm P} = 1$. This condition does not impose any other restriction on functional dependence of τ from λ. This short discussion brings to a conclusion that each tangent vector at a point P of a manifold defines an entire class of equivalence of curves passing through this point.

3.3.5
Tangent space

A set of tangent vectors defines a vector space at each point P of manifold $\mathcal{M}$. This vector space at each point of the manifold is called a *tangent space*. In order to prove this statement, let us take two different curves, C_τ and $\mathcal{L}_\lambda$, passing through the point P. At this point the two curves have two tangent vectors

$$\boldsymbol{V} = \frac{d}{d\tau} = \frac{dx^\alpha}{d\tau}\frac{\partial}{\partial x^\alpha} \,, \tag{3.45a}$$

$$\boldsymbol{U} = \frac{d}{d\lambda} = \frac{dx^\alpha}{d\lambda}\frac{\partial}{\partial x^\alpha} \,, \tag{3.45b}$$

where one has used local coordinates x^α. Now, one picks up two arbitrary *constant* numbers, a and b, and form a linear combination

$$a\boldsymbol{V} + b\boldsymbol{U} = \left(a\frac{dx^\alpha}{d\tau} + b\frac{dx^\alpha}{d\lambda}\right)\frac{\partial}{\partial x^\alpha} \,. \tag{3.46}$$

It can be interpreted as a new vector

$$\boldsymbol{W} = \frac{d}{d\sigma} = \frac{dx^\alpha}{d\sigma}\frac{\partial}{\partial x^\alpha} \,, \tag{3.47}$$

that is tangent to a third curve $\mathcal{S}_\sigma$ passing through the point P, and it has coordinates

$$\frac{dx^\alpha}{d\sigma} = a\frac{dx^\alpha}{d\tau} + b\frac{dx^\alpha}{d\lambda} \,. \tag{3.48}$$

This proves that operators of differentiation along all possible curves passing through point P of the manifold, form a vector space at this point [21]. Dimension of the

21) We emphasize that this vector space operates with linear combinations of vectors with *constant* numerical coefficients: a, b, etc. in order to comply with the Leibnitz rule 3.44.

vector space coincides with that of the manifold. Making use of the formalism of this section, it is straightforward to prove that all axioms of vector space introduced in section 2.3.1 remain valid in a tangent space.

It is important to understand that each point of the manifold $\mathcal{M}$ has its own tangent space. Different tangent spaces do not overlap, and are algebraically independent. Vectors residing in one tangent space have no relation to vectors from another tangent space, and they can not be compared to each other. Comparison between tangent spaces can be established if one introduces an additional geometric structure on manifold called *affine connection*. We shall discuss it in next section 3.4.

Spacetime manifold of general relativity and its tangent space are four-dimensional. Each tangent space in this case has four basis vectors $\boldsymbol{e}_\alpha$ $(\alpha = 0, 1, 2, 3)$ which are functions of the point on the manifold to which the tangent space is attached. Any vector $\boldsymbol{V}$ that belongs to the tangent space can be decomposed in basis components

$$\boldsymbol{V} = V^\alpha \boldsymbol{e}_\alpha \ . \tag{3.49}$$

Especially convenient for many applications is the case of a coordinate basis which basis vectors $\partial/\partial x^\alpha$ are tangent to the coordinate lines of the local coordinate grid. Coordinate basis has been used, for example, for decomposition of vector $\boldsymbol{V}$ in coordinate components in equation (3.45a). In particular, coordinate components of the coordinate basis vector $\partial/\partial x^\alpha$ coincide with the Kronecker symbol δ^β_α, where the index $\alpha = (0, 1, 2, 3)$ numerates the basis vectors, and the index $\beta = (0, 1, 2, 3)$ numerates the components of the vector in this basis. More explicitly, the components of the coordinate basis vectors in the chosen coordinates are

$$\frac{\partial}{\partial x^0} = (1, 0, 0, 0) \ , \quad \frac{\partial}{\partial x^1} = (0, 1, 0, 0) \ , \quad \frac{\partial}{\partial x^2} = (0, 0, 1, 0) \ , \quad \frac{\partial}{\partial x^3} = (0, 0, 0, 1) \ , \tag{3.50}$$

Had another basis, $\boldsymbol{e}_{\hat\alpha} = \Lambda^\beta{}_{\hat\alpha} \boldsymbol{e}_\beta$, been chosen in the tangent space, the vector $\boldsymbol{V}$ remained the same but its components would change

$$\boldsymbol{V} = V^{\hat\alpha} \boldsymbol{e}_{\hat\alpha} \ . \tag{3.51}$$

Comparing equations (3.49) and (3.51) yields the law of transformation of the vector components

$$V^\alpha = \Lambda^\alpha{}_{\hat\beta} V^{\hat\beta} \ . \tag{3.52}$$

Equation (3.52) is known as the vector transformation law. The matrix of transformation depends on the point of the manifold, and changes smoothly from one tangent space to another.

Vector field $\boldsymbol{V}$ on a manifold is a rule that assigns a specific numerical value to a tangent vector at each tangent space over manifold. Different vector fields define different tangent vectors. In local basis vector field $\boldsymbol{V}$ is defined by its components V^α which are smooth functions on the manifold.

3.3.6 Covectors and cotangent space

One has discussed covectors in section 2.4.2. They naturally appear in tensor algebra as linear functions mapping vectors from a vector space to numbers in the Euclidean space R^1. Now, one has to generalize this concept to the case of a differential manifold. It can be done similarly to what one did with definition of a vector in section 3.3.4. Vector at point P of a manifold was defined as a class of equivalence of smooth curves passing through this point. It led to the identification of vectors with the operators of differentiation along the curves.

Let us consider an arbitrary set of smooth functions f, g, h,..., in a neighborhood of a point P on a manifold. We shall call two functions, f and g, equivalent if they have the same first-order behavior near P. In local coordinates x^α the equivalence of the two functions means that their values and the values of their first partial derivatives coincide at the point P: $f(x_0) = g(x_0)$ and $(\partial f/\partial x^\alpha)_{x=x_0} = (\partial g/\partial x^\alpha)_{x=x_0}$, where x_0 are the local coordinates of the point P. Differential 1-form or covariant tangent vector or, simply, covector $\tilde{\boldsymbol{V}}$ at the point P of the manifold is a class of equivalence of all smooth functions at P. Covectors are identified with the differentials of the smooth functions.

The most important example of a covector is a gradient, denoted as $\tilde{\boldsymbol{d}}$. It is defined in the local coordinates of a neighborhood of a point P by the rule

$$\tilde{\boldsymbol{d}} f \equiv \frac{\partial f}{\partial x^\alpha} dx^\alpha \,, \tag{3.53}$$

where four numbers $\partial f/\partial x^\alpha$ define the coordinate components of the gradient which are contracted with the increments of the coordinates, dx^α, being tangent to the hypersurface of a constant value of function f passing through the point P. Though function f explicitly enters equation (3.53), the components of the gradient do not depend on its particular choice because the gradient is defined as a class of equivalence of all functions defined in the neighborhood of the point P. Hence, function f can be removed from the definition of gradient in the same way as a particular curve was removed from the definition of a tangent vector.

Covectors are linear functions on vectors. Gradient of function f is a linear function on vectors defined at a point P of a manifold as

$$\tilde{\boldsymbol{d}} f(\boldsymbol{V}) \equiv \boldsymbol{V}[f] \,, \tag{3.54}$$

where $\boldsymbol{V} = d/d\lambda$ is an arbitrary vector from the tangent space at point P, and the notation $\boldsymbol{V}[f] \equiv df/d\lambda$. All possible covectors at the point P of the manifold form a vector space which is called cotangent (or dual) to the tangent space of vectors. Because covectors are linear functions on vectors and vice versa, the dimension of the cotangent space coincides with that of the tangent space [22].

One can introduce a covector basis consisting of four linearly-independent covectors $\tilde{\omega}^\alpha$. They are defined in such a way that their values on four basis vectors $\boldsymbol{e}_\beta$

22) In general relativity this dimension is $n = 4$, which one assumes everywhere in this book.

yield a Kronecker symbol

$$\tilde{\boldsymbol{\omega}}^{\alpha}(\boldsymbol{e}_{\beta}) = \delta^{\alpha}_{\beta} \,. \tag{3.55}$$

In particular, in local coordinates x^{α} the basis covectors are $\tilde{\boldsymbol{d}}x^{\alpha}$. Their components are defined by equation

$$\tilde{\boldsymbol{d}}x^{\alpha}\left(\frac{\partial}{\partial x^{\beta}}\right) = \delta^{\alpha}_{\beta} \,, \tag{3.56}$$

where the index $\alpha = (0, 1, 2, 3)$ numerates the basis covectors, and the index $\beta = (0, 1, 2, 3)$ numerates the components of the covector in this basis. Solving equations (3.56) gives us,

$$\tilde{\boldsymbol{d}}x^{0} = (1, 0, 0, 0) \,, \qquad \tilde{\boldsymbol{d}}x^{1} = (0, 1, 0, 0) \,, \qquad \tilde{\boldsymbol{d}}x^{2} = (0, 0, 1, 0) \,, \qquad \tilde{\boldsymbol{d}}x^{3} = (0, 0, 0, 1) \,, \tag{3.57}$$

that should not be confused with the components of the basis vectors in equation (3.50).

Any covector $\tilde{\boldsymbol{W}}$ that belongs to the cotangent space can be decomposed in basis components

$$\tilde{\boldsymbol{W}} = W_{\alpha}\tilde{\boldsymbol{\omega}}^{\alpha} \,. \tag{3.58}$$

If another basis $\tilde{\boldsymbol{\omega}}^{\hat{\alpha}} = \Lambda^{\hat{\alpha}}{}_{\beta}\tilde{\boldsymbol{\omega}}^{\beta}$ were introduced in the cotangent space, the covector $\tilde{\boldsymbol{W}}$ remained the same but its components changed

$$\tilde{\boldsymbol{W}} = W_{\hat{\alpha}}\tilde{\boldsymbol{\omega}}^{\hat{\alpha}} \,. \tag{3.59}$$

Comparing equations (3.58) and (3.59) yields the law of transformation of the covector components

$$W_{\alpha} = \Lambda^{\hat{\beta}}{}_{\alpha}W_{\hat{\beta}} \,. \tag{3.60}$$

One can see that the law of transformation of covectors is inverse as compared with the law of transformation of vectors (3.52) because the matrices of transformations entering these laws are mutually inverse, $\Lambda^{\hat{\beta}}{}_{\alpha}\Lambda^{\alpha}{}_{\hat{\gamma}} = \delta^{\hat{\beta}}_{\hat{\gamma}}$.

Finally, one defines a covector field $\tilde{\boldsymbol{W}}$ on a manifold as a rule that assigns a specific numerical value to a covector at each cotangent space over manifold. Different covector fields define different covectors. In local basis the covector field $\tilde{\boldsymbol{W}}$ is defined by its components W_{α} which are smooth functions on the manifold.

3.3.7
Tensors

One has already discussed the definition and algebra of tensors belonging to the Minkowski spacetime in section 2.4. Each tangent and cotangent space of a manifold of general relativity has the same algebraic properties as the Minkowski spacetime of special relativity. Tensors residing in different tangent spaces form a tensor field over the manifold. Tensor fields are informally called tensors so that the word 'field' is often omitted. Examples are the metric tensor or the Riemann curvature tensor which will be discussed in section 3.7. If a basis is chosen, each component of a tensor (field) must be differentiable function as discussed in section 3.3.3. However, as one has learned in section 3.2.2, a partial derivative of a tensor does not make a tensor. Hence, the concept of differentiation of tensors on a manifold requires introduction of an additional geometric object on the manifold called affine connection, that allows one to define a new derivative having transformation properties of a tensor and to compare tensors located in different tangent spaces. The affine connection appeared rather formally in the Minkowski spacetime in curvilinear coordinates in the disguise of the Christoffel symbol $\Gamma^{\alpha}_{\beta\gamma}$, but it was not really required in the Minkowski spacetime because all of its tangent spaces are equivalent by definition. This is not the case of spacetime manifolds of general relativity. We shall introduce the reader to the concept of the affine connection in section 3.4 after discussing properties of the metric tensor.

3.3.8 The metric tensor

Metric tensor on manifold is a symmetric tensor field $\mathbb{G}$ that is introduced to endow the manifold with specific metrical properties of length, angles, etc. that are discussed in what follows. The manifold with the metric tensor is called the Riemannian or pseudo-Riemannian depending on whether the metric tensor is locally reduced to the Euclidean or Minkowski metric $\eta_{\alpha\beta} = \mathrm{diag}(-1, 1, 1, 1)$. Spacetime manifold in general theory of relativity and any viable alternative theory of gravity is pseudo-Riemannian with the metric tensor components denoted as $g_{\alpha\beta}$ in a local coordinate chart,

$$g_{\mu\nu} = \mathbb{G}\left(\boldsymbol{e}_\mu, \boldsymbol{e}_\nu\right) . \tag{3.61}$$

The metric tensor $\mathbb{G}$ is a bilinear form which is decomposed with respect to the basis vectors as follows (see section 2.4.3)

$$\mathbb{G} = g_{\mu\nu}\tilde{\omega}^\mu \otimes \tilde{\omega}^\nu . \tag{3.62}$$

3.3.8.1 *Operation of rising and lowering indices*

The tangent space and the cotangent space at a point P of a manifold $\mathcal{M}$ are both real vector spaces of the same dimension and therefore there must exist a mutual relationship, called duality or isomorphism, between their elements. This isomorphism is a rule of transformation of vectors to covectors and vice versa, and it can be established on a manifold in several different ways. In any case, its mathematical definition involves additional geometric structure on the manifold. In general relativity the role of the isomorphic structure between tangent and cotangent spaces is given to a pseudo-Riemannian metric $\mathbb{G}$ of spacetime manifold which is defined as a covariant tensor field of type $(0, 2)$ at each point of the manifold.

By definition, the metric tensor $\mathbb{G}$ converts a tangent vector $\boldsymbol{V}$ to a covector

$$\tilde{\boldsymbol{V}} \equiv \mathbb{G}(\boldsymbol{V}) , \tag{3.63}$$

by means of calculation of a value of the metric tensor on the vector $\boldsymbol{V}$. Let us pick up an arbitrary local dual basis consisting of vectors $\boldsymbol{e}_a$ and covectors $\tilde{\omega}^\alpha$. In this basis, vector $\boldsymbol{V} = V^\alpha \boldsymbol{e}_\alpha$ and the corresponding covector $\tilde{\boldsymbol{V}} = V_\alpha \tilde{\omega}^\alpha$, both residing at the same point of the manifold. Equation (3.63) is then equivalent to the following

$$V_a = g_{\alpha\beta} V^\beta , \tag{3.64}$$

where $g_{\alpha\beta}$ are components of the metric tensor in the given basis. Notice that one has used the same letter V for designation of the covector components with the index lowered. This is consistent with the definition of the isomorphism - one really works with a single geometric object transforming its components from one space to another by means of the metric tensor.

Definition of isomorphism between the tangent space and the cotangent space requires that the inverse operation must be valid that must give us the same vector. It requires introduction of a contravariant metric tensor of type $(2, 0)$ (cometric) with components $g^{\alpha\beta}$ that are inverse to $g_{\alpha\beta}$. The cometric and metric components are related to each other by equation

$$g_{\alpha\beta} g^{\beta\gamma} = \delta^\gamma_\alpha . \tag{3.65}$$

The cometric reverts components V_a of the covector as defined in equation (3.64), back to its original form

$$V^a = g^{\alpha\beta} V_\beta . \tag{3.66}$$

Mathematical operation defined in equations (3.64) and (3.66) are respectively called lowering and rising index (with the metric tensor). This operation can be applied to any tensor field at each point of the spacetime manifold and it works in the same way as discussed in section 2.4.4.3.

3.3.8.2 ***Magnitude of a vector and an angle between vectors***

The correspondence between the tangent and cotangent spaces established with the help of the metric tensor makes it possible to use it for calculation of a magnitude of a vector and an angle between two vectors at any point of the manifold. Magnitude of a vector $\boldsymbol{V}$ is defined as

$$||\boldsymbol{V}|| = \sqrt{|\mathbb{G}(\boldsymbol{V}, \boldsymbol{V})|} , \tag{3.67}$$

or in component notations

$$||\boldsymbol{V}|| = \sqrt{|g_{\alpha\beta} V^\alpha V^\beta|} = \sqrt{|V_\alpha V^\alpha|} , \tag{3.68}$$

which should be compared with equation (2.43). The angle, let say Φ, between two vectors $\boldsymbol{V}$ and $\boldsymbol{U}$ is formally defined as

$$\cos\Phi = \frac{\mathbb{G}(\boldsymbol{V}, \boldsymbol{U})}{||\boldsymbol{V}||\ ||\boldsymbol{U}||} , \tag{3.69}$$

or in component notations

$$\cos\Phi = \frac{g_{\alpha\beta}V^\alpha U^\beta}{\sqrt{|V_\alpha V^\alpha|}\sqrt{|U_\beta U^\beta|}} . \tag{3.70}$$

Since the metric on the spacetime manifold is pseudo-Riemannian, a care should be taken in calculation of the angle between two vectors to avoid singularity that may arise if one of the vectors is null-like, that is have magnitude equal to zero. One has to emphasize that in practice one can measure the angle only between spacelike vectors. The concept of the angular measurements in relativistic astrometry is discussed in section 7.

3.3.8.3 *The Riemann normal coordinates*

The metric tensor on a manifold is an analytic tensor field which can be expanded in a Taylor series around any point P of the manifold. Let us focus on a four-dimensional spacetime manifold of general relativity and introduce a local chart in the neighborhood of a point P with coordinates x^α centered at the point P, so that $x^\alpha(\mathrm{P}) = 0$. The components of the metric tensor in the vicinity of the point P are analytic functions that can be represented in the form of a Taylor series as a polynomial of the coordinates

$$g_{\alpha\beta}(x) = g_{\alpha\beta} + g_{\alpha\beta,\mu}x^\mu + \frac{1}{2!}g_{\alpha\beta,\mu\nu}x^\mu x^\nu + O\left(x^3\right) , \tag{3.71}$$

where the coefficients of the expansion are taken at point P, and the comma standing in front of indices denotes partial derivatives with respect to a corresponding coordinate,

$$g_{\alpha\beta,\mu} = \left[\frac{\partial g_{\alpha\beta}(x)}{\partial x^\mu}\right]_{x^\alpha=0} , \qquad g_{\alpha\beta,\mu\nu} = \left[\frac{\partial^2 g_{\alpha\beta}(x)}{\partial x^\mu \partial x^\nu}\right]_{x^\alpha=0} , \tag{3.72}$$

and so on.

We shall prove that by performing a coordinate transformation to a new local coordinates $w^{\hat\alpha}$, one can always make the new components of the metric equal to the Minkowski metric, $\hat g_{\hat\alpha\hat\beta} = \eta_{\hat\alpha\hat\beta}$, and *all* its first partial derivatives equal to zero, $\hat g_{\hat\alpha\hat\beta,\hat\mu} = 0$. The second derivatives of the metric, $\hat g_{\hat\alpha\hat\beta,\hat\mu\hat\nu} \neq 0$, in the most general case. In other words, in the new coordinates the Taylor expansion of the metric coefficients reads

$$\hat g_{\hat\alpha\hat\beta}(w) = \eta_{\hat\alpha\hat\beta} + \frac{1}{2!}\hat g_{\hat\alpha\hat\beta,\hat\mu\hat\nu}w^{\hat\mu}w^{\hat\nu} + O\left(w^3\right) . \tag{3.73}$$

The coordinates $w^{\hat\alpha}$ are called the *Riemann normal coordinates*. They can be thought of as a local realization of the orthogonal coordinates of the Minkowski spacetime.

Because the new coordinates, $w^{\hat\alpha}$, must be the analytic functions of the old coordinates, x^α, one can look for the coordinate transformation in the form of the Taylor expansion around the point P,

$$w^{\hat\alpha} = A^{\hat\alpha}{}_\mu x^\mu + \frac{1}{2!}B^{\hat\alpha}{}_{\mu\nu}x^\mu x^\nu + \frac{1}{3!}C^{\hat\alpha}{}_{\mu\nu\rho}x^\mu x^\nu x^\rho + O\left(x^3\right) , \tag{3.74}$$

where the expansion coefficients $A^{\hat{\alpha}}{}_{\mu}$, $B^{\hat{\alpha}}{}_{\mu\nu}$, $C^{\hat{\alpha}}{}_{\mu\nu\rho}$, and so on, are constant matrices. Transformation (3.74) suggests that the origins of both local coordinates coincide.

Transformation law of the metric tensor is

$$g_{\alpha\beta}(x) = \hat{g}_{\hat{\mu}\hat{\nu}}(w)\frac{\partial w^{\hat{\mu}}}{\partial x^{\alpha}}\frac{\partial w^{\hat{\nu}}}{\partial x^{\beta}} , \tag{3.75}$$

where $\hat{g}_{\hat{\mu}\hat{\nu}}(w)$ are the new components of the metric expressed in the new coordinates $w^{\hat{\alpha}}$ in accordance with equation (3.73). We calculate the partial derivatives of the coordinates with the help of equation (3.74), substitute equation (3.71) to the left side of the transformation law and equate coefficients of the polynomials in both sides of equation (3.75). It gives us the following system of algebraic equations

$$g_{\alpha\beta} = \eta_{\hat{\mu}\hat{\nu}}A^{\hat{\mu}}{}_{\alpha}A^{\hat{\nu}}{}_{\beta} , \tag{3.76a}$$

$$g_{\alpha\beta,\gamma} = \eta_{\hat{\mu}\hat{\nu}}\left(A^{\hat{\mu}}{}_{\alpha}B^{\hat{\nu}}{}_{\beta\gamma} + A^{\hat{\mu}}{}_{\beta}B^{\hat{\nu}}{}_{\alpha\gamma}\right) , \tag{3.76b}$$

$$\begin{aligned} g_{\alpha\beta,\gamma\delta} = {} & \hat{g}_{\hat{\mu}\hat{\nu},\hat{\rho}\hat{\sigma}}A^{\hat{\mu}}{}_{\alpha}A^{\hat{\nu}}{}_{\beta}A^{\hat{\rho}}{}_{\gamma}A^{\hat{\sigma}}{}_{\delta} + \\ & \eta_{\hat{\mu}\hat{\nu}}\left(A^{\hat{\mu}}{}_{\alpha}C^{\hat{\nu}}{}_{\beta\gamma\delta} + A^{\hat{\mu}}{}_{\beta}C^{\hat{\nu}}{}_{\alpha\gamma\delta} + B^{\hat{\mu}}{}_{\alpha\gamma}B^{\hat{\nu}}{}_{\beta\sigma} + B^{\hat{\mu}}{}_{\alpha\sigma}B^{\hat{\nu}}{}_{\beta\gamma}\right) , \end{aligned} \tag{3.76c}$$

for finding the coefficients of the coordinate transformation (3.74) from the known values of the metric tensor and its derivatives.

Equation (3.76a) is over-determined and can be always solved because the number of the independent components of $g_{\alpha\beta}$ is 10 (as the metric tensor is symmetric and the spacetime is four-dimensional) and the number of unknown elements of the matrix $A^{\hat{\mu}}{}_{\alpha}$ is $4 \times 4 = 16$. Hence, there is 6 degrees of freedom which corresponds to three Lorentz boosts and three spatial rotations - the transformations that leave the Minkowski metric invariant.

Left side of equation (3.76b) has $4 \times 10 = 40$ independent partial derivatives of the metric tensor, $g_{\alpha\beta,\gamma}$, and exactly the same number of the unknown coefficients $B^{\hat{\nu}}{}_{\beta\gamma}$ in its right side. Hence, it has a unique solution allowing us to annihilate all first derivatives of the metric tensor. This makes all the Christoffel symbols equal to zero in the normal coordinates as they are expressed in terms of the first derivatives of the metric tensor as shown in equation (3.34). The Christoffel symbols are associated with a homogeneous gravitational force. The fact that all of them can vanish at the origin of the normal coordinates is a consequence of the physical principle of equivalence.

The number of the second-order partial derivatives of the metric tensor, $g_{\alpha\beta,\gamma\delta}$, in the left side of equation (3.76c) is $10 \times 10 = 100$ while the number of the unknown coefficients $C^{\hat{\nu}}{}_{\beta\gamma\delta}$ in its ride side is only equal to $4 \times 20 = 80$. Since $80 < 100$, equation (3.76c) is under-determined and can not be solved in the most general case. As demonstrated later in section 3.7, the remaining 20 partial derivatives, $g_{\alpha\beta,\gamma\delta}$, of the metric tensor that are left over and do not match the transformation coefficients $C^{\hat{\nu}}{}_{\beta\gamma\delta}$, constitute a tensor of curvature of spacetime that is equal to zero only in the Minkowski spacetime. One can conclude that each (curved) metrical manifold is locally flat in the sense that the metric tensor can be reduced to the Minkowski metric up to the terms being quadratic with respect to the local coordinates. This property of the normal coordinates makes them physically equivalent to a reference frame falling freely in inhomogeneous gravitational field (see section 3.8.2).

It should be clearly stated that the procedure of construction of the normal coordinates can be applied to a differential manifold with the affine connection defined more generally than the Christoffel symbols. Indeed, the compatibility of the metric tensor with the affine connection has not been used in order to make all 40 components of the affine connection equal to zero at the origin of the normal coordinates. For this purpose the law of transformation (3.105) of the affine connection is sufficient [Schutz, 1995]. Normal coordinates represent a useful theoretical tool in proving various theorem of differential geometry on curved manifolds. The reason is that if some geometric property of tensor is satisfied in the normal coordinates, it will be valid in arbitrary coordinates due to the covariant nature of tensors and tensor transformations. Relativistic celestial mechanics uses more general approach to build the normal coordinates around massive bodies (planets, stars) composing N-body system. We discuss it in chapter 4.

3.4
Affine Connection on Manifold

In an elementary (Euclidean or Minkowski) vector space vectors can be moved about freely as long as one takes care not to change the magnitude or the direction of a vector. When one turns to the study of vectors on manifolds, however, it becomes apparent that one can no longer take this freedom of moving vectors from one point of the manifold to another as self-evident. Vectors are confined to their tangent space and the basic operations with vectors are only defined for vectors residing at the same point of the manifold.

To move a vector from one point to another, one needs to specify how this is to be done. It is achieved with the introduction of an affine connection that is a prescription for comparing vectors based at the tangent spaces attached to different points of a manifold. The connection consists of a set of linear transformations which tell us how to propagate vectors based at a particular point to infinitesimally nearby points. Unlike in elementary vector space where there is only one way of moving a vector from one point to another, there are many ways of moving vectors around differentiable manifold, so one needs to specify which connection one is using before one can move vectors from point to point.

This act of moving a vector from point to point is called parallel transport in analogy with the operation of elementary vector analysis which it generalizes. The parallel transport is a basic element in definition of the rule of differentiation of vectors (and tensors) on manifold. The derivative based on the parallel transport is called a covariant derivative as it preserves tensorial properties of the object under differentiation.

This section is intended to give the reader a handle on the geometric ideas underlying covariance and covariant derivative. A century ago, the geometric ideas communicated to Einstein by Marcel Grossmann, enabled Einstein to shape up the geometric approach to gravity. Nowadays, grasping mathematical extensions of Einstein's theory is impossible without understanding the concept of manifold, tangent space, the metric tensor and other geometric constructions, among which one of the most important is the affine connection.

Originally, the symbols $\Gamma^{\alpha}_{\beta\gamma}$ were introduced by Elwin Bruno Christoffel in the 1870s, to compensate the non-tensorial nature of the derivative from a vector field components expressed in curvilinear coordinates. Christoffel surely realized that the such corrected differential of a vector field transforms as a vector when the coordinates are changed arbitrary. However, a complete understanding of the mathematical universality of the compensating symbols $\Gamma^{\alpha}_{\beta\gamma}$ was achieved only by the end of the XIXth century. This was due to Tullio Levi-Civita who discovered that the description of parallel transport of a vector along a curved surface demands introduction of connection coefficients matching (connecting) components of the vector in different tangent spaces. It did not escape Levi-Civita's attention that the connection coefficients coincided with the Christoffel symbols $\Gamma^{\alpha}_{\beta\gamma}$. At first glance, the reason for this coincidence was not readily apparent, because Levi-Civita's connection originated from the surface's bending, while the Christoffel symbols emerged due to the distortion of the coordinate grid on a plane surface. It however did not take Levi-Civita long to provide a general framework within which it became clear why the two approaches, originating in seemingly different contexts, should overlap.

Above, section 3.2.2.1 advocated the Christoffel's viewpoint on the affine connection and covariant differentiation. This viewpoint is rather limited and does not embrace all specific features of a manifold endowed with an affine connection. Below the modern, axiomatic description of the affine connection will be presented which is originally due to Élie Cartan and Jean-Louis Koszul [1950]. In this approach the affine connection is considered from a rather general but abstract point of view. Later on, the affine connection will be linked to the metric tensor which will re-establish its major role in geometrized theories of gravity including general relativity. The axiomatic approach to affine connection was further extended by Charles Ehresmann [1951] (a student of É. Cartan) in relation to fiber bundles [23].

We would like to mention that in earlier developments of differential geometry on manifold, the affine connection was often introduced with the help of the projection operator from an external Euclidean space of higher dimensions to which the manifold under consideration was embedded. This approach has been abandoned in favor of the intrinsic definition because one cannot speak of moving 'outside' the geometric object. Moreover, the intrinsic point of view is more flexible as it admits many different types of connections (Cartan, Koszul, Ehresmann, etc.) on various differential manifolds.

3.4.1
Axiomatic definition of the affine connection

We consider a manifold $\mathcal{M}$ with the metric tensor $g_{\alpha\beta}$ defined at each point of it. As discussed in section 2.3.4, the metric is a symmetric tensor field of rank 2, which defines the interval ds between two infinitesimally close events.

The affine connection is an additional geometric structure on the manifold $\mathcal{M}$, defining the law of parallel transport of tensors from one point of the manifold to another, and the rule of the covariant differentiation. In the most general case, affine connection is not related to the metric tensor. However, in general theory of relativity (and in many other geometrized theories of gravitation) the connection is uniquely linked to the metric tensor and is expressed in terms of its components $g_{\alpha\beta}$. Such connection is called metrical and is also known under the name of the Levi-Civita connection. If a basis is chosen the components of the Levi-Civita connection are denoted $\Gamma^{\alpha}_{\beta\gamma}$. In a coordinate basis the components of the Levi-Civita connection are called the Christoffel symbols.

In general, an affine connection is a linear operator of differentiation, $\boldsymbol{\nabla}$, having two arguments

$$\boldsymbol{\nabla} \equiv \boldsymbol{\nabla}_{(\text{argument \#2})}(\text{argument \#1}) \,, \tag{3.77}$$

23) A fiber bundle is intuitively a space which locally "looks" like a product of two other spaces, but globally may have a different topological structure [Schutz, 1995].

where the slot for argument #1 is for a tensor field, which is differentiated, and the subscript slot for argument #2 is for a vector (and only for a vector), along which the derivative is taken. For example, if one differentiate a tensor $\mathbb{T}$ without specifying the vector in the argument #2, the result of the differentiation is written as $\boldsymbol{\nabla}\mathbb{T}$. If one decides to take the derivative of the same tensor along a vector $\boldsymbol{X}$, the result of this operation is written as $\boldsymbol{\nabla}_{\boldsymbol{X}}\mathbb{T}$. We emphasize that the argument #2 can be used only for a *vector*. It does not work for covector or any other tensor. The affine connection defines mapping of a C^r-differentiable tensor field $\mathbb{T}$ of type (k, l) to a C^{r-1}-differentiable tensor field $\boldsymbol{\nabla}\mathbb{T}$ of type $(k, l+1)$ [24]. The operator $\boldsymbol{\nabla}$ satisfies at each point P of the manifold $\mathcal{M}$ the following five axioms:

A1. Algebraic linearity, in regard to the argument #1

$$\boldsymbol{\nabla}(\alpha\mathbb{T} + \beta\mathbb{S}) = \alpha\boldsymbol{\nabla}\mathbb{T} + \beta\boldsymbol{\nabla}\mathbb{S}\ , \tag{3.78}$$

where $\mathbb{T}$ and $\mathbb{S}$ are tensor fields of the same type, while α and β are constant numbers. The linearity is violated if α and β are scalar functions due to the Leibnitz rule (axiom A3). It makes the affine connection not a tensor.

A2. Linearity, in regard to the argument #2

$$\boldsymbol{\nabla}_{f\boldsymbol{X}+g\boldsymbol{Y}}\mathbb{T} = f\boldsymbol{\nabla}_{\boldsymbol{X}}\mathbb{T} + g\boldsymbol{\nabla}_{\boldsymbol{Y}}\mathbb{T}\ , \tag{3.79}$$

where $\boldsymbol{X}$ and $\boldsymbol{Y}$ are vectors, while f and g are C^1-differentiable functions. This axiom tells us that $\boldsymbol{\nabla}_{\boldsymbol{X}}\mathbb{T}$ at point P depends on the value of $\boldsymbol{X}$ only at P. One does not need to know behavior of the vector field $\boldsymbol{X}$ in the neighborhood of P.

A3. The Leibniz rule for tensors

$$\boldsymbol{\nabla}(\mathbb{T}\otimes\mathbb{S}) = (\boldsymbol{\nabla}\mathbb{T})\otimes\mathbb{S} + \mathbb{T}\otimes\boldsymbol{\nabla}\mathbb{S}\ , \tag{3.80}$$

and for scalars

$$\boldsymbol{\nabla}(f\mathbb{T}) = f\boldsymbol{\nabla}\mathbb{T} + \mathbb{T}\boldsymbol{\nabla}f\ . \tag{3.81}$$

A4. Commutativity with the operation of transvection of a tensor (defined in section 2.4.4.4)

A5. Coincidence with the operator of gradient, when acting on an arbitrary scalar field f, that is

$$\boldsymbol{\nabla}f = \tilde{\boldsymbol{d}}f\ . \tag{3.82}$$

where the operator of gradient, $\tilde{\boldsymbol{d}}f$ has been defined in equations (3.53), (3.54).

24) Tensor of type (k, l) has k contravariant and l covariant indices. C^r-differentiable field means that it has all derivatives up to the order r.

3.4.2 Components of the connection

We consider a neighborhood of a point P on spacetime manifold $\mathcal{M}$. Let us introduce a vector basis $\boldsymbol{e}_\alpha$ and the dual covector basis $\tilde{\boldsymbol{\omega}}^\alpha$ which satisfy to equation

$$\boldsymbol{e}_\alpha(\tilde{\boldsymbol{\omega}}^\beta) = \tilde{\boldsymbol{\omega}}^\beta(\boldsymbol{e}_\alpha) = \delta^\beta_\alpha \,. \tag{3.83}$$

Since one is not working in the Minkowski spacetime, the vector basis can not be fixed on the entire manifold $\mathcal{M}$ - it depends on point and changes as one goes from one point of the manifold to another. Components $\mathfrak{G}^\alpha_{\beta\gamma}$ of the affine connection $\boldsymbol{\nabla}$ quantify the change of the basis vectors as one moves on the manifold from point to point. More specifically, in the neighborhood of the point P with the basis chosen, the components of the affine connection are defined by

$$\boldsymbol{\nabla}\boldsymbol{e}_\alpha = \mathfrak{G}^\beta_{\alpha\gamma}\boldsymbol{e}_\beta \otimes \tilde{\boldsymbol{\omega}}^\gamma \,, \tag{3.84}$$

where $\mathfrak{G}^\beta_{\alpha\gamma}$ has in the most general case, $4^3 = 64$ components, each of which may have different values at different points of the manifold. By virtue of axiom A3, and the condition of the duality of vector and covector bases, one obtains from equation (3.84)

$$\boldsymbol{\nabla}\tilde{\boldsymbol{\omega}}^\alpha = -\mathfrak{G}^\alpha_{\beta\gamma}\tilde{\boldsymbol{\omega}}^\beta \otimes \tilde{\boldsymbol{\omega}}^\gamma \,, \tag{3.85}$$

where $\mathfrak{G}^\alpha_{\beta\gamma}$ are the same connection coefficients as in equation (3.84).

Equations (3.84), (3.85) do not specify the direction along which the derivative is taken. Let us specify the change of the vector field $\boldsymbol{e}_\alpha$ along a basis vector $\boldsymbol{e}_\beta$. We shall use an abbreviation

$$\boldsymbol{\nabla}_{\boldsymbol{e}_\beta} \equiv \boldsymbol{\nabla}_\beta \,, \tag{3.86}$$

for the covariant derivative along vector $\boldsymbol{e}_\beta$. Making use of the definition (3.77) of the affine connection in equation (3.84), one obtains

$$\boldsymbol{\nabla}_\beta\boldsymbol{e}_\alpha = \mathfrak{G}^\gamma_{\alpha\beta}\boldsymbol{e}_\gamma \,, \tag{3.87}$$

where one has used the property (3.83) of the duality of vector and covector bases and axiom A2 of the linearity of the operator $\boldsymbol{\nabla}$ with respect to its vector argument. One draws attention of the reader to the order of indices α and β in equation (3.87) - it is fixed by the definition and is not arbitrary since in the most general case, the connection coefficients $\mathfrak{G}^\gamma_{\alpha\beta}$ are not symmetric with respect to α and β. Proceeding similarly, one obtains from equation (3.85) the derivative of the covector $\tilde{\boldsymbol{\omega}}^\beta$ along the basis vector $\boldsymbol{e}_a$

$$\boldsymbol{\nabla}_\beta\tilde{\boldsymbol{\omega}}^\alpha = -\mathfrak{G}^\alpha_{\gamma\beta}\tilde{\boldsymbol{\omega}}^\gamma \tag{3.88}$$

Axioms of section 3.4.1 and equations (3.87), (3.88) allow one to calculate the components of the covariant derivative of an arbitrary tensor field $\mathbb{S}$.

3.4.3
Covariant derivative of tensors

A tensor field $\mathbb{S}$ of type (k,l) is decomposed in the basis vectors and covectors as follows (see section 2.4.4.1)

$$\mathbb{S} = S^{\alpha_1\alpha_2...\alpha_k}_{\beta_1\beta_2...\beta_l}\boldsymbol{e}_{\alpha_1}\otimes\boldsymbol{e}_{\alpha_2}\otimes...\otimes\boldsymbol{e}_{\alpha_k}\otimes\tilde{\boldsymbol{\omega}}^{\beta_1}\otimes\tilde{\boldsymbol{\omega}}^{\beta_2}\otimes...\otimes\tilde{\boldsymbol{\omega}}^{\beta_l}\ , \tag{3.89}$$

where $S^{\alpha_1\alpha_2...\alpha_k}_{\beta_1\beta_2...\beta_l}$ are differentiable scalar fields of the same order. Covariant derivative of the tensor $\mathbb{S}$ along a basis vector $\boldsymbol{e}_\mu$ is a new tensor of type $(k,l+1)$ denoted as $\boldsymbol{\nabla}_\mu\mathbb{S}$, and it has components $\boldsymbol{\nabla}_\mu S^{\alpha_1\alpha_2...\alpha_k}_{\beta_1\beta_2...\beta_l}$ such that [25)]

$$\boldsymbol{\nabla}_\mu\mathbb{S} \equiv \boldsymbol{\nabla}_\mu S^{\alpha_1\alpha_2...\alpha_k}_{\beta_1\beta_2...\beta_l}\boldsymbol{e}_{\alpha_1}\otimes\boldsymbol{e}_{\alpha_2}\otimes...\otimes\boldsymbol{e}_{\alpha_k}\otimes\tilde{\boldsymbol{\omega}}^{\beta_1}\otimes\tilde{\boldsymbol{\omega}}^{\beta_2}\otimes...\otimes\tilde{\boldsymbol{\omega}}^{\beta_l}\otimes\tilde{\boldsymbol{\omega}}^{\mu}\ . \tag{3.90}$$

Exact expression for the components is obtained after direct application of the operator $\boldsymbol{\nabla}$ to both sides of equation (3.89) and making use of the axioms 1-5, and equations (3.87), (3.88). One gets

$$\begin{aligned}\boldsymbol{\nabla}_\mu S^{\alpha_1\alpha_2...\alpha_k}_{\beta_1\beta_2...\beta_l} &= \partial_\mu S^{\alpha_1\alpha_2...\alpha_k}_{\beta_1\beta_2...\beta_l}\\ &+ \mathfrak{G}^{\alpha_1}_{\nu\mu}S^{\nu\alpha_2...\alpha_k}_{\beta_1\beta_2...\beta_l}+...+\mathfrak{G}^{\alpha_k}_{\nu\mu}S^{\alpha_1\alpha_2...\nu}_{\beta_1\beta_2...\beta_l}\\ &- \mathfrak{G}^{\nu}_{\beta_1\mu}S^{\alpha_1\alpha_2...\alpha_k}_{\nu\beta_2...\beta_l}-...-\mathfrak{G}^{\nu}_{\beta_l\mu}S^{\alpha_1\alpha_2...\alpha_k}_{\beta_1\beta_2...\nu}\ ,\end{aligned} \tag{3.91}$$

where

$$\partial_\mu S^{\alpha_1\alpha_2...\alpha_k}_{\beta_1\beta_2...\beta_l} \equiv \boldsymbol{e}_\mu\left[S^{\alpha_1\alpha_2...\alpha_k}_{\beta_1\beta_2...\beta_l}\right]\ , \tag{3.92}$$

is the value of the tensor component $S^{\alpha_1\alpha_2...\alpha_k}_{\beta_1\beta_2...\beta_l}$ on a basis vector $\boldsymbol{e}_\mu$ that is a partial derivative of $S^{\alpha_1\alpha_2...\alpha_k}_{\beta_1\beta_2...\beta_l}$ along the direction of this vector. In case of a coordinate basis $\boldsymbol{e}_\alpha = \partial/\partial x^\alpha$ and $\tilde{\boldsymbol{\omega}}^\alpha = \boldsymbol{d}x^\alpha$, the partial derivative $\partial_\mu = \partial/\partial x^\mu$.

Equation (3.91) makes it more clear how the covariant derivative works for a tensor of type (k,l). It consists of the partial derivative of the components of the tensor, and $k+l$ terms describing the coupling of the tensor' components with the coefficients of the affine connection. From these terms, k contravariant indices couples with the affine connection with the sign plus, and l covariant indices couples with the affine connection with the sign minus. The covariant derivative tensor is of type $k,l+1$. It is worth emphasizing that the order of indices in the connection coefficients is fixed because one operates with the most general form of the affine connection and have not yet imposed any limitations on its symmetry. Only the metric connection in coordinate bases is symmetric as explained below in section 3.5.4.

3.4.4
Parallel transport of tensors

3.4.4.1 *Equation of the parallel transport*

Let $C(\tau)$ be a differentiable curve on a manifold $\mathcal{M}$ with τ being a parameter. By definition, a tensor $\mathbb{T}$ of an arbitrary rank is subject to a parallel transport along the

25) Another common notation for the covariant derivative of a tensor is with a semicolon, $S^{\alpha_1\alpha_2...\alpha_k}_{\beta_1\beta_2...\beta_l;\mu}$, which is not used in the book.

curve if its covariant derivative along a tangent vector $\boldsymbol{V} = \boldsymbol{V}(\tau)$ to the curve, is zero [26)]

$$\boldsymbol{\nabla}_{\boldsymbol{V}}\mathbb{T} = 0\,. \tag{3.93}$$

If a basis $\boldsymbol{e}_\alpha$ is chosen along the curve, the tangent vector $\boldsymbol{V} = V^\alpha \boldsymbol{e}_a$, and the equation of the parallel transport of the tensor components becomes

$$V^\alpha \boldsymbol{\nabla}_\alpha \mathbb{T} = 0\,, \tag{3.94}$$

where the covariant derivative $\boldsymbol{\nabla}_\alpha \mathbb{T}$ has been defined in equation (3.91). The result of the parallel transport of a tensor from one point of the manifold $\mathcal{M}$ to another will in general depend on the choice of the curve connecting these two points [27)]. If it does not, the manifold is called *flat*, and its *curvature* is zero (see Figure 3.7 for further details).

Let us now consider a tensor field $\mathbb{W}$ on a manifold $\mathcal{M}$ and two points P and Q on the curve $C(\tau)$ with a tangent vector $\boldsymbol{V}$. Tensor field $\mathbb{W}$ has a value $\mathbb{W}(\mathrm{P})$ at point P, and a value $\mathbb{W}(\mathrm{Q})$ at point Q. We take the value $\mathbb{W}(\mathrm{P})$ as an initial condition for the equation of the parallel transport of the tensor $\mathbb{W}$ along the given curve, and solve it. It generates a new tensor field $\mathbb{W}^*$ on the curve that obeys the ordinary differential equation $\boldsymbol{\nabla}_{\boldsymbol{V}}\mathbb{W}^* = 0$. It means that if the point Q is sufficiently close to P, the new tensor $\mathbb{W}^*(\mathrm{Q}) = \mathbb{W}^*(\mathrm{P}) = \mathbb{W}(\mathrm{P})$. In general, the new field $\mathbb{W}^*(\mathrm{Q}) \neq \mathbb{W}(\mathrm{Q})$, and the covariant derivative $\boldsymbol{\nabla}_{\boldsymbol{V}}\mathbb{W} \neq 0$. In some cases, however, the parallel-transported tensor $\mathbb{W}^*(\mathrm{Q})$ will coincide with the value of the tensor $\mathbb{W}(\mathrm{Q})$ at Q. Such tensor field is called covariantly constant along the curve, and for such a tensor $\mathbb{W}(\mathrm{P}) = \mathbb{W}(\mathrm{Q})$. Notice that if a tensor field is covariantly constant along one curve, it may not remain constant along another curve. This is because covariant derivatives do not commute as explained below in section 3.7.1. Figure 3.1 visualizes the idea of the parallel transport of a vector $\boldsymbol{W}$ along a differentiable curve with a tangent vector $\boldsymbol{V}$.

An important example of a covariantly constant tensor field is delivered by a metric tensor which defines the metric properties (length, angles, etc.) in each tangent space of a manifold. We discuss this property of the metric tensor below in section 3.5.4. Covariantly constant vector fields are generated by a curves commonly known as geodesics.

3.4.4.2 ***Geodesics***

An important case of a covariantly constant field is obtained when equation (3.94) is applied to the curve's tangent vector $\boldsymbol{V}$ itself

$$\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{V} = 0\,, \tag{3.95}$$

If the transportation curve is parameterized with a parameter τ, the coordinate components of the tangent vector $V^\alpha(\tau) = dx^\alpha/d\tau$ at each point of the curve. The equa-

26) Equation $\boldsymbol{\nabla}_{\boldsymbol{V}}\mathbb{T} = \lambda \boldsymbol{T}$ with λ being a smooth function on the transportation curve, is not accepted in general relativity as it does not preserves lengths of vectors, and is not consistent with the definition of the Levi-Civita connection introduced later in section 3.5.4.

27) This "noncommutation" of the parallel transport is explained in section 3.7.1.

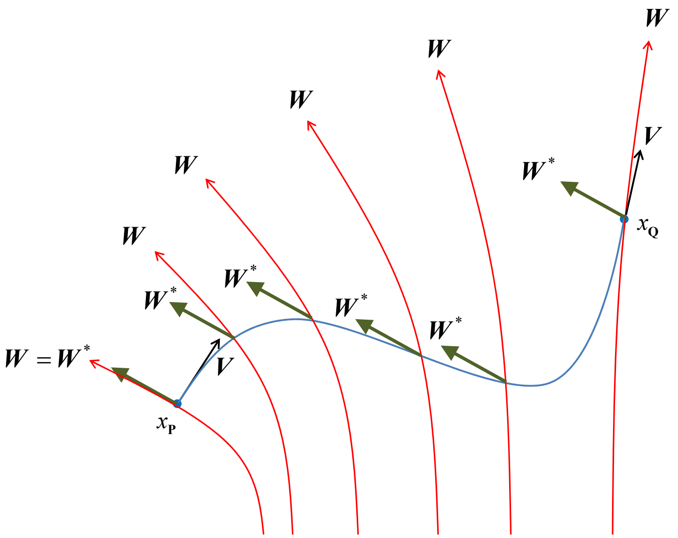

Figure 3.1 A vector $\boldsymbol{W}$ is parallel transported along a smooth curve from a point P to a point Q. It generates a new tensor field $\boldsymbol{W}^*$ along the curve. The field $\boldsymbol{W}^*$ obeys equation $\boldsymbol{\nabla}_{\boldsymbol{V}}\mathbb{W}^* = 0$ and coincides with $\boldsymbol{W}$ at the point P. However, in the most general case the field $\boldsymbol{W}^*$ is different from the field $\boldsymbol{W}$ in other points of the transportation curve. In the situation, when $\boldsymbol{W}^* = \boldsymbol{W}$ along the curve, the field $\boldsymbol{W}$ is called covariantly constant along the given curve.

tion (3.95) of the parallel transport of vector $\boldsymbol{V}$ along itself is called the equation of geodesic [28], and its explicit coordinate form reads

$$\frac{dV^\alpha}{d\tau} + \mathfrak{G}^\alpha_{\beta\gamma} V^\beta V^\gamma = 0\,, \tag{3.96}$$

or

$$\frac{d^2x^\alpha}{d\tau^2} + \mathfrak{G}^\alpha_{\beta\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} = 0\,. \tag{3.97}$$

Equation (3.97) shows that the parallel transport of V^a depends only on the symmetric part $\mathfrak{G}^\alpha_{(\beta\gamma)} \equiv (\mathfrak{G}^\alpha_{\beta\gamma} + \mathfrak{G}^\alpha_{\gamma\beta})/2$ of the connection and on the values of V^a on the curve $x^a(\tau)$, so one does not need to know (or to define) the entire vector field V^a around the curve. Solution of this equation is called a *geodesic* worldline or simply a *geodesic*.

The geodesic is a generalization of the concept of a straight line in the Minkowski spacetime, in the sense that tangent lines to geodesic remain parallel as one moves from one to another point along the curve $x^\alpha(\tau)$. It should be noticed, however, that in the (pseudo) Riemannian geometry geodesics may not be the same as the "shortest curves" between two points, though the two concepts are closely related. The point is that geodesics can only locally represent the shortest distance between points. Therefore, geodesics can not be, in general, defined as the curves of a minimal length on topologically non-trivial manifolds. Furthermore, equation of the curve of a minimal length is derived from a variational principle,

$$\delta \int g_{\alpha\beta} dx^\alpha dx^\beta = 0\,, \tag{3.98}$$

with the variations of the coordinates being fixed at the end points. It leads to equation [Misner *et al.*, 1973]

$$\frac{d^2x^\alpha}{d\tau^2} + \Gamma^\alpha_{\beta\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} = 0\,, \tag{3.99}$$

depending only on the Christoffel symbols $\Gamma^\alpha_{\beta\gamma}$ [29]. The Christoffel symbols are a metric-related part of the affine connection $\mathfrak{G}^\alpha_{\beta\gamma}$, which can have other ingredients, as explained in equation (3.121). In general relativity equations of geodesics and the curves of the minimal length, coincide. In Einstein's gravitational physics, geodesics describe the motion of point-like particles under the influence of gravity alone; all other non-gravitational forces are absent if the particle moves along a geodesic.

There is a certain freedom in choosing parameter τ along the geodesic. Equation (3.96) demands that not only vector $\boldsymbol{V}$ remained parallel to itself but its norm did not change: $|\boldsymbol{V}| = \left(g_{\alpha\beta}V^\alpha V^\beta\right)^{1/2}$=const., which ensures the identity (isomorphism) of

28) Equation (3.95) is also called as equation of *auto-parallel transport* [Kleinert and Pelster, 1999; Manoff, 2000].

29) Kleinert and Pelster [1999] explain how to generalized this variational principle in order to take into account the part of the affine connection depending on torsion.

tangent spaces attached to different points of the manifold. If the latter requirement is relaxed the equation of geodesic becomes

$$\frac{d^2x^\alpha}{d\tau^2} + \mathfrak{G}^\alpha_{\beta\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} = \lambda\frac{dx^\alpha}{d\tau} , \tag{3.100}$$

where $\lambda = \lambda(\tau)$ is a smooth but arbitrary function along the geodesic. Equation (3.100) defines a class of equivalence of geodesics with a rather large freedom in choosing parametrization τ along the geodesic. For example, if one chooses another parametrization, $\sigma = \sigma(\tau)$, equation (3.100) becomes

$$\frac{dV^\alpha}{d\sigma} + \mathfrak{G}^\alpha_{\beta\gamma}\frac{dx^\beta}{d\sigma}\frac{dx^\gamma}{d\sigma} = \left(\frac{d\sigma}{d\tau}\right)^{-2}\left(\lambda\frac{d\sigma}{d\tau} - \frac{d^2\sigma}{d\tau^2}\right)\frac{dx^\alpha}{d\sigma} , \tag{3.101}$$

making evident that function λ determines parametrization of the geodesic.

One has adopted in section 3.3.4 to distinguish between two curves having the same geometric image on a manifold (passing through the same points) but parameterized differently. It compels us to constrain the form of the equation of geodesic by fixing parameter λ in equation (3.100). The most convenient choice is $\lambda = 0$ which gets us back to equation (3.97). A parametrization that yields equation (3.97) is called an *affine parametrization*. The affine parameter τ is defined up to a linear transformation $\tau \to a\tau + b$, where a and b are arbitrary constant numbers with $a \neq 0$. The affine parametrization of geodesics is frequently adopted in Einstein's general relativity.

3.4.5
Transformation law for connection components

Let us chose for simplicity a coordinate basis $\boldsymbol{e}_\alpha = \partial/\partial x^a$. The components of the affine connection in a basis $\partial/\partial x^a$ and the dual basis of covectors $\tilde{\omega}^\alpha = \boldsymbol{d}x^\alpha$, are given at the point P of manifold $\mathcal{M}$ by equation (3.84). Let us change the basis in the neighborhood of point P to a new basis

$$\boldsymbol{e}_{\alpha'} = \Lambda^\beta{}_{\alpha'}\boldsymbol{e}_\beta , \qquad \tilde{\omega}^{\alpha'} = \Lambda^{\alpha'}{}_\beta\tilde{\omega}^\beta , \tag{3.102}$$

where $\Lambda^\beta{}_{\alpha'}$ is the matrix of transformation, and $\Lambda^{\alpha'}{}_\beta$ is its inverse

$$\Lambda^{\alpha'}{}_\beta\Lambda^\beta{}_{\gamma'} = \delta^{\alpha'}_{\gamma'} , \qquad \Lambda^\alpha{}_{\beta'}\Lambda^{\beta'}{}_\gamma = \delta^\alpha_\gamma . \tag{3.103}$$

Both the matrix of transformation and its inverse are differentiable in the neighborhood of the point P.

Components of the affine connection in the new basis are defined by equation being similar to equation (3.84), that is

$$\boldsymbol{\nabla}\boldsymbol{e}_{\alpha'} = \mathfrak{G}^{\beta'}_{\alpha'\gamma'}\boldsymbol{e}_{\beta'} \otimes \tilde{\omega}^{\gamma'} . \tag{3.104}$$

Substituting the transformation equations (3.102) to equation (3.104, and making use of the Leibnitz rule for the covariant derivative, one obtains the law of transformation of the affine connection

$$\mathfrak{G}^{\alpha'}_{\beta'\gamma'} = \Lambda^{\alpha'}{}_{\rho}\Lambda^{\mu}{}_{\beta'}\Lambda^{\nu}{}_{\gamma'}\mathfrak{G}^{\rho}_{\mu\nu} + \Lambda^{\alpha'}{}_{\rho}\frac{\partial \Lambda^{\rho}{}_{\beta'}}{\partial x^{\gamma'}} . \tag{3.105}$$

Presence of the second term in the right side of this equation reveals that the affine connection does not transform as a tensor in the most general case. In case of the coordinate basis, equation (3.105) is fully compatible with equation (3.16) of the (inverse) transformation of the Christoffel symbols.

One remarkable consequence of the transformation law (3.105) is that if one has two different affine connections on the same manifold, $\mathfrak{G}^{\alpha}_{\beta\gamma}$ and $\mathfrak{P}^{\alpha}_{\beta\gamma}$, the difference between the connections

$$\mathfrak{D}^{\alpha}_{\beta\gamma} \equiv \mathfrak{G}^{\alpha}_{\beta\gamma} - \mathfrak{P}^{\alpha}_{\beta\gamma} , \tag{3.106}$$

is tensor of a third rank and of type $(1, 2)$.

3.5 The Levi-Civita Connection

After defining the affine connection and deriving the properties of covariant derivative, one will eventually be faced with the following question: how arbitrarily can one set the values of $\mathfrak{G}^{\alpha}_{\beta\gamma}$ on the manifold $\mathcal{M}$ besides the requirement that they must be, at least, smooth functions of the local coordinates? The fact of the matter is that on any differentiable manifold there are infinitely many affine connections. Equation (3.84) defining the coefficients of connection, imposes none restriction on the symmetry of $\mathfrak{G}^{\alpha}_{\beta\gamma}$ nor on a particular structure of them. Hence, in a four-dimensional case the number of the connection coefficients is $4 \times 4 \times 4 = 64$. However, general relativity theory is quite specific about the nature of gravity and associates gravitational potential with a pseudo-Riemannian metric tensor $g_{\alpha\beta}$ that is defined on the spacetime manifold $\mathcal{M}$ from the Einstein's field equations. If the manifold is endowed with a pseudo-Riemannian metric, then, there exists a natural choice of the affine connection, called the Levi-Civita connection $\Gamma^{\alpha}_{\beta\gamma}$ that warrants the identity of algebraic properties of tangent spaces at different points of the manifold. The Levi-Civita connection is selected by imposing the torsion-free property on the connection coefficients and linking it to the metric tensor by demanding that the covariant derivative of the metric tensor (called *nonmetricity*) vanishes. Levi-Civita connection is also known as the metrical connection. In a coordinate basis the Levi-Civita connection coincides with the Christoffel symbols. In orthonormal (tetrad) vector basis the components of the Levi-Civita connection are called the Ricci rotation coefficients. Further discussion of the metric connection requires introduction of new geometric objects associated with vectors fields and the covariant derivative.

3.5.1 Commutator of two vector fields

Let a neighborhood of a point P be covered by local coordinates with a basis $\partial/\partial x^\alpha$. Consider in this neighborhood two vector fields, $\boldsymbol{V} = V^\alpha \partial/\partial x^\alpha$ and $\boldsymbol{U} = U^\alpha \partial/\partial x^\alpha$ with components V^α, U^α being differentiable functions. Commutator of two vector fields is defined as an operator of partial differentiation

$$[\boldsymbol{V}, \boldsymbol{U}] \equiv \boldsymbol{V}\boldsymbol{U} - \boldsymbol{U}\boldsymbol{V} \,. \tag{3.107}$$

Components of the commutator can be readily calculated if one applies equation (3.107) to a scalar function f. One has

$$\begin{aligned} [\boldsymbol{V}, \boldsymbol{U}](f) &= V^\beta \frac{\partial}{\partial x^\beta}\left(U^\alpha \frac{\partial f}{\partial x^\alpha}\right) - U^\beta \frac{\partial}{\partial x^\beta}\left(V^\alpha \frac{\partial f}{\partial x^\alpha}\right) \\ &= \left(V^\beta \frac{\partial U^\alpha}{\partial x^\beta} - U^\beta \frac{\partial V^\alpha}{\partial x^\beta}\right) f \,. \end{aligned} \tag{3.108}$$

because second partial derivatives of a scalar function are always commutative. Thus, the commutator is a linear vector operator with coefficients $W^\alpha = [\boldsymbol{V}, \boldsymbol{U}](\tilde{\boldsymbol{d}}x^\alpha)$ given in the coordinate basis by formula

$$W^\alpha = V^\beta \frac{\partial U^\alpha}{\partial x^\beta} - U^\beta \frac{\partial V^\alpha}{\partial x^\beta} \,. \tag{3.109}$$

In the case of an arbitrary basis $\boldsymbol{e}_\alpha$, the commutator of the basis vectors is defined by

$$\left[\boldsymbol{e}_\alpha, \boldsymbol{e}_\beta\right] = C^\gamma_{\alpha\beta} \boldsymbol{e}_\gamma \,, \tag{3.110}$$

where the $C^\gamma_{\alpha\beta} = -C^\gamma_{\beta\alpha}$ are anti-symmetric commutation coefficients. The number of algebraically-independent coefficients of $C^\gamma_{\alpha\beta}$ is $4 \times 6 = 24$.

In arbitrary (non-coordinate) basis $\tilde{\omega}^\alpha$ the components of the commutator of the fields, $\boldsymbol{V}$ and $\boldsymbol{U}$, are

$$[\boldsymbol{V}, \boldsymbol{U}](\tilde{\omega}^\alpha) = V^\beta \boldsymbol{e}_\beta(U^\alpha) - U^\beta \boldsymbol{e}_\beta(V^\alpha) + C^\alpha_{\beta\gamma} V^\beta U^\gamma \,. \tag{3.111}$$

In any coordinate basis $C^\alpha_{\beta\gamma} = 0$ because coordinate grid consists by definition from vector fields that are mutually constant on each other. Conversely, if one chooses the vector field $\boldsymbol{V}$ constant on the lines of the field $\boldsymbol{U}$, and visa versa, the commutator of these two fields will be zero: $[\boldsymbol{V}, \boldsymbol{U}] = 0$. A set of any N vector fields with mutually vanishing commutators are constant with respect to each other, and hence can be used to build a local coordinate grid on the N-dimensional manifold. We emphasize that the property of the vector fields to form a coordinate grid is related to a neighborhood of a point on the manifold but not to the point itself as the commutator of the fields depends both on their values at the point and on the values of their partial derivatives.

Geometric interpretation of commutator of two vector fields is given in the Figure 3.2 which shows the coordinate lines corresponding to Cartesian and polar coordinates in the plane, and two paths corresponding to the motion along the vector fields. If commutator of two vector fields vanishes, the fields are called holonomic. In opposite case, the two fields are called non-holonomic. Further discussion of geometric interpretation of commutator of two vector fields is given below in section 3.6.3 as well as in the Figure 3.5.

3.5.2
Torsion tensor

Next step towards the metric connection demands a comparison of a differentiation operator $\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{U} - \boldsymbol{\nabla}_{\boldsymbol{U}}\boldsymbol{V}$ for two vectors fields, $\boldsymbol{V}$ and $\boldsymbol{U}$, against their commutator $[\boldsymbol{V},\boldsymbol{U}]$. According to definition of the covariant derivative and the commutator, both the difference $\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{U} - \boldsymbol{\nabla}_{\boldsymbol{U}}\boldsymbol{V}$ and the commutator are linear vector fields. Therefore, their linear combination must be a tensor called tensor of torsion or simply torsion . Torsion $\mathbb{T}$ is a linear function defined on two vector fields, $\boldsymbol{V}$ and $\boldsymbol{U}$,

$$\mathbb{T}(\boldsymbol{V},\boldsymbol{U}) \equiv \boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{U} - \boldsymbol{\nabla}_{\boldsymbol{U}}\boldsymbol{V} - [\boldsymbol{V},\boldsymbol{U}] \ , \tag{3.112}$$

which takes its values on covectors [Schutz, 1995]. Hence, torsion is an anti-symmetric tensor of type $(1,2)$: $\mathbb{T}(\boldsymbol{V},\boldsymbol{U}) = -\mathbb{T}(\boldsymbol{U},\boldsymbol{V})$. Components of the torsion are defined as the value of the tensor $\mathbb{T}$ on the basis vectors and covectors

$$\mathbb{T} = T^{\alpha}_{\beta\gamma}\boldsymbol{e}_{\alpha} \otimes \tilde{\boldsymbol{\omega}}^{\beta} \otimes \tilde{\boldsymbol{\omega}}^{\gamma} \ , \tag{3.113}$$

Making use of definitions (3.84), (3.110) and (3.112), one gets

$$T^{\alpha}_{\beta\gamma} = \mathfrak{G}^{\alpha}_{\gamma\beta} - \mathfrak{G}^{\alpha}_{\beta\gamma} - C^{\alpha}_{\beta\gamma} \ . \tag{3.114}$$

where $\mathfrak{G}^{\alpha}_{\gamma\beta}$ is the affine connection and $C^{\alpha}_{\beta\gamma}$ are the commutation coefficients of the basis.

Let us take a smooth scalar function f and account that the value of a basis vector field, let say, $\boldsymbol{e}_{\alpha}$ on this function is a partial derivative along this vector, $\boldsymbol{e}_{\alpha}[f] = \partial_{\alpha} f$ in accordance with equation (3.42). On the other hand, a partial derivative from a scalar function coincides with its covariant derivative

$$\partial_{\alpha} f = \boldsymbol{\nabla}_{\alpha} f \ . \tag{3.115}$$

Taking into account this fact, definition (3.112) of torsion can be written down as a commutator of two covariant derivatives acting on the scalar function,

$$\left[\boldsymbol{\nabla}_{\beta}, \boldsymbol{\nabla}_{\gamma}\right] f \equiv \left(\boldsymbol{\nabla}_{\beta}\boldsymbol{\nabla}_{\gamma} - \boldsymbol{\nabla}_{\gamma}\boldsymbol{\nabla}_{\beta}\right) f = \left(T^{\alpha}_{\beta\gamma} + C^{\alpha}_{\beta\gamma}\right) \boldsymbol{\nabla}_{\alpha} f \ . \tag{3.116}$$

This equation explicitly demonstrates that torsion is a linear differential operator. It tells us that the commutator of two covariant derivatives is a linear operator when

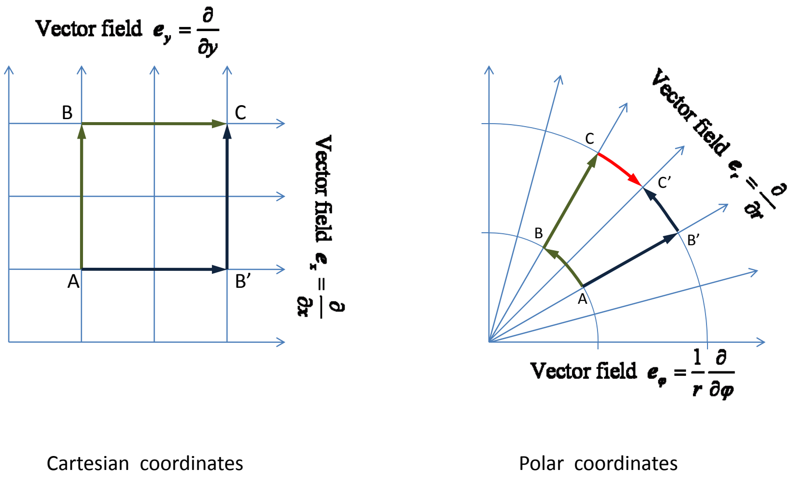

Figure 3.2 Vector fields $\boldsymbol{e}_x = \partial/\partial x$ and $\boldsymbol{e}_y = \partial/\partial y$ corresponding to Cartesian coordinates x and y, commute, $\left[\boldsymbol{e}_x, \boldsymbol{e}_y\right] = 0$, that is two paths ABC and AB'C makes a complete closure, and the two vector fields are holonomic. On the other hand, vector fields $\boldsymbol{e}_r = \partial/\partial r$ and $\boldsymbol{e}_\varphi = (1/r)(\partial/\partial\varphi)$ corresponding to polar coordinates r and φ, do not commute, $\left[\boldsymbol{e}_r, \boldsymbol{e}_\varphi\right] = -(1/r)\boldsymbol{e}_\varphi$, that is moving along path ABC and along AB'C' does not make a closure and the fields form a non-holonomic basis. The non-holonomy is evaluated by the commutator, $\left[\boldsymbol{e}_r, \boldsymbol{e}_\varphi\right]$, that is a vector directed from point C to point C' along the field $\boldsymbol{e}_\varphi$.

it acts on scalar functions [30]. Equation (3.116) also allows us to give a geometric interpretation of torsion. Specifically, torsion of a manifold is characterized by the absence of parallelograms, i.e. if a vector is parallel transported along a closed loop it will coincide with the original vector but if it is parallel transported along four legs of a parallelogram the torsion of the manifold will result in a displacement of the resulting vector from its original position as explained in the Figure 3.3.

The affine connection is called torsion-free if the tensor of torsion vanishes on the entire manifold: $T^\alpha_{\beta\gamma} \equiv 0$. The torsion-free connection has a non-vanishing antisymmetric part in any non-coordinate (non-holonomic) basis

$$\mathfrak{G}^\alpha_{[\beta\gamma]} = -\frac{1}{2}C^\alpha_{\beta\gamma} \,, \tag{3.117}$$

Because all commutation coefficients $C^\alpha_{\beta\gamma}$ vanish in any coordinate basis, the torsion-free connection is fully-symmetric in the coordinate basis: $\mathfrak{G}^\alpha_{\beta\gamma} = \mathfrak{G}^\alpha_{\gamma\beta}$. This symmetry reduces the number of independent connection coefficients at each point of the manifold $\mathcal{M}$ from $4^3 = 64$ to $64 - 24 = 40$.

Initially, torsion looked unimportant to Einstein who set it to zero, because it did not appear necessary to provide a gravitation theory with a consistent set of equations for the metric gravitational field. Einstein returned to the gravity theory with torsion later, and methodically explored its role in a unified field theory which he consistently worked on in pursuing the program of geometrization of physics [Sauer, 2010]. This Einstein-Cartan theory[Trautman, 2006] and its modern version of teleparallelism explore the ideas associated with the role the torsion may play in gravity [Hehl and Kerlick, 1978; Hehl and Obukhov, 2007; Hehl and Weinberg, 2007; Kleinert, 2008]. One is not going to thoroughly discuss the effects of torsion in this book, and assume in most cases that the connection is torsion-free [31].

3.5.3
Nonmetricity tensor

A tensor of nonmetricity $\mathbb{Q}$, is defined as a covariant derivative of the metric tensor [Kleinert, 2008]

$$\mathbb{Q} \equiv -\boldsymbol{\nabla}\mathbb{G} \,. \tag{3.118}$$

Nonmetricity is not zero in the most general metric-affine geometries [Hehl and Kerlick, 1978; Heinicke *et al.*, 2005], though it is equal to zero in general relativity and in the metric-based theories of gravity [Will, 1993, 2006]. The nonmetricity is a tensor of the type (1,2) which has the following components

$$\mathbb{Q} = Q^\alpha_{\beta\gamma}\boldsymbol{e}_\alpha \otimes \tilde{\omega}^\beta \otimes \tilde{\omega}^\gamma \,, \tag{3.119}$$

30) We shall demonstrate in section 3.7.1 that the commutator of two covariant derivatives remains a linear operator if the scalar function is replaced with a vector. It leads to the concept of curvature of a manifold.

31) Differentiable manifolds with non-vanishing torsion play an essential role in the theory of dislocations and disclinations in crystals as well as in the gauge theories of gravity that go beyond general relativity [Hehl and Obukhov, 2007; Kleinert, 1989].

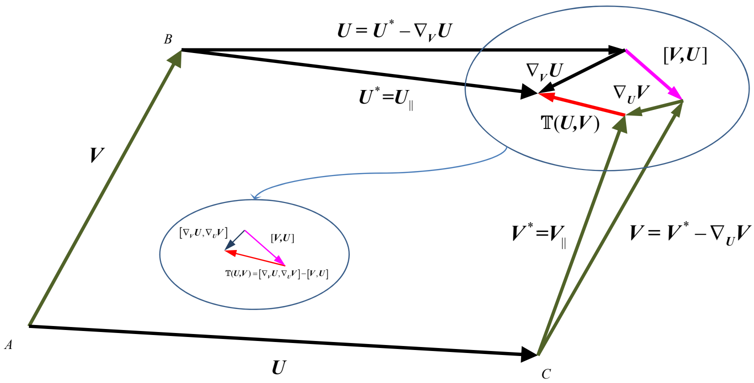

Figure 3.3 Two vector fields $\boldsymbol{U}$ and $\boldsymbol{V}$ are given. Tangent vector $\boldsymbol{U}$ is parallel-transported along vector $\boldsymbol{V}$ from point A to point B. It results in vector $\boldsymbol{U}^* = \boldsymbol{U}_{\parallel}$ at point B which compares with tangent vector of the vector field $\boldsymbol{U}$ at this point: $\boldsymbol{U} = \boldsymbol{U}^* - \boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{U}$. Tangent vector $\boldsymbol{V}$ is parallel-transported along vector $\boldsymbol{U}$ from point A to point C. It results in vector $\boldsymbol{V}^* = \boldsymbol{V}_{\parallel}$ at point C which compares with tangent vector of the vector field $\boldsymbol{V}$ at this point: $\boldsymbol{V} = \boldsymbol{V}^* - \boldsymbol{\nabla}_{\boldsymbol{U}}\boldsymbol{V}$. The tips of two vectors, $\boldsymbol{V}^*$ and $\boldsymbol{U}^*$, do not make a closure on the manifolds with torsion. The tips of two vectors, $\boldsymbol{U} = \boldsymbol{U}^* - \boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{U}$ and $\boldsymbol{V} = \boldsymbol{V}^* - \boldsymbol{\nabla}_{\boldsymbol{U}}\boldsymbol{V}$, do not make a closure if $\boldsymbol{U}$ and $\boldsymbol{V}$ do not commute. The gap resulting in the *closure failure of parallelogram*, is a vector consisting of an algebraic sum of the commutator $[\boldsymbol{V}, \boldsymbol{U}]$ of the two vectors and the torsion $\mathbb{T}(\boldsymbol{V}, \boldsymbol{U})$. Torsion is a geometric object being intrinsic to the differential structure of manifold, and it does not depend on vector fields. At the same time, the commutator of two vector fields depends exclusively on the nature of the fields, and vanishes if the fields commute.

where $Q^{\alpha}_{\beta\gamma} = g^{\alpha\mu} Q_{\mu\beta\gamma}$ is symmetric with respect to indices β and γ, $Q^{\alpha}_{\beta\gamma} = Q^{\alpha}_{\gamma\beta}$, and

$$Q_{\mu\beta\gamma} = -\boldsymbol{\nabla}_{\mu} g_{\beta\gamma} \ . \tag{3.120}$$

The nonmetricity tensor of a special type, $Q_{\alpha\beta\gamma} = A_{\alpha} g_{\beta\gamma}$, with A_{α} being a vector of electromagnetic field, was considered by Weyl in the attempt to unify gravity with electromagnetism. This attempt was criticized by Einstein on the ground that if the idea of Weyl's nonmetricity were correct the principle of equivalence would be violated. In such a case, the behavior of clocks would depend on their history, thus, making the atomic spectra unstable which contradicted the existing empirical evidence [O'Raifeartaigh and Straumann, 2000].

It is also possible to introduce the nonmetricity in the form of a linear combination of gradients of a scalar function f as, for example, it appears in the most typical example of, so-called, Palatini $f(R)$ gravity without torsion, where f is a function of the Ricci scalar R [Sotiriou and Liberati, 2007]. This kind of nonmetricity can be eliminated by converting it to an algebraic function of the Ricci tensor that is expressed in terms of the Levi-Civita connection only [Sotiriou, 2009]. In any case, the nonmetricity either complicates the gravity theory or introduces other geometric objects besides the metric tensor. It also leads to existence of a preferred-frame effects [Heinicke *et al.*, 2005; Kleinert, 2008]. Einstein excluded the nonmetricity from the general-relativistic theory of gravity. Thus, general relativity and other metric-based theories of gravity [Will, 1993] operate solely with the Levi-Civita connection, $\Gamma^{\alpha}_{\mu\nu}$, which is linked to the metric tensor, $g_{\alpha\beta}$, as explained in the next section.

3.5.4
Linking the connection with the metric structure

The affine connection exists on manifold $\mathcal{M}$ independently of any other geometric structures. It establishes the equivalence between a bundle of tangent spaces to the manifold. The equivalence means that a mutual relationship defined in a tangent space between any two of its elements must be preserved in the process of parallel transport of tensors from one tangent space to another. One particularly important relationship of this kind is a dot-product of two vectors, let say $\boldsymbol{V}$ and $\boldsymbol{U}$, that is defined with the help of a metric tensor $\mathbb{G}$. Components $g_{\beta\gamma}$ of the metric tensor on the manifold $\mathcal{M}$ are scalar functions defined by equation (2.126) as a numerical value of tensor $\mathbb{G}$ calculated on two basis vectors, $\boldsymbol{e}_{\beta}$ and $\boldsymbol{e}_{\gamma}$.

Any affine connection $\mathfrak{G}_{\alpha\beta\gamma} = g_{\alpha\mu}\mathfrak{G}^{\mu}_{\beta\gamma}$ is linearly decomposed in four, geometrically-different pieces [Heinicke *et al.*, 2005; Kleinert, 1989, 2008]

$$\mathfrak{G}_{\alpha\beta\gamma} = \Gamma_{\alpha\beta\gamma} + \mathfrak{K}_{\alpha\beta\gamma} + \mathfrak{D}_{\alpha\beta\gamma} + \mathfrak{N}_{\alpha\beta\gamma} \ , \tag{3.121}$$

where

$$\mathfrak{K}_{\alpha\beta\gamma} = \frac{1}{2}\left(T_{\gamma\alpha\beta} + T_{\beta\alpha\gamma} - T_{\alpha\beta\gamma}\right) \ , \tag{3.122}$$

are components of the contortion tensor depending on torsion $T_{\alpha\beta\gamma} = g_{\alpha\mu}T^{\mu}_{\beta\gamma} =$

$-g_{\alpha\mu}T^{\mu}_{\ \gamma\beta}$,

$$\mathfrak{D}_{\alpha\beta\gamma} = \frac{1}{2}\left(Q_{\gamma\alpha\beta} + Q_{\beta\alpha\gamma} - Q_{\alpha\beta\gamma}\right) , \tag{3.123}$$

are components of the deviation tensor depending on nonmetricity $Q_{\alpha\beta\gamma} = Q_{\alpha\gamma\beta}$,

$$\mathfrak{R}_{\alpha\beta\gamma} = \frac{1}{2}\left(C_{\gamma\alpha\beta} + C_{\beta\alpha\gamma} - C_{\alpha\beta\gamma}\right) , \tag{3.124}$$

are the Ricci rotation coefficients depending on the commutation coefficients $C_{\alpha\beta\gamma} = g_{\alpha\mu}C^{\mu}_{\ \beta\gamma} = -g_{\alpha\mu}C^{\mu}_{\ \gamma\beta}$, and $\Gamma_{\alpha\beta\gamma} = g_{\alpha\mu}\Gamma^{\mu}_{\ \beta\gamma}$ is the Levi-Civita connection that is not a tensor.

Compatibility of the affine connection $\mathfrak{G}_{\alpha\beta\gamma}$ with the dot-product of two vectors under the parallel transport demands the metric tensor to be covariantly constant along *any* curve on the manifold, that is the nonmetricity tensor must vanish at any point of the manifold. Thus, one imposes the *metricity condition*

$$\mathbb{Q} \equiv 0 . \tag{3.125}$$

We shall also demand that the connection is torsion-free assuming that $T_{\alpha\beta\gamma} = 0$ everywhere on the manifold. This requirement is not obligatory and the condition (3.125) does not depend on the value of torsion. However, torsion is not accepted in general relativity and in other metric-based theories of gravity, which is the reason for its elimination from further consideration.

Taking into account definition (3.119) of nonmetricity and equation (3.91) of the covariant derivative of tensors, the condition (3.125) can be re-formulated in terms of the components of the metric tensor, $g_{\alpha\beta}$, and the affine connection $\mathfrak{G}^{\alpha}_{\ \beta\gamma}$ as follows

$$\boldsymbol{\nabla}_{\alpha} g_{\beta\gamma} = \partial_{\alpha} g_{\beta\gamma} - g_{\mu\gamma}\mathfrak{G}^{\mu}_{\ \beta\alpha} - g_{\mu\beta}\mathfrak{G}^{\mu}_{\ \gamma\alpha} = 0 , \tag{3.126}$$

where $\partial_{\alpha} g_{\beta\gamma} \equiv \boldsymbol{e}_{\alpha}\left[g_{\beta\gamma}\right]$ denotes a partial derivative of the metric tensor components along the direction of a vector basis $\boldsymbol{e}_{\alpha}$, and $\mathfrak{G}^{\mu}_{\ \beta\alpha} = g^{\mu\nu}\mathfrak{G}_{\nu\beta\alpha}$. Equation (3.126) fixes the Levi-Civita connection coefficients and links it to the metric tensor and its partial derivatives.

Now, one substitutes the right side of expression (3.121) for the affine connection $\mathfrak{G}^{\alpha}_{\ \beta\gamma}$ into equation (3.126), make permutations of indices and form two other similar equations. Their linear combination allows us to express the Levi-Civita connection in the following form

$$\Gamma_{\alpha\beta\gamma} = \frac{1}{2}\left(\partial_{\gamma} g_{\alpha\beta} + \partial_{\beta} g_{\alpha\gamma} - \partial_{\alpha} g_{\beta\gamma}\right) + \mathfrak{R}_{\alpha\beta\gamma} , \tag{3.127}$$

where

$$\Gamma_{\alpha\beta\gamma} \equiv g_{\alpha\mu}\Gamma^{\mu}_{\ \beta\gamma} . \tag{3.128}$$

In the coordinate bases expression for the Levi-Civita connection is simplified because all commutation coefficients vanish ($C_{\alpha\beta\gamma} = 0$), and the derivatives along the basis vectors are reduced to regular partial derivatives along coordinate axes, $\partial_a = \partial/\partial x^\alpha$. The resulting form of the Levi-Civita connection referred to a coordinate basis is called the Christoffel symbols

$$\Gamma^\alpha_{\beta\gamma} = \frac{1}{2} g^{\alpha\lambda} \left(\frac{\partial g_{\lambda\beta}}{\partial x^\gamma} + \frac{\partial g_{\lambda\gamma}}{\partial x^\beta} - \frac{\partial g_{\beta\gamma}}{\partial x^\lambda} \right) , \tag{3.129}$$

which were introduced previously in a different way in equation (3.34). Present derivation however did not rely upon prior existence of any coordinate system and/or Minkowski metric.

Coordinate bases are not orthonormal in a general situation, in the sense that the coefficients $g_{\alpha\beta}$ of the metric tensor are functions of its position on manifold $\mathcal{M}$ and are not reduced to the Minkowski metric $\eta_{\alpha\beta}$ globally. On the other hand, one can always introduce an orthonormal tetrad basis on manifold $\mathcal{M}$ such that at each point of the manifold the coefficients of the metric tensor are reduced to the Minkowski metric $\eta_{\alpha\beta}$, and their first partial derivatives are zero, $\partial_a g_{\beta\gamma} = 0$. An example of a tetrad basis is delivered by the normal coordinates explained in section 3.3.8.3. In such basis the Christoffel symbols vanish, and the Levi-Civita connection is fully defined by the Ricci rotation coefficients

$$\Gamma_{\alpha\beta\gamma} = \Re_{\alpha\beta\gamma} = \frac{1}{2} \left(C_{\gamma\alpha\beta} + C_{\beta\alpha\gamma} - C_{\alpha\beta\gamma} \right) . \tag{3.130}$$

The tetrad bases are especially important in mathematical theory of black holes [Chandrasekhar, 1983] but are not a popular tool in relativistic celestial mechanics and astrometry that mostly rely upon non-orthogonal coordinate bases since they are more convenient for practical calculations and in data reduction algorithms.

3.6 Lie Derivative

3.6.1 A vector flow

Let us consider a manifold $\mathcal{M}$ and a smooth vector field $\boldsymbol{V}$ on it. Let x^α be local coordinates at a neighborhood of a point P. In this coordinate chart the vector $\boldsymbol{V}(x)$ is given by its components: $\boldsymbol{V}(x) = V^\alpha(x)\partial/\partial x^\alpha$. The flow, F^s_V, of the vector field is defined as solution of a system of ordinary differential equations

$$\frac{dy^\alpha}{ds} = V^\alpha(y) , \tag{3.131}$$

where $y = y(s)$, and s is a parameter along the direction of the vector field. The most famous example of a vector flow is the motion of fluid in hydrodynamics. Each fluid element has its own velocity and all together they form the velocity flow of the fluid. An ensemble of test particles moving along geodesics defined by equation (3.96), give another example of a vector flow in relativistic mechanics of Einstein's theory of gravity. Each vector flow defines a congruence of the integral curves of the corresponding vector field.

The curves $y^\alpha(s)$ are called the integral lines of the vector flow F^s_V which, thus, defines a one-parameter mapping (diffeomorphism) of a local subset of a manifold $\mathcal{M}$ to itself: $x^\alpha \mapsto y^\alpha = F^s_V x^\alpha$ (see Figure 3.4). In local coordinates each integral line of the vector flow is an analytic function and it can be expanded into Taylor series around the value $s = 0$, which yields $y^\alpha(s) = y^\alpha(0) + s\; dy^\alpha(0)/ds + O(s^2)$. Taking into account that for $s = 0$ the values of integral lines constitute a set of points coinciding with the original local coordinates, $y^\alpha(0) = x^\alpha$, and that according to equation (3.131) the derivative $dy^\alpha(0)/ds = V^\alpha(y(0)) = V^\alpha(x)$, one can write down the diffeomorphism F^s_V in the following form

$$y^\alpha = x^\alpha + s\,V^\alpha(x) , \tag{3.132}$$

where one has discarded all terms of the higher order in s. This equation can be interpreted as a parametric transformation between the local coordinates x^α and y^α. Then, the Jacobian matrix of transformation is

$$\Lambda^\alpha{}_\beta(s) = \frac{\partial y^\alpha}{\partial x^b} = \delta^\alpha_\beta + s\,\frac{\partial V^\alpha}{\partial x^\beta} , \tag{3.133}$$

and the matrix of the inverse transformation

$$\tilde{\Lambda}^\beta{}_\alpha(s) = \frac{\partial x^\beta}{\partial y^\alpha} = \delta^\beta_\alpha - s\,\frac{\partial V^\beta}{\partial x^\alpha} . \tag{3.134}$$

In the linear approximation with respect to the parameter s, the direct and inverse Jacobian matrices are related as follows

$$\tilde{\Lambda}^\beta{}_\alpha(s) = \Lambda^\beta{}_\alpha(-s) , \tag{3.135}$$

that is sufficient for most of the applications.

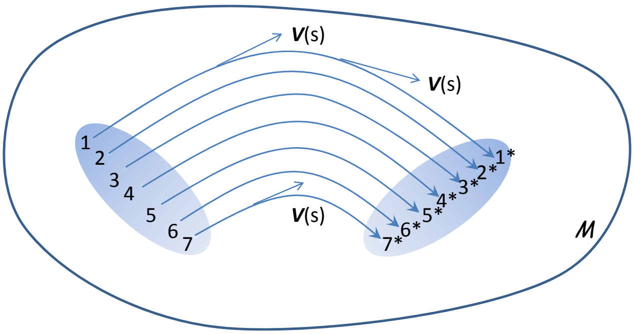

Figure 3.4 Local diffeomorphism of a manifold $\mathcal{M}$ to itself by means of a vector flow defined by the congruence of integral curves $y^\alpha(s)$ of a vector field $\boldsymbol{V}(s)$. Only a few lines of the vector flow are shown. Parametric distance along the curves $1 \to 1^*$, $2 \to 2^*$, etc., is the same in correspondence with the definition of the vector flow.

3.6.2
The directional derivative of a function

Let us consider a differentiable function $f = f(x)$ on the manifold $\mathcal{M}$. We define a derivative $\pounds_{\boldsymbol{V}}$ of the function f in the direction of the vector field $\boldsymbol{V}$ by the following rule [32)]

$$\pounds_{\boldsymbol{V}} f(x^\alpha) \equiv \lim_{s\to 0}\left[\frac{df\left(F^s_V x^\alpha\right)}{ds}\right] , \tag{3.136}$$

where $F^s_V x^\alpha$ is the flow of the vector field $\boldsymbol{V}$. Performing mathematical operations in equation (3.136), one concludes that in the local coordinates x^α the directional derivative is

$$\pounds_{\boldsymbol{V}} f = V^\alpha \frac{\partial f}{\partial x^\alpha} . \tag{3.137}$$

It coincides with the definition of the value of the vector field $\boldsymbol{V}$ on function f

$$\boldsymbol{V}[f] \equiv V^\alpha \frac{\partial f}{\partial x^\alpha} , \tag{3.138}$$

which (due to the duality of vectors and covectors) is also a value of the gradient $\tilde{\boldsymbol{d}} f$ of function f on the vector field $\boldsymbol{V}$, that is $\boldsymbol{V}[f] = \tilde{\boldsymbol{d}} f(\boldsymbol{V})$ according to definition (3.54). Hence, one has

$$\pounds_{\boldsymbol{V}} f = \tilde{\boldsymbol{d}} f(\boldsymbol{V}) , \tag{3.139}$$

that establishes equivalence between the directional derivative of a function and its gradient.

3.6.3
Geometric interpretation of the commutator of two vector fields

Suppose that one is given two vector fields, $\boldsymbol{V}$ and $\boldsymbol{U}$ on a manifold $\mathcal{M}$ which corresponds to two vector flows, F^s_V and F^t_U, where s and t are parameters of the corresponding flows. Let us evaluate the degree of non-commutativity of the two flows. To this end, one considers two images of the point P obtained by mapping its coordinates $x^\alpha = x^\alpha_{\rm P}$ to points $x^\alpha_{\rm B} = F^s_V F^t_U x^\alpha$ and $x^\alpha_D = F^t_U F^s_V x^\alpha$ along the two flows, and compute the difference between them by comparing the values of some analytic function $f(x)$ at these two points (see Figure 3.5)

$$\Delta(x^\alpha_{\rm B}, x^\alpha_{\rm D}) = f\left(x^\alpha_{\rm D}\right) - f\left(x^\alpha_{\rm B}\right) . \tag{3.140}$$

One can expand function $f(x)$ in the Taylor series around the point P, where $t = s = 0$. The first terms of the expansion of $f(x^\alpha_{\rm B})$ and $f(x^\alpha_{\rm D})$ will cancel each other in the difference $\Delta(x^\alpha_{\rm B}, x^\alpha_{\rm D})$. The same cancelation will take place for all other terms being

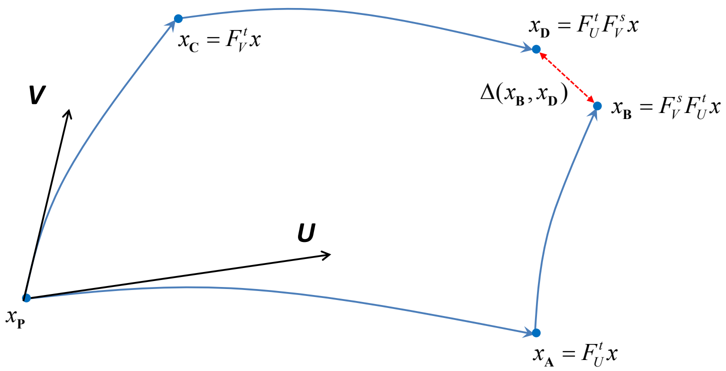

Figure 3.5 Two integral lines of the vector fields, $\boldsymbol{V}$ and $\boldsymbol{U}$, mapping a point P with coordinates $x^\alpha = x^\alpha_{\mathrm{P}}$ to two other points, $x^\alpha_{\mathrm{B}} = F^s_V F^t_U x^\alpha$ and $x^\alpha_D = F^t_U F^s_V x^\alpha$, are shown. The two vector fields are commutative (that is $[\boldsymbol{U}, \boldsymbol{V}] = 0$), if and only if, the difference $\Delta(x^\alpha_{\mathrm{B}}, x^\alpha_{\mathrm{D}})$ between the points is equal to zero.

proportional to the powers of either t or s. Therefore, the first non-vanishing term in the Taylor expansion of $\Delta(x^\alpha_\mathrm{B}, x^\alpha_\mathrm{D})$ will be proportional to the product of st.

We shall calculate this principal term that is defined as a mixed, second-order partial derivative [Arnold, 1995, page 210]

$$\Delta(x^\alpha_\mathrm{B}, x^\alpha_\mathrm{D}) = \lim_{\substack{t\to 0\\ s\to 0}} \left[\frac{\partial^2 f\left(F^t_U F^s_V x^\alpha\right)}{\partial s \partial t} - \frac{\partial^2 f\left(F^s_V F^t_U x^\alpha\right)}{\partial t \partial s}\right] . \tag{3.141}$$

By definition (3.136) of the directional derivative, one has

$$\lim_{s\to 0}\left[\frac{\partial f\left(F^s_V F^t_U x^\alpha\right)}{\partial s}\right] = £_{\boldsymbol{V}} f\left(F^t_U x^\alpha\right) . \tag{3.142}$$

If one denotes the function $£_{\boldsymbol{V}} f$ by ψ, then for the same reasons, one obtains

$$\lim_{t\to 0}\left[\frac{\partial \psi\left(F^t_U x^\alpha\right)}{\partial t}\right] = £_{\boldsymbol{U}} \psi\left(x^\alpha\right) . \tag{3.143}$$

Hence,

$$\lim_{\substack{t\to 0\\ s\to 0}}\left[\frac{\partial^2 f\left(F^s_V F^t_U x^\alpha\right)}{\partial t \partial s}\right] = £_{\boldsymbol{U}} £_{\boldsymbol{V}} f(x^\alpha) . \tag{3.144}$$

Proceeding in a similar way, one arrives to the result

$$\Delta(x^\alpha_\mathrm{B}, x^\alpha_\mathrm{D}) = \left[£_{\boldsymbol{V}} £_{\boldsymbol{U}} - £_{\boldsymbol{U}} £_{\boldsymbol{V}}\right] f(x^\alpha) . \tag{3.145}$$

If one now uses the local coordinates x^α to compute $\Delta(x^\alpha_\mathrm{B}, x^\alpha_\mathrm{D})$, one obtains

$$\Delta(x^\alpha_\mathrm{B}, x^\alpha_\mathrm{D}) = \left(V^\beta \frac{\partial U^\alpha}{\partial x^\beta} - U^\beta \frac{\partial V^\alpha}{\partial x^\beta}\right) \frac{\partial f}{\partial x^\alpha} , \tag{3.146}$$

which means that the operator $£_{\boldsymbol{V}} £_{\boldsymbol{U}} - £_{\boldsymbol{U}} £_{\boldsymbol{V}}$ is a linear differential operator corresponding to some vector field $\boldsymbol{W}$. By comparing equation (3.146) with equation (3.109), one immediately arrives to the conclusion that the field $\boldsymbol{W} = [\boldsymbol{V}, \boldsymbol{U}]$ is a commutator of the two vector fields, $\boldsymbol{V}$ and $\boldsymbol{U}$, and

$$£_{\boldsymbol{V}} £_{\boldsymbol{U}} - £_{\boldsymbol{U}} £_{\boldsymbol{V}} = £_{[\boldsymbol{V},\boldsymbol{U}]} = [\boldsymbol{V}, \boldsymbol{U}] . \tag{3.147}$$

Finally, one has

$$\Delta(x^\alpha_\mathrm{B}, x^\alpha_\mathrm{D}) = [\boldsymbol{V}, \boldsymbol{U}] f(x^\alpha) . \tag{3.148}$$

This equation allows us to formulate an important theorem:

32) Derivative $£_{\boldsymbol{V}}$ is also called the directional derivative.

the two flows F^s_V and F^t_U commute, if and only if, the commutator of the corresponding vectors fields $[\boldsymbol{V},\boldsymbol{U}]$ is equal to zero.

Indeed, if the two points, $x^\alpha_{\rm B} = F^s_V F^t_U x^a$ and $x^\alpha_{\rm D} = F^s_U F^t_V x^a$, coincide, the commutator $[\boldsymbol{V},\boldsymbol{U}] = 0$, making the difference $\Delta(x^\alpha_{\rm B}, x^\alpha_{\rm D}) = 0$, due to equation (3.148). On the other hand, if $[\boldsymbol{V},\boldsymbol{U}] = 0$, one has $\Delta(x^\alpha_{\rm B}, x^\alpha_{\rm D}) = 0$ for any analytic function $f(x)$ in a neighborhood of the point P. It implies, $f\left(x^\alpha_{\rm B}\right) = f\left(x^\alpha_{\rm D}\right)$. Since, the function $f\left(x^\alpha_{\rm B}\right)$ is analytic, its Taylor expansion around point D is convergent and leads to equation

$$\frac{\partial f(x_{\rm D})}{\partial x^\alpha}\left(x^\alpha_{\rm B} - x^\alpha_{\rm D}\right) + \frac{1}{2!}\frac{\partial^2 f(x_{\rm D})}{\partial x^\alpha \partial x^\beta}\left(x^\alpha_{\rm B} - x^\alpha_{\rm D}\right)\left(x^\beta_{\rm B} - x^\beta_{\rm D}\right)^2 + \ldots = 0\,, \tag{3.149}$$

where the ellipses denote terms of higher order in the infinite Taylor series. Because the series is convergent, and the coefficients of the series are not all equal to zero, one concludes that the points $x^\alpha_{\rm B}$ and $x^\alpha_{\rm D}$ must coincide.

An important example of commuting vectors field is delivered by the grid of any coordinate chart. Indeed, according to definition, each coordinate line keeps the value of other coordinates constant in the domain of its applicability. If one moves from point P along one coordinate line at a parametric distance s, and then along another coordinate line at a parametric distance t, one arrives to the same point on manifold $\mathcal{M}$ as if one moved, first, along the second line at the parametric distance t, and then along the first line at the parametric distance s. This explains why coordinates are so convenient in practical applications to celestial mechanics, astrometry and geodesy: each point of spacetime manifold $\mathcal{M}$ is characterized by a unique set of four parameters. Precise navigation is also inconceivable without coordinates.

3.6.4 Definition of the Lie derivative

Let us consider a tensor field $\mathbb{S}$ and a vector field $\boldsymbol{V}$ that generates a one-parametric vector flow F^s_V. The Lie derivative of the tensor field is a generalization of the concept of the directional derivative of a scalar function as explained in equation (3.136). Lie derivative maps tensor $\mathbb{S}$ to a tensor $\pounds_{\boldsymbol{V}}\mathbb{S}$ of the same type, and is defined by the following equation [33)]

$$\pounds_{\boldsymbol{V}}\mathbb{S}(x^\alpha) \equiv \lim_{s\to 0}\left[\frac{d}{ds}F^s_V\mathbb{S}\left(F^s_V x^\alpha\right)\right]\,. \tag{3.150}$$

In order to obtain components of a Lie derivative, one introduces a local coordinate basis $\partial/\partial x^\alpha$, where the tensor $\mathbb{S}$ of type (k,l) has components $S^{\alpha_1\alpha_2\ldots\alpha_k}_{\beta_1\beta_2\ldots\beta_l}(x^\alpha)$, the vector field $\boldsymbol{V}$ has components $V^\alpha(x^\alpha)$, and the matrix of transformation corresponding to the diffeomorphism is given by equations (3.133), (3.134). Projection of equation

33) The requirement that the object $\mathbb{S}$ is a tensor can be relaxed. In fact, definition (3.150) is applied for arbitrary geometric object with a known law of coordinate transformation of its components. We emphasize that though the object may be not a tensor, its Lie derivative will be a tensor. This remarkable fact can be checked by direct calculation [Mitskevich, 1969; Popova and Petrov, 1988].

(3.150) on the basis tensor composed of the tensor product of the basis vectors and covectors (see section 2.4.4.1), yields

$$\pounds_{\boldsymbol{V}} S^{\alpha_1\alpha_2\ldots\alpha_k}_{\beta_1\beta_2\ldots\beta_l}(x^\alpha) = \lim_{s\to 0}\left\{\frac{d}{ds}\left[F^s_V\mathbb{S}\right]^{\alpha_1\alpha_2\ldots\alpha_k}_{\beta_1\beta_2\ldots\beta_l}\left(F^s_V x^\alpha\right)\right\}\,, \tag{3.151}$$

where the components of tensor $F^s_V\mathbb{S}$ are obtained by transformation of tensor $\mathbb{S}$ from the point with coordinates $F^s_V x^\alpha$ to the original point [34] with coordinates x^α. More specifically,

$$\begin{aligned}\left[F^s_V\mathbb{S}\right]^{\alpha_1\alpha_2\ldots\alpha_k}_{\beta_1\beta_2\ldots\beta_l}\left(F^s_V x^\alpha\right) &\equiv \Lambda^{\alpha_1}{}_{\mu_1}(s)\Lambda^{\alpha_2}{}_{\mu_2}(s)\ldots\Lambda^{\alpha_k}{}_{\mu_k}(s)\\ &\times \Lambda^{\nu_1}{}_{\beta_1}(-s)\Lambda^{\nu_2}{}_{\beta_2}(-s)\ldots\Lambda^{\nu_l}{}_{\beta_l}(-s)\\ &\times S^{\mu_1\mu_2\ldots\mu_k}_{\nu_1\nu_2\ldots\nu_l}\left(x^\alpha + s\,V^\alpha(x)\right) + O(s^2)\,.\end{aligned} \tag{3.152}$$

Substituting equation (3.152) to definition (3.151) and performing corresponding operations, one obtains for Lie derivative of a tensor [35]

$$\begin{aligned}\pounds_{\boldsymbol{V}} S^{\alpha_1\alpha_2\ldots\alpha_k}_{\beta_1\beta_2\ldots\beta_l} &= V^\mu\partial_\mu S^{\alpha_1\alpha_2\ldots\alpha_k}_{\beta_1\beta_2\ldots\beta_l}\\ &+ \partial_{\beta_1}V^\mu S^{\alpha_1\alpha_2\ldots\alpha_k}_{\mu\beta_2\ldots\beta_l} + \ldots + \partial_{\beta_l}V^\mu S^{\alpha_1\alpha_2\ldots\alpha_k}_{\beta_1\beta_2\ldots\mu}\\ &- \partial_\mu V^{\alpha_1} S^{\mu\alpha_2\ldots\alpha_k}_{\beta_1\beta_2\ldots\beta_l} - \ldots - \partial_\mu V^{\alpha_k} S^{\alpha_1\alpha_2\ldots\mu}_{\beta_1\beta_2\ldots\beta_l}\,.\end{aligned} \tag{3.153}$$

In particular, this equation reveals that the Lie derivative of a vector field $\boldsymbol{U} = U^\alpha\partial/\partial x^\alpha$ is a commutator of the two vector fields

$$\pounds_{\boldsymbol{V}}\boldsymbol{U} = [\boldsymbol{V},\boldsymbol{U}] = -\pounds_{\boldsymbol{U}}\boldsymbol{V}\,. \tag{3.154}$$

This equation points out that the operator of the Lie derivative is not a linear operator with respect to the vector field $\boldsymbol{V}$ which defines direction of the derivative. This statement can be easily verified if one re-scales $\boldsymbol{V}$ to $f\boldsymbol{V}$, where f is a scalar function, and takes the Lie derivative along $f\boldsymbol{V}$,

$$\pounds_{f\boldsymbol{V}}\boldsymbol{U} = [f\boldsymbol{V},\boldsymbol{U}] = f\pounds_{\boldsymbol{V}}\boldsymbol{U} - \boldsymbol{V}\pounds_{\boldsymbol{U}}f\,. \tag{3.155}$$

One can see that the right side of this equation includes one more term depending on the Lie derivative $\pounds_{\boldsymbol{U}}f$ of the scalar function f. Only in case when f =const., the Lie derivative is a linear operator with respect to the direction of differentiation.

It should be clearly understood that the Lie derivative is conceptually different from the covariant derivative. Indeed, the operation of covariant differentiation along a vector demands a value of the vector at the point of differentiation [36] while taking the Lie derivative requires knowledge of the entire vector field (the vector flow) in the neighborhood of this point. This fact is reflected in equation (3.153) that demonstrates that the Lie derivative of a tensor along a vector field $\boldsymbol{V}$ requires taking partial

34) This operation is also called *pull-back*.
35) Lie derivative from a non-tensorial geometric object may include higher-order partial derivatives from vector filed V^α that defines the vector flow [Mitskevich, 1969].
36) Covariant derivative is a linear operator with respect to its "subscript" argument as explained in section 3.4.1.

derivatives of the vector field and does not involve any operation with the connection which is not a part of the definition of Lie derivative. Hence, Lie derivative can be defined on a manifold even if it has not been endowed with an affine connection.

Nevertheless, if manifold $\mathcal{M}$ has a torsion-free (Levi-Civita) connection it can be introduced to the definition of Lie derivative (3.153). Indeed, by making use of definition of the covariant derivative of tensor (3.91) and taking into account that the torsion-free connection in a coordinate basis is reduced to a (locally) symmetric Christoffel symbols, $\Gamma^{\alpha}_{\beta\gamma} = \Gamma^{\alpha}_{\gamma\beta}$, one can prove that the Lie derivative of a tensor of a (p,q) type, is given by equation

$$
\begin{aligned}
\pounds_{\boldsymbol{V}} S^{\alpha_1\alpha_2\ldots\alpha_p}_{\beta_1\beta_2\ldots\beta_q} &= V^{\mu}\nabla_{\mu} S^{\alpha_1\alpha_2\ldots\alpha_p}_{\beta_1\beta_2\ldots\beta_q} \\
&+ \nabla_{\beta_1} V^{\mu} S^{\alpha_1\alpha_2\ldots\alpha_p}_{\mu\beta_2\ldots\beta_q} + \ldots + \nabla_{\beta_q} V^{\mu} S^{\alpha_1\alpha_2\ldots\alpha_p}_{\beta_l\ldots\beta_{q-1}\mu} \\
&- \nabla_{\mu} V^{\alpha_1} S^{\mu\alpha_2\ldots\alpha_p}_{\beta_1\beta_2\ldots\beta_q} - \ldots - \nabla_{\mu} V^{\alpha_p} S^{\alpha_1\ldots\alpha_{p-1}\mu}_{\beta_1\beta_2\ldots\beta_q} \,,
\end{aligned}
\tag{3.156}
$$

which is obtained from equation (3.153) by formal replacing all partial derivatives with their covariant counterparts. Validity of equation (3.156) can be checked by inspection. We emphasize however that equation (3.156) is valid, if and only if, the connection is metrical. In case of a generic affine connection the Lie derivative must contain additional terms proportional to torsion and nonmetricity.

3.6.5
Lie transport of tensors

One has already seen in section 3.4.4 that tensor can be parallel-transported from one point of a manifold $\mathcal{M}$ to another along a smooth curve by making use of an affine connection $\Gamma^{\alpha}_{\beta\gamma}$ that defines a covariant derivative. There is however another way for transportation tensors on the manifold based on a linear diffeomorphism (vector flow) between tangent spaces.

We define Lie transport of a tensor $\mathbb{T}$ along a *congruence* of integral curves of a vector field $\boldsymbol{V}$ by a condition that its Lie derivative is zero

$$
\pounds_{\boldsymbol{V}} \mathbb{T} = 0 \,. \tag{3.157}
$$

Let each integral curve from the congruence be parameterized by parameter τ. We consider motion along the congruence from the parametric value τ_0 to $\tau_0 + \tau$ that defines a vector flow F^{τ}_{V} as described in section 3.6.1. Let manifold $\mathcal{M}$ has a tensor field $\mathbb{S}$. It has values $\mathbb{S}(\tau_0)$ and $\mathbb{S}(\tau_0 + \tau)$ at each point of the congruence with the values of parameter τ_0 and $\tau_0 + \tau$ respectively. Lie transport of the tensor field generates a new tensor field $\hat{\mathbb{S}}$ along the congruence that is a solution of equation $\pounds_{\boldsymbol{V}} \hat{\mathbb{S}} = 0$. The value of the new tensor field at the points of the congruence with the parameter's value $\tau_0 + \tau$, is $\hat{\mathbb{S}}(\tau_0 + \tau) = \hat{\mathbb{S}}(\tau_0) = \mathbb{S}(\tau_0)$ by the definition of the Lie transport and the initial condition. In general case, $\hat{\mathbb{S}}(\tau_0 + \tau) \neq \mathbb{S}(\tau_0 + \tau)$. In some cases, however, the Lie-transported tensor field $\hat{\mathbb{S}}$ will coincide with the value of the tensor field $\mathbb{S}$ at each point of the congruence. Such tensor field is called Lie covariant with respect to a vector field $\boldsymbol{V}$. It preserves its value along each integral line of $\boldsymbol{V}$: $\mathbb{S}(\tau) = \mathbb{S}(\tau_0 + \tau)$. Figure 3.6 visualizes the idea of the Lie transport and Lie invariance in case of a vector field $\boldsymbol{S}$ that is Lie-transported along a congruence defined by the integral curves of a vector field $\boldsymbol{V}(\tau)$.

If a tensor field $\mathbb{S}$ has a physical importance and, in addition, is Lie invariant with respect to a vector field $\boldsymbol{V}$, the field $\boldsymbol{V}$ is also physically meaningful. For example, a vector field $\boldsymbol{V} = V^\alpha \boldsymbol{e}_\alpha$ that makes a metric tensor $\mathbb{G}$ Lie invariant on spacetime manifold

$$\pounds_{\boldsymbol{V}} \mathbb{G} = 0 \,, \tag{3.158}$$

or in components

$$\boldsymbol{\nabla}_\alpha V_\beta + \boldsymbol{\nabla}_\beta V_a = 0 \,, \tag{3.159}$$

is known as a *Killing* vector indicating the existence of a symmetry of the spacetime (that means the metric tensor does not change along the Killing vector) with respect to a certain class of transformations. The transformations generated by Killing vectors are associated with conservation laws in dynamics of test particles [Misner *et al.*, 1973; Wald, 1984]. Timelike Killing vector yields conservation of energy; spacelike ones give conservation of momentum in the direction of the Killing vector, etc.

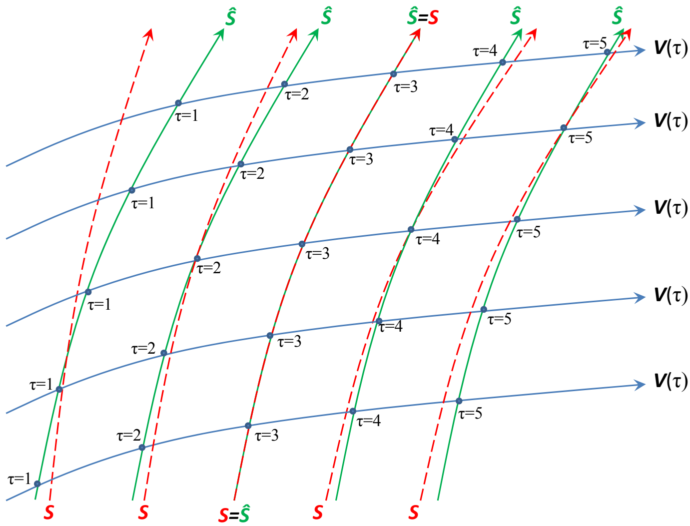

Figure 3.6 A vector S is Lie transported along a congruence of a vector field V parameterized with τ. The Lie transport generates a new vector field $\hat{S}$ along the integral curves of the vector field V. Since $£_V \hat{S} = [V, \hat{S}] = 0$, each integral line of the vector field $\hat{S}$ passes through the same value of the parameter τ of the congruence. In the most general case, the integral lines of the vector field S and the Lie-transported vector field $\hat{S}$ coincide only for one value of the parameter τ (in figure $\tau = 3$) that is the initial condition for the solution of the equation of the Lie transport. If vector fields S and $\hat{S}$ coincide for all values of the parameter along the congruence, the vector field S is called Lie invariant.

3.7
The Riemann Tensor and Curvature of Manifold

In previous section one has introduced the concept of manifold $\mathcal{M}$ equipped with an affine connection $\mathfrak{G}^{\alpha}_{\beta\gamma}$ that defines the covariant derivative and the rule of parallel transport of vectors and tensors from one point of the manifold to another. The connection is a geometric structure that global behavior is characterizes by two invariants - torsion and curvature. The torsion is determined by the anti-symmetric part of the affine connection but one has postulated that it vanishes identically. We shall also assume the nonmetricity of the connection vanishes. Hence, the affine connection is reduced to the Levi-Civita connection. Its geometric property - the curvature of manifold - is measuring the extent to which a parallel transport of a vector around closed loops fails to preserve its identity. It is the curvature which makes the parallel transport on a manifold path-dependent and two covariant derivatives noncommuting. Curvature of a manifold is quantified by the Riemann tensor which plays a central role both in the differential geometry and in any viable theory of gravity.

3.7.1
Noncommutation of covariant derivatives

Let us consider two vector fields, $\boldsymbol{V}$ and $\boldsymbol{U}$, on a differentiable manifold $\mathcal{M}$ endowed with an affine connection associated with a covariant derivative $\boldsymbol{\nabla}$. Existence of a metric tensor on the manifold is not required as it is unimportant for mathematics of this section. We define a commutator operator of the covariant derivatives

$$\left[\boldsymbol{\nabla}_{\boldsymbol{V}},\boldsymbol{\nabla}_{\boldsymbol{U}}\right]=\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{\nabla}_{\boldsymbol{U}}-\boldsymbol{\nabla}_{\boldsymbol{U}}\boldsymbol{\nabla}_{\boldsymbol{V}}\;, \tag{3.160}$$

and introduce an operator of a covariant derivative $\boldsymbol{\nabla}_{[\boldsymbol{V},\boldsymbol{U}]}$ along the commutator of the two vector fields. The operator (3.160) looks like a second-order covariant derivative which does not form a tensor. This observation is only partially correct - the operator (3.160) is not a tensor but form a first order covariant derivative. It is remarkable that a small modification allows one to build an actual tensor out of the commutator after subtracting the covariant derivative along the field $[\boldsymbol{V},\boldsymbol{U}]$

$$\mathbb{R}(\boldsymbol{V},\boldsymbol{U})\equiv\left[\boldsymbol{\nabla}_{\boldsymbol{V}},\boldsymbol{\nabla}_{\boldsymbol{U}}\right]-\boldsymbol{\nabla}_{[\boldsymbol{V},\boldsymbol{U}]}\;. \tag{3.161}$$

This geometric object is called the *curvature operator*. Let us prove that it belongs to the tangent space of the manifold $\mathcal{M}$ that is its calculation does not involve differentiation of the vector fields $\boldsymbol{V}$ and $\boldsymbol{U}$ and/or a geometric object upon which it acts.

First of all, one notices that the curvature operator is antisymmetric

$$\mathbb{R}(\boldsymbol{V},\boldsymbol{U})=-\mathbb{R}(\boldsymbol{U},\boldsymbol{V})\;. \tag{3.162}$$

so that the two vector arguments are interchangeable.

At second step, one proves that the curvature operator is linear with respect to its two arguments in the sense that for any differentiable function f

$$\mathbb{R}(f\boldsymbol{V},\boldsymbol{U}) = f\mathbb{R}(\boldsymbol{V},\boldsymbol{U}) \,. \tag{3.163}$$

Indeed, one has

$$\begin{aligned}[\boldsymbol{\nabla}_{f\boldsymbol{V}},\boldsymbol{\nabla}_{\boldsymbol{U}}] &= \boldsymbol{\nabla}_{f\boldsymbol{V}}\boldsymbol{\nabla}_{\boldsymbol{U}} - \boldsymbol{\nabla}_{\boldsymbol{U}}\boldsymbol{\nabla}_{f\boldsymbol{V}} \\ &= f\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{\nabla}_{\boldsymbol{U}} - \boldsymbol{\nabla}_{\boldsymbol{U}}\left(f\boldsymbol{\nabla}_{\boldsymbol{V}}\right) \\ &= f[\boldsymbol{\nabla}_{\boldsymbol{V}},\boldsymbol{\nabla}_{\boldsymbol{U}}] - \boldsymbol{U}[f]\boldsymbol{\nabla}_{\boldsymbol{V}} \,,\end{aligned} \tag{3.164}$$

where one has applied axioms of the affine connection from section 3.4 and the fact that the directional derivative $\boldsymbol{U}[f]$ is a scalar function. On the other hand,

$$\boldsymbol{\nabla}_{[f\boldsymbol{V},\boldsymbol{U}]} = \boldsymbol{\nabla}_{f\boldsymbol{V}\boldsymbol{U}-\boldsymbol{U}(f\boldsymbol{V})} = f\boldsymbol{\nabla}_{[\boldsymbol{V},\boldsymbol{U}]} - \boldsymbol{U}[f]\boldsymbol{\nabla}_{\boldsymbol{V}} \,, \tag{3.165}$$

where one has again used the axioms of the covariant derivatives. After adding up equations (3.164) and (3.165) one arrives to equation (3.163), *q.e.d.*

The last step, is to prove that the curvature operator is linear when it acts on a product of any differentiable function f and a vector field $\boldsymbol{W}$

$$\mathbb{R}(\boldsymbol{V},\boldsymbol{U})(f\boldsymbol{W}) = f\mathbb{R}(\boldsymbol{V},\boldsymbol{U})\boldsymbol{W} \,. \tag{3.166}$$

After performing the covariant differentiation, one has

$$[\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{\nabla}_{\boldsymbol{U}}](f\boldsymbol{W}) = f[\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{\nabla}_{\boldsymbol{U}}]\boldsymbol{W} + \boldsymbol{W}\boldsymbol{\nabla}_{[\boldsymbol{V},\boldsymbol{U}]}f \,, \tag{3.167}$$

$$\boldsymbol{\nabla}_{[\boldsymbol{V},\boldsymbol{U}]}(f\boldsymbol{W}) = f\boldsymbol{\nabla}_{[\boldsymbol{V},\boldsymbol{U}]}\boldsymbol{W} + \boldsymbol{W}\boldsymbol{\nabla}_{[\boldsymbol{V},\boldsymbol{U}]}f \,. \tag{3.168}$$

Subtracting equation (3.168) from (3.167) yields equation (3.166), *q.e.d.*

A tensor corresponding to the curvature operator is called the Riemann curvature tensor . Its components in an arbitrary basis are defined by equation

$$\mathbb{R}(\boldsymbol{e}_\gamma,\boldsymbol{e}_\delta)\boldsymbol{e}_\beta = R^\alpha{}_{\beta\gamma\delta}\boldsymbol{e}_\alpha \,, \tag{3.169}$$

where one has used the same sign convention as in the textbook by Misner *et al.* [1973]. In case of a Levi-Civita connection and in a coordinate basis the components of the Riemann tensor are

$$R^\alpha{}_{\beta\gamma\delta} = \frac{\partial\Gamma^\alpha_{\beta\delta}}{\partial x^\gamma} - \frac{\partial\Gamma^\alpha_{\beta\gamma}}{\partial x^\delta} + \Gamma^\alpha_{\mu\gamma}\Gamma^\mu_{\beta\delta} - \Gamma^\alpha_{\mu\delta}\Gamma^\mu_{\beta\gamma} \,, \tag{3.170}$$

where $\Gamma^\alpha_{\beta\gamma} = \Gamma^\alpha_{\gamma\beta}$ are the Christoffel symbols.

The above consideration reveals that if the Riemann tensor of a manifold is not equal to zero, the second-order covariant derivatives on a curved manifold do not commute, $\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{\nabla}_{\boldsymbol{U}} \neq \boldsymbol{\nabla}_{\boldsymbol{U}}\boldsymbol{\nabla}_{\boldsymbol{V}}$, as contrasted to the standard second-order partial derivatives which always commute, $\partial_\alpha\partial_\beta = \partial_\beta\partial_\alpha$. We emphasize however that the precise definition of the curvature operator entails a covariant derivative along the commutator of two vector fields as shown in equation (3.161). It means that two covariant derivatives from an arbitrary tensor taken along two non-commuting (non-holonomic) vector fields $\boldsymbol{V}$ and $\boldsymbol{U}$ do not commute, even if the Riemann tensor $\mathbb{R} = 0$. This happens because the commutator of the vector fields $[\boldsymbol{V}, \boldsymbol{U}] \neq 0$ [37].

3.7.2 The dependence of the parallel transport on the path

If the Riemann tensor of a manifold is different from zero, the result of a parallel transport of a vector will depend on the path of transportation. Indeed, let a vector $\boldsymbol{W}$ be parallel transported from a point P along the vector fields $\boldsymbol{V}$ and $\boldsymbol{U}$, that is

$$\boldsymbol{\nabla}_{\boldsymbol{V}} \boldsymbol{W} = 0 , \qquad \boldsymbol{\nabla}_{\boldsymbol{U}} \boldsymbol{W} = 0 . \tag{3.171}$$

It will arrive to a point D with a value $\boldsymbol{W}_{\rm PCD}$. Let now vector $\boldsymbol{W}$ be parallel transported from the same point P but, first, along $\boldsymbol{U}$ and,then, along $\boldsymbol{V}$ at the same parametric distances. Its destination point B will not coincide with the point D if the commutator $[\boldsymbol{V}, \boldsymbol{U}] \neq 0$, as shown in Figure 3.7. In order to compare the two results of the parallel transport, the vector $\boldsymbol{W}$ must be additionally parallel-transported,

$$\boldsymbol{\nabla}_{[\boldsymbol{V},\boldsymbol{U}]} \boldsymbol{W} = 0 , \tag{3.172}$$

along the direction of the non-vanishing commutator of the vector fields from the point B to the point D. After the parallel transport along this path vector $\boldsymbol{W}$ takes a value $\boldsymbol{W}_{\rm PCBD}$ at the point D. The difference, $\delta\boldsymbol{W} = \boldsymbol{W}_{\rm PCD} - \boldsymbol{W}_{\rm PCBD}$, between the two results of the parallel transport of vector $\boldsymbol{W}$ is given by

$$\delta\boldsymbol{W} = \left[\boldsymbol{\nabla}_{\boldsymbol{V}}, \boldsymbol{\nabla}_{\boldsymbol{U}}\right] \boldsymbol{W} - \boldsymbol{\nabla}_{[\boldsymbol{V},\boldsymbol{U}]} \boldsymbol{W} = \mathbb{R}(\boldsymbol{V}, \boldsymbol{U}) \boldsymbol{W} . \tag{3.173}$$

Equation (3.173) tells us that if a vector is parallel transported from one point of a manifold to another along two different paths it must have two different values at the point of destination if the Riemann tensor of the manifold is different from zero. Accordingly, parallel transportation of a vector (or a form, or a tensor) along a closed curve may render a vector (form, tensor) different from the original one. This property of the affine connection is known as a holonomy of an affine connection.

3.7.3 The holonomy of a connection

Let us consider a manifold and an infinitesimal contour around a point P bounded by a curve $\Gamma(\lambda)$ with a tangent vector $\boldsymbol{V} = d/d\lambda$ where λ is a parameter along the curve taking values from $\lambda = 0$ to $\lambda = \lambda_0$. Let a vector $\boldsymbol{W}$ be parallel transported along the curve so that at each point of the contour a covariant derivative $\boldsymbol{\nabla}_{\boldsymbol{V}} \boldsymbol{W} = 0$. Equation of the parallel transport can be solved and its solution represents a vector

37) Misunderstanding of this peculiarity of the non-holonomic vector bases has led some researchers [Wu and Ruan, 2003] to curious mathematical mistakes and misinterpretation of the foundations of Einstein's theory of general relativity.

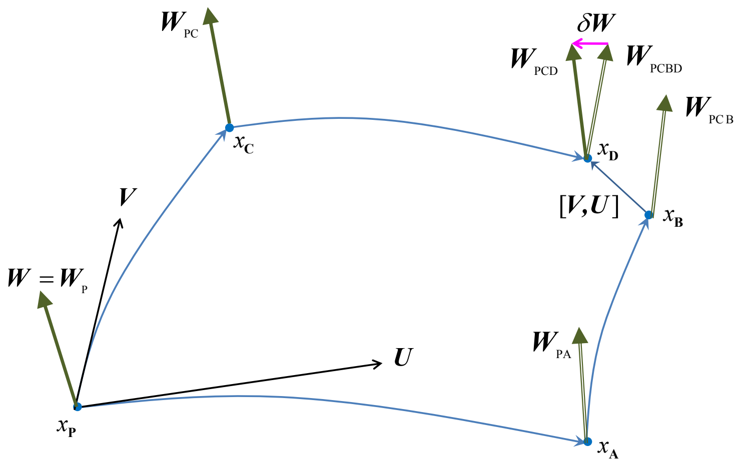

Figure 3.7 A vector $\boldsymbol{W}$ is parallel transported along infinitesimally-small paths PABD and PCD from a point P to a point D. The parallel transport generates a vector field $\boldsymbol{W}$ along the integral curves of the vector field $\boldsymbol{V}$ and $\boldsymbol{U}$. The difference $\delta\boldsymbol{W} = \boldsymbol{W}_{\rm PCD} - \boldsymbol{W}_{\rm PCBD}$ between the two results of the parallel transport, is proportional to the value of the Riemann tensor taken at the point P. One has $\delta\boldsymbol{W} = 0$, if and only if, the manifold is flat.

field $\boldsymbol{W}(\lambda)$ on the curve. This vector field is nonholonomic in the sense that after one moves vector $\boldsymbol{W}$ around the contour and returns it to a starting point, the difference $\delta\boldsymbol{W} = \boldsymbol{W}(\lambda_0) - \boldsymbol{W}(0) \neq 0$ (see Figure 3.8).

To quantify the holonomy, let us integrate a parallel-transported vector $\boldsymbol{W}$ along the contour. It is convenient to introduce a local coordinate chart x^α around the point P on the manifold, and integrate the covariant components W_α of the vector that are obtained with the metric tensor isomorphism, $W_\alpha = g_{\alpha\beta}W^\beta$. Suppose, one starts the integration at point $x_0^\alpha = x^\alpha(\lambda = 0)$ on the contour and return to the same point at another value of $\lambda = \lambda_0$. It gives us a change of the covector $\delta W_\alpha = W_\alpha(\lambda_0) - W_\alpha(0)$ accumulated over a walk about the contour,

$$\delta W_\alpha = \oint dW_\alpha = \oint \frac{dW_\alpha}{d\lambda} d\lambda = \oint \Gamma^\beta_{\alpha\gamma} W_\beta \frac{dx^\gamma}{d\lambda} d\lambda\,, \tag{3.174}$$

where one has used the equation of the parallel transport of the covector components W_α. To carry out the integration, one will make use of the fact that the contour $\Gamma(\lambda)$ is infinitesimally small. It will allow us to expand both functions in the integrand, $\Gamma^\beta_{\alpha\gamma}(x)$ and $W_\beta(x)$, about the initial point x_0^α to the leading order in $x^\alpha - x_0^\alpha$, where $x^\alpha = x^\alpha(\lambda)$,

$$\Gamma^\beta_{\alpha\gamma}(x) \;=\; \Gamma^\beta_{\alpha\gamma}(x_0) + (x^\delta - x_0^\delta)\Gamma^\beta_{\alpha\gamma,\delta}(x_0) + O\left(x - x_0\right)^2\,, \tag{3.175}$$

$$W_\beta(x) \;=\; W_\beta(x_0) + (x^\delta - x_0^\delta)\Gamma^\mu_{\beta\delta}(x_0)W_\mu(x_0) \;+ O\left(x - x_0\right)^2\,, \tag{3.176}$$

where comma denotes a partial derivative with respect to a corresponding coordinate, $\Gamma^\beta_{\alpha\gamma,\delta} = \partial\Gamma^\beta_{\alpha\gamma}/\partial x^\delta$, and one has again used the equation of the parallel transport to express the increment dW_β in terms of the Christoffel symbols.

Insertion of these in equation (3.174) results in

$$\begin{aligned}\delta W_\alpha \;&=\; \Gamma^\beta_{\alpha\gamma}(x_0)W_\beta(x_0)\oint dx^\gamma \\ &+\; \left[\Gamma^\beta_{\alpha\gamma,\delta}(x_0) + \Gamma^\mu_{\alpha\gamma}(x_0)\Gamma^\beta_{\mu\delta}(x_0)\right] W_\beta(x_0)\oint\left(x^\delta - x_0^\delta\right)dx^\gamma\,,\end{aligned} \tag{3.177}$$

where all quadratic and higher-order terms in the integrand have been neglected as they provide a negligible contribution. The terms being proportional to $\oint dx^\gamma$, vanish after integration over the closed contour. The integration in equation (3.177) will then be reduced to

$$\delta W_\alpha = \left[\Gamma^\beta_{\alpha\gamma,\delta}(x_0) + \Gamma^\mu_{\alpha\gamma}(x_0)\Gamma^\beta_{\mu\delta}(x_0)\right] W_\beta(x_0) S^{\delta\gamma}\,, \tag{3.178}$$

where

$$S^{\delta\gamma} = \oint x^\delta dx^\gamma = -\oint x^\gamma dx^\delta\,, \tag{3.179}$$

is the area encircled by the contour $\Gamma(\lambda)$. Because the area is an antisymmetric object, $S^{\delta\gamma} = -S^{\gamma\delta}$, one can recast equation (3.178) to

$$\delta W_\alpha = \frac{1}{2} R^\beta{}_{\alpha\delta\gamma} W_\beta S^{\delta\gamma}\,, \tag{3.180}$$

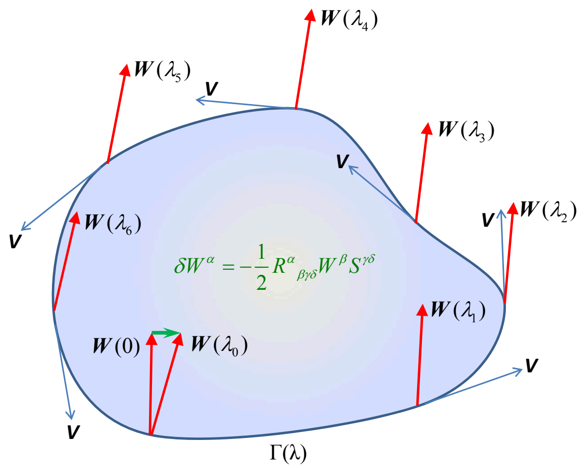

Figure 3.8 A vector $\boldsymbol{W}$ is parallel transported (counterclockwise) around an infinitesimal contour $\Gamma(\lambda)$. Nonholonomy of the affine connection is a failure of the parallel transport to preserve identity. The nonholonomy is measured by the variation of the vector $\delta\boldsymbol{W}$ that is proportional to a curvature tensor $R^\alpha{}_{\beta\gamma\delta}$ of a manifold and the area of the contour $S^{\gamma\delta}$.

where $R^{\beta}{}_{\alpha\delta\gamma}$ is the Riemann tensor given by equation (3.170). The holonomy equation (3.180) for the covector is equivalent to

$$\delta W^{\alpha} = -\frac{1}{2} R^{\alpha}{}_{\beta\delta\gamma} W^{\beta} S^{\delta\gamma} , \tag{3.181}$$

where one has used the symmetry of the Riemann tensor explained in the next section 3.7.6.

This result is trivially extended to a finite-sized contour by the standard trick borrowed from the proof of the Stokes theorem presented in any textbook on calculus: cover the surface inside the contour with an infinite set of infinitesimally small rectangles, and notice that the integral over the entire contour is equal to the sum of such integrals over the boundaries of all the rectangles. In this sum, all terms will cancel, except those corresponding to the outer boundary. This will enable us to rewrite formula (3.181) for the finite-size contour

$$\delta W^{\alpha} = -\frac{1}{2} \iint R^{\alpha}{}_{\beta\delta\gamma} W^{\beta} dS^{\delta\gamma} , \tag{3.182}$$

where the integral is over the surface encircled by the finite-size contour with the surface element of the integration $dS^{\delta\gamma} = -dS^{\gamma\delta}$. One sees that after a parallel transportation around the finite-size contour a vector retains its original value, if and only if, the Riemann tensor $R^{\alpha}{}_{\beta\gamma\delta} = 0$, that is the manifold is flat. It corresponds to the case of a holonomic connection. An alternative proof of equation (3.182) can be found, for example, in Schutz [1995].

3.7.4
The Riemann tensor as a measure of flatness

Among the most valuable features of the Riemann tensor is that it helps us to quantify the notion of flatness of a manifold. Manifold is called *flat* if it can be covered entirely with a Cartesian coordinates. We shall now build a mathematical criterium of flatness: *a manifold is flat if and only if the Riemann tensor* $\mathbb{R}$ *vanishes everywhere.*

Proving this criterium in one direction is easy. If the manifold is flat, and covered globally with Cartesian coordinates, then (in these coordinates) the affine connection is zero everywhere. Accordingly, $\mathbb{R} = 0$ on the entire manifold. A flat manifold can, of course, be covered with atlas consisting of non-Cartesian charts with curvilinear local coordinates. In such coordinates the Christoffel symbols are not zero. Nevertheless, as the existence of the global Cartesian coordinates was admitted, there exists a coordinate transformation from the global (Cartesian) to local (curvilinear) coordinates. Hence, the Riemann tensor calculated in the curvilinear coordinates will vanish, as can be easily seen from the tensor transformation law: if a tensor is nil in one coordinate system, it will remain nil in any other.

Now one has to prove the inverse statement: vanishing of the Riemann tensor entails flatness. In other words, $\mathbb{R} = 0$ is a sufficient condition for existence of the Cartesian coordinates covering the *entire* manifold. The proof will employ equation (3.16) describing transformation of the Christoffel symbols from coordinates $x^{\alpha'}$ to the others x^α. Let us reformulate equation (3.16) in the following form

$$\frac{\partial^2 x^{\beta'}}{\partial x^\mu \partial x^\nu} = \frac{\partial x^{\beta'}}{\partial x^\alpha}\Gamma^\alpha_{\mu\nu} - \frac{\partial x^{\lambda'}}{\partial x^\mu}\frac{\partial x^{\rho'}}{\partial x^\nu}\Gamma^{\beta'}_{\lambda'\rho'} , \tag{3.183}$$

and assume that coordinates $x^{\alpha'}$ are chosen in such a way that the Christoffel symbols $\Gamma^{\beta'}_{\lambda'\rho'} = 0$ in a neighborhood of a point P of the manifold. This can be always done, at least at one point of the neighborhood, as follows from the procedure of construction of the normal coordinates described in section 3.3.8.3. However, now one demands more than that - one assumes that not only $\Gamma^{\beta'}_{\lambda'\rho'} = 0$ at point P but at least all first derivatives of the Christoffel symbols also vanish: $\Gamma^{\beta'}_{\lambda'\rho',\mu'} = 0$ at this point. This is a rather restrictive demand as it imposes a ceratin limitation on equation (3.183).

Indeed, under these assumptions, the transformation equation (3.183) is reduced to a system of homogeneous linear equations in partial derivatives

$$\frac{\partial \Lambda^{\beta'}{}_\mu}{\partial x^\nu} = \Gamma^\alpha_{\mu\nu}\Lambda^{\beta'}{}_\alpha , \tag{3.184}$$

for a matrix of the coordinate transformation, $\Lambda^{\beta'}{}_\mu \equiv \partial x^{\beta'}/\partial x^\nu$, that obeys 6×4=24 restrictions

$$\frac{\partial \Lambda^{\beta'}{}_\mu}{\partial x^\nu} = \frac{\partial \Lambda^{\beta'}{}_\nu}{\partial x^\mu} , \qquad (\mu \neq \nu) . \tag{3.185}$$

The number of independent components of the matrix $\Lambda^{\beta'}{}_\mu$ is 16, while the number of the independent equations is 64-24=40. It means that the system of equations (3.184) is over-determined.

The Frobenius theorem gives necessary and sufficient conditions for finding a maximal set of independent solutions of an over-determined system of partial differential equations (3.184). Specifically, the second partial derivatives of the matrix $\Lambda^{\beta'}{}_\mu$ must commute

$$\frac{\partial^2 \Lambda^{\beta'}{}_\mu}{\partial x^\nu \partial x^\gamma} = \frac{\partial^2 \Lambda^{\beta'}{}_\mu}{\partial x^\gamma \partial x^\nu} . \tag{3.186}$$

Taking the corresponding partial derivatives from both sides of equation (3.184) and substituting them to equation (3.186) one obtains

$$\Lambda^{\beta'}{}_\alpha \left(\Gamma^\alpha_{\mu\nu,\gamma} - \Gamma^\alpha_{\mu\gamma,\nu} + \Gamma^\rho_{\mu\nu}\Gamma^\alpha_{\rho\gamma} - \Gamma^\rho_{\mu\gamma}\Gamma^\alpha_{\rho\nu}\right) = 0 , \tag{3.187}$$

or accounting for definition (3.170) of the Riemann tensor,

$$\Lambda^{\beta'}{}_\alpha R^\alpha{}_{\mu\gamma\nu} = 0 . \tag{3.188}$$

This equation tells us that the condition $\Gamma^{\beta'}_{\lambda'\rho'} = 0$ can be satisfied in a neighborhood of a point P, if and only if, the Riemann tensor $R^\alpha{}_{\mu\gamma\nu} = 0$. Because all local coordinate charts overlap on the manifold, and the transition functions between the charts are analytic, one comes to the conclusion that if the Riemann tensor is nil at least at one chart, it must be nil on the entire manifold, and the global coordinates can be constructed in which the Christoffel symbols vanish. In these coordinates the metric tensor is apparently constant as follows from the definition (3.34), and can be chosen coinciding with the Minkowski metric of flat spacetime.

It is worth noticing that the symmetry relations (3.185) can be also considered as consequences of the symmetry of the Christoffel symbols, $\Gamma^\alpha_{\mu\nu} = \Gamma^\alpha_{\nu\mu}$ that are standing in the right side of equation (3.184). It is possible to relax this restriction and to work with a non-holonomic basis admitting a non-symmetric affine connection that can also include the torsion tensor. The connection made up of the torsion, is known as the Weitzenböck connection [Weitzenböck, 1923]. The above-given proof of flatness of the manifold will be still valid for the non-holonomic bases and for the Weitzenböck connection. It demonstrates rather remarkably that the concept of flat manifold is more general than the standard Minkowski spacetime in the sense that it can admit co-existence of the Minkowski metric $\eta_{\alpha\beta}$ along with the additional geometric structure - the affine connection made up of torsion. This property of the flat manifold is a mathematical basis for a teleparallel theory of gravity [Blome *et al.*, 2010; Ferraro, R. and Fiorini, F. , 2008; Itin and Marolf, 2001]. Unlike general relativity, gravity in teleparallelism is not due to the curvature of spacetime (which is equal to zero) but due to the torsion. Proponents of the teleparallel point of view maintains the idea that curvature and torsion are simply alternative ways of describing the gravitational field. Any gravitational phenomenon that can be interpreted in terms of curvature, therefore, can also be interpreted in terms of torsion so that there exist equivalence between general relativity and teleparallel gravity [Aldrovandi and Pereira, 2007]. On the other hand, more general gravity theories, like Einstein-Cartan [Trautman, 2006] and the metric-affine groups [Heinicke *et al.*, 2005; Sotiriou and Liberati, 2007], consider curvature and torsion as representing independent degrees of freedom. In these theories, differently from teleparallel gravity, torsion might become relevant only when spins are important. This could be the case either at the microscopic level or near a rapidly rotating neutron stars or black holes. According to these models, therefore, since torsion represents additional degrees of freedom in relation to curvature, new physical phenomena should be expected from its presence. Mao *et al.* [2007] claimed that the effects of torsion may be expected in the solar system and possible measured by means of the Gravity Probe B (GP-B) experiment. This statement was confronted by Hehl and Obukhov [2007] on the basis that the GP-B gyroscope as a whole, has no uncompensated elementary spin which is crucial in coupling of torsion with gravitational field.

3.7.5
The Jacobi equation and the geodesics deviation

Consider a smooth, one-parameter congruence (vector flow) of geodesics, C_σ, numbered with aid of parameter σ. Each geodesic is a curve parameterized with an affine parameter τ. For each fixed σ, the appropriate geodesic has its tangent vector, $\boldsymbol{V} \equiv \boldsymbol{V}_\sigma(\tau)$, that satisfies equation (3.95): $\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{V} = 0$. Let us construct (see Figure 3.9) a so-called *Jacobi vector field*, $\boldsymbol{J} \equiv \boldsymbol{J}_\tau(\sigma)$, consisting of a congruence of curves each of which passing through the points of the geodesics having the same value of the parameter τ. Each curve from the Jacobi field is numbered with the parameter τ and is parameterized with the parameter σ being reckoned along it. By the construction, the Jacobi field is Lie transported along the congruence of the geodesics and obeys equation (3.157): $\pounds_{\boldsymbol{V}}\boldsymbol{J} = 0$. This, in its own turn, guarantees that the two vector fields commute, $[\boldsymbol{V}, \boldsymbol{J}] = 0$.

The Jacobi vector field evolves as it is Lie transported along the geodesic congruence. Its evolution is associated with the change of the parameter τ. Let us consider how much the Jacobi field changes as the parameter increases from the initial value $\tau = 0$ to some other sufficiently small value τ. Because one assumes all geometric objects to be analytic functions on the manifold, one can expand the Jacobi field $\boldsymbol{J} \equiv \boldsymbol{J}_\tau(\sigma)$ in a Taylor series around the initial value. We shall assume that the manifold is torsion-free and it is endowed with a Levi-Civita connection. Then, the Taylor expansion of the Jacobi field reads

$$\boldsymbol{J} = \boldsymbol{J}_0 + \tau\left(\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{J}\right)_{\tau=0} + \frac{\tau^2}{2}\left(\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{J}\right)_{\tau=0} + O\left(\tau^3\right) . \tag{3.189}$$

The first derivative of the Jacobi vector field $\boldsymbol{J}$ taken along the tangent vector $\boldsymbol{V}$ measures the relative speed of deviation of geodesics from one another. It depends on the initial conditions and may be different from zero. However, it characterizes a pure kinematic behavior of the geodesic congruence and does not reflect a dynamics of the relative separation of the geodesics as one moves along their congruence. The dynamics is revealed by the second derivative of $\boldsymbol{J}$ along the tangent vector $\boldsymbol{V}$ that yields the relative acceleration of the geodesics. Hence, one focuses on the calculation of the second derivative of $\boldsymbol{J}$ in equation (3.189).

As one has assumed that the manifold is torsion-free, the covariant derivatives of the Jacobi field $\boldsymbol{J}$ and the tangent vector field $\boldsymbol{V}$ must commute (see section 3.5.2): $\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{J} = \boldsymbol{\nabla}_{\boldsymbol{J}}\boldsymbol{V}$. Hence, the second derivative of $\boldsymbol{J}$ can be written as

$$\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{J} = \boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{\nabla}_{\boldsymbol{J}}\boldsymbol{V} = \left[\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{\nabla}_{\boldsymbol{J}}\right]\boldsymbol{V} , \tag{3.190}$$

where one has omitted a term $\boldsymbol{\nabla}_{\boldsymbol{J}}\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{V}$ as it equals to zero due to the geodesic equation $\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{V} = 0$ being valid at any point of the geodesic congruence.

The last term in equation (3.190) is the curvature operator $\mathbb{R}(\boldsymbol{V}, \boldsymbol{J})\boldsymbol{V}$ defined in equation (3.161), where the derivative along the commutator of two vector fields must be omitted due to the fact that the Jacobi field $\boldsymbol{J}$ commutes with the tangent vector $\boldsymbol{V}$ to the geodesics. Eventually, equation (3.190) can be recast to the final form

$$\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{\nabla}_{\boldsymbol{V}}\boldsymbol{J} + \mathbb{R}(\boldsymbol{J}, \boldsymbol{V})\boldsymbol{V} = 0 , \tag{3.191}$$

or in components

$$V^\mu V^\nu \boldsymbol{\nabla}_\mu \boldsymbol{\nabla}_\nu J^\alpha + R^\alpha{}_{\beta\gamma\delta} V^\beta V^\delta J^\gamma = 0 . \tag{3.192}$$

This equation is called the Jacobi equation in differential geometry [Fecko, 2006]. It is also known as an equation of the deviation of geodesics in general theory of relativity [de Felice and Clarke, 1990; Misner *et al.*, 1973; Wald, 1984]. The key observation ensuing from this formula is that the relative acceleration of nearby geodesics vanishes, if and only if, so does the Riemann tensor. The Jacobi equation is used for measuring the change in deviation of geodesics of two test particles in gravitational-wave detectors [LIGO, 2010] and in precise gradiometry as described in section 8.7 of the present book.

We emphasize that the Jacobi equation (3.191) has been derived under assumption that the manifold is torsion-free. Torsion will bring about additional terms to the equation of deviation of geodesics [Kleinert, 2008].

3.7.6 Properties of the Riemann tensor

3.7.6.1 *Algebraic symmetries*

Let us assume in this subsection that the dimension of the manifold $\mathcal{M}$ is n. Einstein's spacetime, $n = 4$, will be a particular case. We shall also assume that the manifold is pseudo-Riemannian with the components of the metric tensor $g_{\alpha\beta}$ in a local chart. The metric tensor allows us to lower an index in the definition (3.170) of the Riemann tensor, $R_{\alpha\beta\gamma\delta} = g_{\alpha\mu}R^{\mu}{}_{\beta\gamma\delta}$. The goal of this subsection is to study the symmetries of the Riemann tensor $R_{\alpha\beta\gamma\delta}$ which determine the number of its algebraically independent components.

Let us consider the Riemann tensor at a point P on the manifold $\mathcal{M}$. The symmetries of the Riemann tensor do not depend on the choice of coordinates. Hence, one can chose the normal coordinates with the origin at the point P. In these coordinates, the Christoffel symbols vanish at the point P and the Riemann tensor is reduced to a linear combination of the second derivatives from the metric tensor [Misner *et al.*, 1973]

$$R_{\alpha\beta\gamma\delta} = \frac{1}{2}\left(g_{\alpha\delta,\beta\gamma} - g_{\beta\delta,\alpha\gamma} - g_{\alpha\gamma,\beta\delta} + g_{\beta\gamma,\alpha\delta}\right) , \tag{3.193}$$

where comma with two indices after it denotes two partial derivatives with respect to the indices. Algebraic symmetries of the Riemann tensor follow immediately from equation (3.193). They are:

1. Antisymmetry with respect to two indices

$$R_{\alpha\beta\gamma\delta} = -R_{\beta\alpha\gamma\delta} = -R_{\alpha\beta\delta\gamma} = R_{\beta\alpha\delta\gamma} , \tag{3.194}$$

2. Symmetry with respect to a pair of indices

$$R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta} , \tag{3.195}$$

3. Cyclicity with respect to a permutation of any three indices, for example,

$$R_{\alpha\beta\gamma\delta} + R_{\alpha\gamma\delta\beta} + R_{\alpha\delta\beta\gamma} \equiv 0 . \tag{3.196}$$

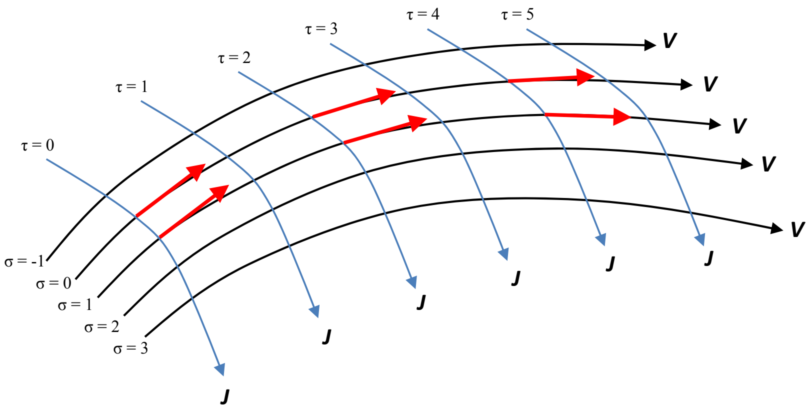

Figure 3.9 A vector field $\boldsymbol{V} = \boldsymbol{V}_\sigma(\tau)$ is formed by geodesics numbered with a parameter σ. The Jacobi vector field $\boldsymbol{J} \equiv \boldsymbol{J}_\tau(\sigma)$ is formed by the curves passing through the same values of the parameter τ on the geodesics. Two infinitesimally-close geodesics have their tangent vectors parallel to each other at an initial value of the parameter $\tau = 0$. As parameter τ increases the two geodesics deviate from each other and their tangent vectors gradually cease to be parallel.

Total number of components of the Riemann tensor on the manifold of dimension n is equal to n^4. However, the afore mentioned symmetries reduce the number of independent components of the Riemann tensor. The antisymmetry property tells us that there exist only $k = n(n-1)/2$ independent pairs (α,β) and the same number $k = n(n-1)/2$ of pairs (γ,δ). This gives, altogether, k^2 choices of the two pairs constituting the total subscript. However, the symmetry property reduces the number of choices of the pairs by $k(k-1)/2$, leaving us with only $k(k+1)/2$ choices. Their number will be restricted further by the cyclicity property. The number of extra constraints imposed by this property is equal to the number of choices of four different indices [38] out of n, which is: $n!/(n-3)!4! = n(n-1)(n-2)(n-3)/24$. Naturally, the number of these constraints vanishes for $n < 4$.

All in all, the number $N_R(n)$ of independent components of the curvature tensor $R_{\alpha\beta\gamma\delta}$ turns out to be equal

$$N_R(n) = \frac{1}{8}n(n-1)(n^2-n+2) - \frac{1}{24}n(n-1)(n-2)(n-3) = \frac{n^2(n^2-1)}{12} \,. \quad (3.197)$$

One can easily see that the Riemann tensor is identically zero in case of a one-dimensional manifold. On two-dimensional manifolds, the number of nonvanishing component of the curvature tensor is one, R_{1212}. Three-dimensional manifolds have six independent components of the Riemann tensor. A non-trivial case appears in four-dimensional manifold, $n = 4$, where the number of independent nonvanishing components of the Riemann tensor is 20. This is exactly the number of terms which prevented to nullify the second derivatives of the metric tensor in construction of the normal coordinates in section 3.3.8.3. It explains why the normal coordinates can not cover the entire four-dimensional manifold in the most general case – the curvature of the manifold leads to appearance of coordinate singularities and collapse of the grid of the normal coordinates [39].

3.7.6.2 *The Weyl tensor and the Ricci decomposition*

An important role in differential geometry is played by the Weyl conformal tensor $C_{\alpha\beta\gamma\delta}$, named after Hermann Weyl who showed that this tensor measures the deviation of a pseudo-Riemannian manifold from conformal flatness. In other words, if $C_{\alpha\beta\gamma\delta} = 0$, the manifold is conformally equivalent to a flat manifold, and its metric tensor $g_{\alpha\beta}$ can be expressed as a product of a scale factor (function) Ω^2 and the Minkowski metric: $g_{\alpha\beta} = \Omega^2\eta_{\alpha\beta}$.

On pseudo-Riemannian manifold of dimension n the Weyl tensor is obtained as an algebraic part of the Riemann tensor having the same symmetries as the Riemann tensor but, in addition, being trace-free with respect to transvection of any two

38) In fact, the symmetry property B makes the components $R_{\alpha\beta\gamma\delta}$ of the Riemann tensor fully antisymmetric with respect to a permutation of all four indices.
39) This remarks is applicable to any coordinates in curved manifolds.

indices

$$C_{\alpha\beta\gamma\delta} = R_{\alpha\beta\gamma\delta} + \frac{1}{n-2}\left(g_{\alpha\delta}R_{\gamma\beta} - g_{\alpha\gamma}R_{\delta\beta} - g_{\beta\delta}R_{\gamma\alpha} + g_{\beta\gamma}R_{\delta\alpha}\right) + \frac{R}{(n-1)(n-2)}\left(g_{\alpha\gamma}g_{\delta\beta} - g_{\alpha\delta}g_{\gamma\beta}\right) , \tag{3.198}$$

where

$$R_{\alpha\beta} = R^{\gamma}{}_{\alpha\gamma\beta} = g^{\alpha\gamma}R_{\alpha\beta\gamma\delta} , \tag{3.199}$$

is the Ricci curvature tensor, and

$$R = R^{\alpha}{}_{\alpha} = g^{\alpha\beta}R_{\alpha\beta} , \tag{3.200}$$

is the curvature scalar. An explicit expression for the Ricci tensor, obtained from transvection of indices in equation (3.170) in accordance to equation (3.199), reads

$$R_{\beta\delta} = \frac{\partial\Gamma^{\alpha}_{\beta\delta}}{\partial x^{\alpha}} - \frac{\partial\Gamma^{\alpha}_{\beta\alpha}}{\partial x^{\delta}} + \Gamma^{\alpha}_{\mu\alpha}\Gamma^{\mu}_{\beta\delta} - \Gamma^{\alpha}_{\mu\delta}\Gamma^{\mu}_{\beta\alpha} . \tag{3.201}$$

The Ricci tensor is symmetric as evident from equation (3.195) and can be independently confirmed by inspection of equation (3.201). Equation (3.198) is also known as the Ricci decomposition.

Equation (3.198) points out that the Weyl tensor in general relativity provides curvature to the spacetime when the Ricci tensor is zero, that is in vacuum. This equation also confirms the trace-free nature of the Weyl tensor,

$$C^{\alpha}{}_{\beta\alpha\delta} = 0 . \tag{3.202}$$

where one has taken into account that on a manifold with dimension n, the transvection of the metric tensor $g^{\alpha\beta}g_{\alpha\beta} = \delta^{\alpha}_{\alpha} = n$. Equation (3.202) gives $n(n+1)/2$ additional constraints on the components of the Weyl tensor $C_{\alpha\beta\gamma\delta}$ as compared with the Riemann tensor $R_{\alpha\beta\gamma\delta}$. Therefore, the number $N_C(n)$ of algebraically-independent components of the Weyl tensor is [Weinberg, 1972]

$$N_C(n) = \frac{1}{12}n^2(n^2-1) - \frac{1}{2}n(n+1) = \frac{1}{12}n(n+1)(n+2)(n-3) , \qquad (n \geq 3) . \tag{3.203}$$

This shows that on 3-dimensional manifolds the Weyl tensor vanishes identically so that the Riemann tensor can be expressed in terms of the Ricci tensor and Ricci scalar. In four-dimensional spacetime of general relativity the number of algebraically-independent components of the Weyl tensor is 10. Possible algebraic symmetries of the non-vanishing components of the Weyl tensor at each point in a pseudo-Riemannian manifold were classified by A. Z. Petrov [1954, 1969] [40)]

40) Concise introduction to Petrov's classification can be found in a review article by Frolov [1979].

The linear combination

$$G_{\alpha\beta} \equiv R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R \,, \tag{3.204}$$

is called *the Einstein tensor* . Its definition does not depend on the dimension of a manifold. The trace of the Einstein tensor

$$G^{\alpha}{}_{\alpha} = \frac{2-n}{2} R \,, \tag{3.205}$$

where n is dimension of the pseudo-Riemannian manifold. In general relativity $n = 4$, and $G^{\alpha}{}_{\alpha} = -R$ is the negative of the Ricci scalar. The Einstein tensor occupies the central place in general relativity. It appears in Einstein's equation for gravitational field.

3.7.6.3 *The Bianchi identities*

The Riemann tensor obeys important differential identities, in addition to the algebraic identities discussed above. In order to derive them, let us again assume the connection to be torsion-free and employ the normal coordinates with the origin at a point P of a manifold. First, one will calculate a covariant derivative from the Riemann tensor at the point P. One has in the normal coordinates

$$\nabla_{\alpha} R_{\beta\gamma\mu\nu} = R_{\beta\gamma\mu\nu,\alpha} \,, \tag{3.206}$$

where the comma before the index denotes a partial derivative, $f_{,\alpha} \equiv \partial f / \partial x^{\alpha}$, and all terms with the Christoffel symbols vanish as all of them are equal to zero at the point P. This property is again used for calculating the partial derivative from various terms entering the definition of the Riemann tensor in equation (3.206). It brings about

$$\nabla_{\alpha} R_{\beta\gamma\mu\nu} = \frac{1}{2} \left(g_{\beta\nu,\gamma\mu\alpha} - g_{\gamma\nu,\beta\mu\alpha} - g_{\beta\mu,\gamma\nu\alpha} + g_{\gamma\mu,\beta\nu\alpha} \right) . \tag{3.207}$$

Taking a cyclic permutations of indices α, β, γ one obtains the *Bianchi identities*

$$\nabla_{\alpha} R_{\beta\gamma\mu\nu} + \nabla_{\gamma} R_{\alpha\beta\mu\nu} + \nabla_{\beta} R_{\gamma\alpha\mu\nu} \equiv 0 \,. \tag{3.208}$$

These equations have been derived in the normal coordinates but they are covariant, that is valid in any other coordinates due to the tensorial nature of the covariant derivative and the Riemann tensor. Contracting the Bianchi identity (3.208) over indices α and γ, and then over β and μ, and using the axiom that a covariant derivative commutes with the operation of contraction of indices, one arrives at

$$\nabla_{\alpha} G^{\alpha\beta} \equiv 0 \,, \tag{3.209}$$

where $G^{\alpha\beta} = g^{\alpha\mu} g^{\beta\nu} G_{\mu\nu}$ is the contravariant Einstein tensor defined in (3.204).

Geometric interpretation of the Bianchi identity (3.208) has been found by É. Cartan who used a language of differential forms and mobile frames. This language was popularized by J.A. Wheeler in terms of the famous statement "the boundary of a boundary is zero" [Misner *et al.*, 1973, pp. 364-382]. It corresponds in general relativity to the law of conservation of the moment of rotation of a stress-energy tensor of gravitational field associated by Cartan with the Einstein tensor.

3.8 Mathematical and Physical Foundations of General Relativity

The physical principles of general relativity have been thoroughly discussed for decades since the time of its discovery by Einstein [Norton, 1993]. Einstein [1918] pointed out three principles on which the theory rested: (a) principle of covariance, (b) principle of equivalence, and (c) Mach's principle [41]. Subsequent scientific elaboration on the theory has refined and enriched the earlier developments. Nevertheless, there is no a universal agreement between various researchers about either the optimal way of deducing general relativity from a few axioms, or what minimal set of axioms should be specifically taken in order to develop the Einstein's theory of gravity in a self-consistent way [42]. For this reason, the chapter focuses not on developing a minimal set of axioms of general relativity, but rather on discussing the basic principles lying in the foundation of modern gravitational physics.

3.8.1 General covariance on curved manifolds

Section 3.2 discussed the *principle of general covariance* in application to flat (Minkowski) spacetime of special relativity. One has found that by making use of the formalism of covariant derivatives one can extend the form-invariance of the fundamental laws of physics to arbitrary (for example, accelerated or curvilinear) coordinates in contrast to the special covariance principle that is valid only in a subclass of inertial coordinates that move with respect to each other with constant velocities and do not rotate.

The principle of general covariance can be further extended to differential manifolds with curvature. This possibility played an important heuristic role for Einstein in an earlier work on the theory of general relativity. Einstein also recognized that the principle of general covariance alone may be insufficient to build the foundation for gravity theory. Indeed, a pure mathematical extension of the principle of covariance to curved manifolds allows us to write down most of the laws of physics in a covariant form by making use of the language of differential geometry. Loosely speaking, this can be achieved by replacing partial derivatives, $\partial/\partial x^\alpha$, acting upon geometric objects in the equations of special relativity with covariant derivatives $\boldsymbol{\nabla}_\alpha$ [43], and by replacing the Minkowski metric $\eta_{\alpha\beta}$ with the metric $g_{\alpha\beta}$ associated with the curved manifold. This is reasonable as any curved manifolds admits introduction of the normal coordinates in a local neighborhood of any point P at which the affine connection vanishes and the metric tensor is reduced to the Minkowski form $\eta_{\alpha\beta}$ (see section 3.3.8.3).

This observation was made by Erich Kretschmann in 1917. He,however, argued

41) Later on, Einstein excluded this principle as being too uncertain and controversial.

42) Compare, for example, how the gravity principles are approached by Ehlers *et al.* [1972]; Misner *et al.* [1973]; Synge [1964]; Weinberg [1972] and more recently by de Felice and Clarke [1990]; Kleinert [2008]

43) This replacement, in relativistic jargon, is termed as "comma goes to semi-colon" rule.

that besides bringing mathematical shine to the theory, the principle of covariance would reveal no new physics compared to the special relativity principle plus covariance of physical fields. Kretschmann [1917] addressed the skepticism of the covariance principle to the entire theory including general relativity. He insisted that the demand to put a theory to generally covariant form does not limit or restrict the range of acceptable theories of gravity but is simply a challenge to the theorist's mathematical ingenuity. In other words, in order to produce a new theory one needs to supply some other principle(s) besides the principle of general covariance. Initially dissident, this opinion, unfortunately, became popular. As a result of this, the principle of general covariance is often omitted in the modern books on gravity, though back in 1915-1916 it was a source of inspiration for Einstein.

There is nothing wrong in Kretschmann's way of reasoning from a pure mathematical point of view, that explains why it was concurred by majority of physicists [Norton, 1993]. However, extension of Kretschmann's objection to gravitational physics is too far-going endeavor pretending to make the principle of general covariance physically vacuous. Neither Einstein nor other notable physicists could agree with such an extrapolation of Kretschmann's philosophy [Norton, 1993]. Although it is true that the principle of general covariance gives us no unique prescription to build a relativistic theory from its pre-relativistic analogue, it nevertheless, allows us to understand much deeper the physical foundations underlying the theory, and to narrow down the search for the additional principles helping to single out an appropriate variant of the theory. For example, Einstein's judgement was that of two theoretical systems, both of which agree with experience, the one is to be preferred which, from the point of view of the differential calculus, is the simpler and more transparent [Einstein, 1918]. This reminds Occam's razor principle stating that when competing hypotheses are equal in other respects, the hypothesis that introduces the fewest assumptions and postulates the fewest entities while still sufficiently answering the question, is to be selected [44].

Anderson [1967] provided more elaborative arguments in support of Einstein's point of view by arguing that the covariant generalization of the pre-relativistic theory must not involve additional geometric structures besides the dynamic objects like the metric tensor and the affine connection. Any time as such an extraneous, or "absolute" or "kinematic" in Anderson's terminology, geometric object appears in the covariant formulation of the theory, the theory is physically ill-posed and has to be abandoned. According to Anderson and Gautreau [1969] the absolutes are defective in a sense that they affect the behavior of the dynamic geometric objects but are not affected by these objects in turn. For example, the Newtonian theory of gravity can be formulated in covariant form but it must include a timelike congruence of a vector field T^α in addition to the metric tensor, $g_{\alpha\beta}$, and the affine connection, ∇_α which is compatible with metric, $\nabla_\alpha g_{\mu\nu} = 0$. The goal of the absolute vector field T^α is to keep the spacetime sliced in absolute space and time irrespectively of the covariant

44) In a sense Einstein reformulated the first rule of reasoning in philosophy proposed by Isaac Newton: "One is to admit no more causes of natural things than such as are both true and sufficient to explain their appearances" [Newton, 1760].

formulation. Mathematically, this stratification tells us that the metric must be kept purely spacelike in the sense that $g_{\alpha\beta}T^{\beta} = 0$ which assumes that the vector field is compatible with the affine connection $\nabla_{\beta}T^{\alpha} = 0$ [Misner *et al.*, 1973, $12.2]. However, the presence of the absolute (kinematic) object - vector field T^{α} - is unnatural for dynamical theory of spacetime, not to say that such a vector field has never been observed.

Kretschmann [1917] also attacked the principle of general covariance from a different angle by appealing to symmetry-based theoretical consideration [Kox and Eisenstaedt, 2005]. In special relativity, the spacetime is isotropic and homogeneous, so that whatever coordinates are used, the metric tensor transformation admits 10-parametric symmetry group of Lorentz-Poincaré. In general relativity, the group of transformation (diffeomorphisms) of the metric tensor has no any symmetry in the most general case [45]. This argument can be translated to the language of Killing vectors which are solutions of the differential equation (3.158). The most general spacetime of Einstein's gravity theory has no Killing vectors at all as contrasted to special relativity where their number is 10. In the transition from a Lorentz covariant formulation of special relativity to a generally covariant formulation of Einstein's theory of gravity, the covariance group is expanded. However, the symmetry group of the metric is actually reduced from the Lorentz group to the identity group, for the general case.

Kretschmann considered this property of general relativity as a shortcoming but this criticism is unacceptable as it demands the existence of a prior geometry with a symmetric spacetime. However, any prior geometry constitutes an absolute geometric element of the theory in terms of Anderson [1967] while, when one makes an equation generally covariant, the general covariance principle demands that the metric tensor, $g_{\alpha\beta}$, and quantities derivable from it, appear only as dynamic geometric objects. The metric tensor obeys the Einstein field equations and is not supposed to drop out at the end of one's calculations or to restrict the equations; rather, it is exploited to represent gravitational field [Weinberg, 1972].

One sees that general covariance, which emerged already in the special relativity context, was not only a fancy mathematical construction but also an important precedent. As Disraeli said in one of his speeches, "*A precedent embalms a principle*".

3.8.2
General relativity principle links gravity to geometry

In order to incorporate gravity to the geometric structure of spacetime manifold, the principle of general covariance must be amended by accounting for the principle of equivalence explained in section 3.1. General covariance embraces the description of accelerated frames while the principle of equivalence equates the inertial force existing in accelerated frame to the homogeneous gravitational force. Such a com-

45) Symmetric spacetimes like the Schwarzschild black hole or the Friedmann-Robertson-Walker universe, do exist in general relativity. They are normally the exact solutions of Einstein's field equations [Stephani *et al.*, 2003].

bined principle Einstein called the *principle of general relativity* and used it to unify the Newtonian theory of gravity with special relativity, thus, developing relativistic theory of gravity - general relativity.

The general relativity principle tells us that gravitational force exerted on a test particle by external bodies must be identified with the components of the affine connection projected on the worldline of the test particle. This is where geometry and gravity embraces each other for the first time. The reason for this identification is that the affine connection vanishes at the origin of the normal coordinates, and so does the gravitational force measured by an observer being in a state of free fall. Because in the absence of forces a test particle moves along a straight line in special relativity, and an equivalent of a straight line on a curved manifold is geodesic (see section 3.4.4.2), one can draw a conclusion that the test particle must move along a geodesic of spacetime that is a curved differential manifold. The equation of geodesic depends on the affine connection that generalizes the concept of the Newtonian gravitational force. Hence, observing all possible trajectories of test particles moving in a given gravitational field one can measure all components of the affine connection.

Equation of geodesic (3.97) depends only on the symmetric part of the affine connection. Explicit expression (3.121) yields

$$\mathfrak{G}^{\alpha}_{(\beta\gamma)} = \Gamma^{\alpha}_{(\beta\gamma)} + \frac{1}{2}\left(T_{\gamma\alpha\beta} + T_{\beta\alpha\gamma}\right) + \mathfrak{D}^{\alpha}_{(\beta\gamma)}\,, \tag{3.210}$$

that shows that equations of geodesics must, in principle, depend not only on the symmetric part of the Levi-Civita connection $\Gamma^{\alpha}_{(\beta\gamma)}$ but on torsion, $T^{\alpha}_{\beta\gamma}$, and the nonmetricity $Q^{\alpha}_{\beta\gamma}$ as well. General relativity abandons torsion and nonmetricity from the affine connection which is postulated to be the Levi-Civita connection $\Gamma^{\alpha}_{(\beta\gamma)}$ [46]. Nonvanishing nonmetricity violates the isomorphism between tangent spaces which is inconsistent with stability of atomic spectra and International Atomic Time . There is no any observational evidences so far indicating that nonmetricity may differ from zero. As for the torsion, it seems that it couples only to the intrinsic spin of elementary particles but not to the rotational angular momentum of astronomical bodies [47]. General relativity assumes that the affine connection on the spacetime manifold is given by the Levi-Civita connection $\Gamma^{\alpha}_{\beta\gamma}$ which is linked to the metric tensor $g_{\alpha\beta}$.

The Newtonian gravity force, $\boldsymbol{F}$, acting on a test particle is proportional to a gradient of gravitational potential, $\boldsymbol{F} = -m\boldsymbol{\nabla}U$, where m is the gravitational mass of the particle. What is the analogue of the gravitational potential U in general relativity? The principle of equivalence was again a clue to Einstein in order to propose a natural answer to this question. If the components of gravitational force are to be entirely associated with the Christoffel symbols $\Gamma^{\alpha}_{\mu\nu}$, the gravitational potential is to be identified with the metric tensor, $g_{\alpha\beta}$, because the Christoffel symbols are uniquely expressed in the form of a linear combination of its gradient as shown in equation

46) This postulate is also accepted in the PPN formalism [Will, 1993], and is not tested in that theoretical framework.

47) Mao *et al.* [2007] argues that this assumption has a logical loophole which can and should be tested experimentally, and consider nonstandard torsion theories in which torsion can be generated by macroscopic rotating objects.

(3.129). The metric tensor $g_{\alpha\beta}$ has 16 components in four-dimensional spacetime that can be algebraically split in symmetric, $g_{(\alpha\beta)}$, and anti-symmetric, $g_{[\alpha\beta]}$, parts,

$$g_{\alpha\beta} = g_{(\alpha\beta)} + g_{[\alpha\beta]} \ . \tag{3.211}$$

Moffat [1979, 1995] proposed a nonsymmetric gravitational theory (NGT) in which the non-symmetric part of the metric tensor represents a new hypothetical force that may explain the observation of the flat rotation curves of galaxies. NGT is a complicated and rich theory but not easy to work with and may have problems with stability [Janssen and Prokopec, 2007]. It also requires the presence of torsion and nonmetricity in the affine connection that are currently incompatible with observations. Only the symmetric part of the metric tensor has a physical meaning in general relativity and in the most alternative theories of gravity, which postulates that $g_{[\alpha\beta]} \equiv 0$ in order to make the metric tensor compatible with the symmetric Minkowski metric, $\eta_{\alpha\beta}$, of special relativity. Symmetry of the metric tensor reduces the number of its algebraically-independent components from 16 to 10. These ten components of $g_{\alpha\beta}$ are considered as gravitational potentials in general theory of relativity.

Newtonian gravitational physics predicts existence of tidal force characterizing inhomogeneity of gravitational field in a neighborhood of a particle. This concept was introduced in section 1.2.6 where one has shown that the tidal force is proportional to the second (and higher-order) partial derivatives of the Newtonian gravitational potential U. Since the metric tensor $g_{\alpha\beta}$ generalizes the concept of the gravitational potential U from Newtonian physics to general relativity, one has to find out a corresponding geometric object which includes second partial derivatives from the metric tensor. There is only one such object on pseudo-Riemannian manifold - the Riemann tensor, $R_{\alpha\beta\gamma\delta}$, that characterizes the curvature of spacetime. The curvature is an intrinsic geometric property that can not be eliminated at a single point if the manifold is not globally flat. Mathematical expression for the Riemann tensor in terms of the second derivatives of the metric tensor is given by equation (3.193).

Summarizing, one can state that combined with the *main assumption* that the spacetime is a smooth pseudo-Riemannian manifold with metric $g_{\alpha\beta}$, the general relativity principle effectively establishes the following relationship between gravity and geometry:

1) Each point of the curved spacetime manifold admits a tangent space that is locally identified with the Minkowski spacetime of special relativity;
2) Local inertial frames[48)] of the tangent Minkowski spacetime have their origins moving along worldlines of freely-falling, massive test particles;
3) Massive test particles move along timelike geodesics, and massless particles (light) move along null geodesics of the curved spacetime manifold;
4) Locally homogeneous gravitational force exerted on a free-falling particle is identified with the projection of the Christoffel symbols on the worldline of the particle. It can be locally eliminated by transition to a freely-falling inertial frame;

48) We say *frames*, not a *frame*. At the same point of the spacetime, two closely-located freely-falling particles define different local-inertial frames, if they are uniformly moving relative to one another.

5) Tidal gravitational force in the neighborhood of a freely-falling particle is identified with the projection of the Riemann (curvature) tensor on the worldline of the particle. The tidal force causes the relative deviation of geodesics and can *not* be locally eliminated by transition to the freely-falling inertial frame;
6) The metric tensor, $g_{\alpha\beta}$, represents ten gravitational potentials which are a priory unknown dynamic variables that has to be derived from the gravity field equations.

These consequences of the principle of general relativity were contemplated by Einstein on his path from special to general relativity [Lanczos, 1972]. They constitute the mathematical foundation of Einstein's relativistic theory of gravity. Physical foundations of general relativity are based on the phenomenological description of matter, its particular way of coupling to gravity, and the specific formulation of the gravity field (Einstein) equations for the metric tensor.

Mathematically, the general relativity principle states that the metric tensor $g_{\alpha\beta} = g_{\alpha\beta}(w^\mu)$ of the spacetime manifold is reduced to the Minkowski metric $\eta_{\alpha\beta}$ at the origin of the freely-falling reference frame, parameterized with the normal coordinates, w^μ, (see section 3.3.8.3):

$$g_{\alpha\beta} = \eta_{\alpha\beta} + O(|w|^2) \ . \tag{3.212}$$

Herein, the linear terms are absent due to the principle of equivalence, and the residual quadratic terms are proportional to the Riemann tensor which characterizes the inhomogeneity of the gravitational field (tidal effects). It is worthwhile noticing that the quadratic terms ($\sim O(|w|^2)$ in the expansion of the metric tensor in the normal coordinates can take different form. In other words, the normal coordinates are not unique and admit a large freedom when one extrapolates them outside of their origin to the larger domain where the effects of curvature of spacetime become noticeable. For example, one possibility is to build, so-called, Riemann normal coordinates which are made of the bundle of identically-parameterized geodesics radiated out of the coordinate origin located at point P of the spacetime manifold. Calculations reveal that the Taylor expansion of the metric around the origin is [Ehlers, 1973; Misner *et al.*, 1973]:

$$g_{\alpha\beta} = \eta_{\alpha\beta} - \frac{1}{3}R_{\alpha\mu\beta\nu}w^\mu w^\nu + O(|w|^3) \ , \tag{3.213}$$

where the curvature tensor $R_{\alpha\mu\beta\nu}$ must be calculated at the point P. In fact, the normal coordinates can be introduced not only in a single event but along any timelike worldline [Schouten, 1954; Zel′manov and Agakov, 1989]. In the particular case of free fall of the origin of local coordinates, $w^\alpha = (w^0, w^i) = (cu, w^i)$, along timelike geodesic, the metric tensor assumes the following form, accurate to second order in the spatial coordinates [Ni and Zimmermann, 1978b]:

$$g_{00} = -1 - R_{0p0q}w^p w^q + O(|w|^3) \ , \tag{3.214a}$$

$$g_{0i} = -\frac{4}{3}R_{0pjq}w^p w^q + O(|w|^3) \ , \tag{3.214b}$$

$$g_{ij} = \delta_{ij} - \frac{1}{3}R_{ipjq}w^p w^q + O(|w|^3) \ . \tag{3.214c}$$

Here $R_{\alpha\beta\gamma\delta}$ are the components of the Riemann tensor along the geodesic world-line that depends only on the local coordinate time u. Such local coordinates are called *Fermi normal coordinates*. The metric (3.214) corresponds to the case of non-rotating spatial axes. There are other ways to build the normal coordinates which are discussed in more detail in section 4.5.

These mathematical circumstances build a roof beam between the special relativity and the theory of gravity making the principle of general relativity a powerful tool for formulating the laws of physics in the presence of gravitational field. This is achieved in most cases by straightforward replacing in the special-relativistic laws the Minkowski metric, $\eta_{\alpha\beta}$, with the full metric, $g_{\alpha\beta}$, and the partial derivatives, $\partial/\partial x^\alpha$, from the (non-gravitational) field variables with the covariant derivatives, ∇_α, defined in terms of the Christoffel symbols, $\Gamma^\alpha_{\beta\gamma}$, according to equation (3.129) [49]. Notwithstanding simplicity, this way of formulation of physics in the presence of gravity bears uncertainties that arise every time when one has to re-formulate a second-order partial differential equation from Minkowski to curved spacetime. The problem is hidden in non-commutativity of the second-order covariant derivatives caused by the curvature tensor, as explained in section 3.7.1. The general covariance principle does not guide us in choosing the order of the derivatives. Therefore, there is no a unified method of writing the right form of special-relativistic equations on the curved manifold - this ordering-derivatives ambiguity is as bad as the notorious ordering ambiguity of operators in quantum mechanics! The bottom-line is that in each physical setting, description of coupling with curvature is a separate, often involved, problem - see Misner *et al.* [1973, Section 16.3] for examples. In some cases additional arguments like the law of conservation, or the correspondence to the Newtonian gravity equations are helpful.

3.8.3
The equations of motion of test particles

A test particle is a point-like particle whose mass, charge, and spin are assumed to be so small that their effects on external gravitational field are negligible. Let us consider, first, motion of massive neutral particles without spin. This is because spin couples with the Riemann tensor and, though, it does not affect the background gravitational field, its existence produces an external force perturbing motion of the spin particle in a significant way.

Freely-falling test particle moves uniformly as viewed by an observer in a local inertial reference frame associated with the Minkowski spacetime that is tangent to the curved spacetime manifold. The uniformity of the motion means that there is no net force perturbing motion of the test particle in whatever direction. According to the general relativity principle the gravitational force exerted on the particle can be locally eliminated. It means that in the normal coordinates the equation of motion of

49) This phenomenological prescription of replacing the partial to covariant derivatives is called sometimes as "comma-goes-to-semicolon" rule [Misner *et al.*, 1973]. It is also known as the *principle of minimal coupling* of gravity and matter.

the test particle is

$$\frac{d^2 w^\alpha}{d\lambda^2} = 0 , \tag{3.215}$$

where λ is the affine parameter along the particle's worldline. Solution of this equation gives the first integral, $v^\alpha = dx^\alpha/d\lambda = \text{const.}$, and trajectory of the particle in the normal coordinates of the freely-falling frame

$$w^\alpha = w_0^\alpha + v^\alpha \lambda , \tag{3.216}$$

where w_0^α is particle's coordinate at $\lambda = 0$. Equation (3.216) is locally a straight line. The affine parameter can be chosen to coincide with the proper time τ of an ideal clock carried out by the particle (see section 2.5.1). Indeed, the interval $d\tau$ of the proper time between two events separated by the normal coordinate distance, dw^α, on the worldline of the particle is

$$-c^2 d\tau^2 = \eta_{\alpha\beta} dw^a dw^\beta = \eta_{\alpha\beta} v^a v^\beta d\lambda^2 , \tag{3.217}$$

where $\eta_{\alpha\beta} v^a v^\beta = \text{const.}$ Equations can be drastically simplified if one chooses parameter $\lambda = \tau$. Indeed, in such case velocity v^α gets equal to four-velocity of the particle, $u^\alpha = dw^a/d\tau$, that is normalized to $\eta_{\alpha\beta} u^a u^\beta = -c^2$ according to definition (2.163). After accepting the proper-time parametrization of the worldline of the test particle, equation (3.215) can be recast to the following form

$$u^\beta \partial_\beta u^\alpha = 0 , \tag{3.218}$$

where $\partial_\beta = \partial/\partial x^\beta$ is a partial derivative with respect to the normal coordinate x^β.

According to the general relativity principle, equation in local coordinates can be converted to equation in arbitrary coordinates x^α, on the curved manifold by replacing the partial derivative ∂_β with the covariant derivative $\boldsymbol{\nabla}_\beta$. It gives us the covariant equation of motion of test particle

$$u^\beta \boldsymbol{\nabla}_\beta u^\alpha = 0 , \tag{3.219}$$

where the four-velocity $u^\alpha = dx^\alpha/d\tau$ that is exactly the equation of geodesic (3.97) with the affine connection taken in the form of the Christoffel symbols

$$\frac{d^2 x^\alpha}{d\tau^2} + \Gamma^\alpha_{\beta\gamma} \frac{dx^\beta}{d\tau} \frac{dx^\gamma}{d\tau} = 0 . \tag{3.220}$$

This is ordinary differential equation of the second order has the first integral

$$g_{\alpha\beta} u^\alpha u^\beta = -c^2 , \tag{3.221}$$

which is obtained immediately from the condition that the parameter τ is the proper time along the worldline of the particle.

In practical cases of integration [50] of equation (3.220), it is more convenient to use coordinate time, $t = x^0/c$, instead of the proper time τ of the particle. Taking into account the time component of equation (3.220)

$$\frac{d^2t}{d\tau^2} = -\frac{1}{c}\Gamma^0_{\beta\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} , \tag{3.222}$$

and relationships

$$\frac{dx^i}{d\tau} = \frac{dx^i}{dt}\frac{dt}{d\tau} , \qquad \frac{d^2x^i}{d\tau^2} = \frac{d^2x^i}{dt^2}\left(\frac{dt}{d\tau}\right)^2 + \frac{dx^i}{dt}\frac{d^2t}{d\tau^2} , \tag{3.223}$$

yield the spatial component of equation (3.220)

$$\frac{d^2x^i}{dt^2} + \Gamma^i_{\beta\gamma}\frac{dx^\beta}{dt}\frac{dx^\gamma}{dt} = \frac{1}{c}\Gamma^0_{\beta\gamma}\frac{dx^\beta}{dt}\frac{dx^\gamma}{dt}\frac{dx^i}{dt} . \tag{3.224}$$

This form of the equation makes it clear that the coordinate time t is not an affine parameter of the trajectory [51]. Relationship between the coordinate time, t, and the proper time, τ, of the particle is calculated by integrating the first integral (3.221)

$$\tau = \frac{1}{c}\int_{t_0}^{t}\sqrt{-g_{\alpha\beta}\frac{dx^\alpha}{d\sigma}\frac{dx^\beta}{d\sigma}}d\sigma , \tag{3.225}$$

where t_0 is the initial epoch of integration, and σ is the integration parameter substituted for the coordinate time on the interval of integration, $\sigma \in [t_0, t]$.

The fact that the proper time τ can be excluded from the equation of geodesic becomes crucial when one applies the above formalism to photons or any other massless particle. As known from the Maxwell electrodynamics and other field theories, massless particles move with the speed of light. For such particles, the first integral of equations (3.220) becomes

$$g_{\alpha\beta}u^\alpha u^\beta = 0 , \tag{3.226}$$

which tells us that the proper time of massless particles (photons) is undefined on the null-cone hypersurface, and cannot be used. Employing the coordinate time t in the equations of massless particles yields the same equation of motion (3.224) as for massive test particles. We explore the motion of photons in gravitational field in much more detail in section 7.

What can one says about the equation of motion for spinning particles? Scrutiny analysis of this problem performed by previous researches, notably Mathisson [1937], Papapetrou [1951a] and Dixon [1979], showed that spin of a test particle couples with the curvature of spacetime manifold in a non-trivial manner. Due to this coupling the spinning particle does not move along a geodesic worldline [Barker

50) For example, for calculation of numerical ephemerides of satellites or planets of the solar system.
51) The affine parameter makes the right side of the equation of geodesic equal to zero as explained in section (3.4.4.2.

and O'Connell, 1974]. More explicitly, the Mathisson-Papapetrou-Dixon equation for a particle with spin have the following form

$$u^\mu \nabla_\mu p^\alpha = -\frac{1}{2} R^\alpha{}_{\beta\gamma\delta} u^\beta S^{\gamma\delta} , \tag{3.227a}$$

$$u^\mu \nabla_\mu S^{\alpha\beta} = 2 p^{[\alpha} u^{\beta]} , \tag{3.227b}$$

$$S^{\alpha\beta} p_\beta = 0 , \tag{3.227c}$$

where p^α is four-momentum of the particle, and $S^{\alpha\beta} = -S^{\beta\alpha}$ is the anti-symmetric spin tensor. The spin tensor is associated with the particle's intrinsic spin, S^α, as follows

$$S_\alpha = \sqrt{-g}\varepsilon_{\alpha\beta\gamma\delta} S^{\beta\gamma} u^\delta , \tag{3.228}$$

where $g = \det[g_{\alpha\beta}] < 0$, is the determinant of the matrix made of the components of the metric tensor, $\varepsilon_{\alpha\beta\gamma\delta}$ is a forth-rank, completely anti-symmetric Levi-Civita symbol defined by

$$\varepsilon_{\alpha\beta\gamma\delta} = \varepsilon^{\alpha\beta\gamma\delta} = \begin{cases} +1, \text{if } \alpha\beta\gamma\delta \text{ is an even permutation of } 0123, \\ -1, \text{if } \alpha\beta\gamma\delta \text{ is an odd permutation of } 0123, \\ \ \ 0, \text{if } \alpha\beta\gamma\delta \text{ are not all different.} \end{cases} \tag{3.229}$$

Four-momentum p^α of the particle is not proportional to its four-velocity u^α in the most general situation [Dixon, 1979; Ehlers and Rudolph, 1977]. In such case, the right side of equation (3.227b) does not vanish, and the spin of the particle is not parallel transported along its worldline. However, in case when one can use the approximation, $p^\alpha = mu^\alpha$, the equation (3.227b) is reduced to the equation of parallel transport for spin

$$\frac{dS_\alpha}{d\tau} = \Gamma^\beta_{\alpha\gamma} S_\beta u^\gamma . \tag{3.230}$$

This equation predicts that spin of a freely-falling test particle (gyroscope) will precess with respect to a global coordinates ("distant stars") as it moves along its orbit [Schiff, 1960; Weinberg, 1972]. The relativity gyroscope experiment, called Gravity Probe B (GP-B), has been funded by NASA and launched in 2004 to test this prediction. Specific findings and details of the experiment as well as prospects for further improvement of experimental results have been reported in paper by Everitt *et al.* [2009a].

Equation (3.227c) is the so-called Dixon-Tulczyjew spin supplementary condition [Dixon, 1979; Tulczyjew, 1959] that are imposed on spin of the particle to make the system of equations (3.227) fully determined. Dixon and Tulczyjew argued that in case when the four-momentum of the particle is not parallel to its four-velocity, their condition (3.227c) is more appropriate than Pirani condition, $S^{\alpha\beta}u_\beta = 0$, or Corinaldesi-Papapetrou condition $S^{0\alpha} = 0$ which holds in the global coordinates x^α. There is no consensus in literature on this subject despite of multitude of papers devoted to comparison of these conditions and their role in equations of motion. Section 6 will generalize equations of motion (3.227) of a spinning particle to the case of a moving rotating body which is a member of N-body astronomical system like the solar system.

3.8.4
The correspondence principle: the interaction of matter and geometry.

3.8.4.1 *The Newtonian gravitational potential and the metric tensor*

Any new theory should reduce to the well-established theory to which it corresponds, when the new theory is applied to the circumstances for which the less general theory is known to hold. The principle was applied by Einstein to build general relativity theory and must be applied in any alternative theory of gravity to make it consistent with the Newtonian theory in the limit of slow motion and weak gravitational field Will [1993]. This section demonstrates how the geodesic motion of test particles can be implemented to match the metric tensor in general relativity to the Minkowski metric and to the gravitational potential U of the Newtonian theory of gravity.

Geodesic equation (3.220) of freely-falling massive particle can be explicitly written down as follows:

$$\frac{dx^\alpha}{d\tau^2} = -c^2\Gamma^\alpha_{00}\left(\frac{dt}{d\tau}\right)^2 - 2c\Gamma^\alpha_{0i}\frac{dt}{d\tau}\frac{dx^i}{d\tau} - \Gamma^\alpha_{ij}\frac{dx^i}{d\tau}\frac{dx^j}{d\tau} \ , \tag{3.231}$$

where the time derivatives are taken with respect to the proper time τ of the particle. We shall assume that the particle is *slowly* moving which means that the three-velocity of the particle, $v^i = dx^i/dt$, is much smaller in the chosen coordinates, $x^\alpha = (x^0, x^i) = (ct, \boldsymbol{x})$, than the invariant speed c, that is $dx^i/dt \ll c$. The slow-motion approximation entails a strong inequality between the time derivatives taken with respect to the proper time τ along the particle's trajectory

$$\frac{dx^i}{d\tau} \ll c\frac{dt}{d\tau} \ . \tag{3.232}$$

The gravity field potential in general relativity is associated with the metric tensor $g_{\alpha\beta}$ that must always approximate the Minkowski metric $\eta_{\alpha\beta}$ in the local normal coordinates in accordance with equation (3.213). This approximation can be extended on the entire spacetime manifold under assumption that gravitational field is *weak*. Formally, it means that the metric tensor can be decomposed in two parts

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} \ , \tag{3.233}$$

where $h_{\alpha\beta} = h_{\alpha\beta}(t, \boldsymbol{x})$ is a small perturbation of the Minkowski metric in the chosen coordinates x^α, such that any component

$$h_{\alpha\beta} \ll 1 \ . \tag{3.234}$$

The weak-field approximation opens the doors for the perturbational approach in calculation of the Christoffel symbols defined in terms of the metric tensor in equation (3.129). In the leading order over $h_{\alpha\beta}$, the Christoffel symbols read

$$\Gamma^\alpha_{\beta\gamma} = \frac{1}{2}\eta^{\alpha\lambda}\left(\frac{\partial h_{\lambda\beta}}{\partial x^\gamma} + \frac{\partial h_{\lambda\gamma}}{\partial x^\beta} - \frac{\partial h_{\beta\gamma}}{\partial x^\lambda}\right) + O\left(|h_{\alpha\beta}|^2\right) \ , \tag{3.235}$$

We shall further assume that the gravity field is almost *stationary* that is the time derivatives from the metric tensor perturbations normalized to the invariant speed c, are much smaller than partial derivatives with respect to any spatial coordinate

$$\frac{1}{c}\frac{\partial h_{\alpha\beta}}{\partial t} \ll \frac{\partial h_{\alpha\beta}}{\partial x^i} \,. \tag{3.236}$$

Notice that the slow-motion approximation for the test particle and the approximation of the almost stationary gravitational field characterize different objects - the particle and the field, and should be imposed separately.

The field being *stationary* permits us to neglect the time-derivatives from $h_{\alpha\beta}$, while the slow-motion of the particle allows us to neglect its velocity in equation (3.231) which is reduced under these assumptions to

$$\frac{d^2x^\alpha}{d\tau^2} = \frac{c^2}{2}\eta^{\alpha\lambda}\frac{\partial h_{00}}{\partial x^\lambda}\left(\frac{dt}{d\tau}\right)^2 - c\eta^{\alpha\lambda}\frac{\partial h_{0\lambda}}{\partial t}\left(\frac{dt}{d\tau}\right)^2 - \frac{c}{2}\eta^{\alpha\lambda}\frac{\partial h_{0\lambda}}{\partial x^i}\frac{dt}{d\tau}\frac{dx^i}{d\tau} \,, \tag{3.237}$$

or in components

$$\frac{d^2t}{d\tau^2} = \frac{1}{2}\left(\frac{\partial h_{00}}{\partial t}\frac{dt}{d\tau} + \frac{\partial h_{00}}{\partial x^i}\frac{dx^i}{d\tau}\right)\frac{dt}{d\tau} \,, \tag{3.238}$$

$$\frac{d^2x^i}{d\tau^2} = \frac{c^2}{2}\frac{\partial h_{00}}{\partial x^i}\left(\frac{dt}{d\tau}\right)^2 \,. \tag{3.239}$$

Equation (3.238) can be easily integrated taking into account that

$$\frac{\partial h_{00}}{\partial t}\frac{dt}{d\tau} + \frac{\partial h_{00}}{\partial x^i}\frac{dx^i}{d\tau} = \frac{dh_{00}}{d\tau} \,, \tag{3.240}$$

along the particle's trajectory. It leads to

$$\left(\frac{dt}{d\tau}\right)^2 = C\exp h_{00} \,, \tag{3.241}$$

where C is a constant of integration. This constant must be taken equal to a factor $(1 - \boldsymbol{v}^2/c^2)^{-1}$, because in case of vanishing gravity ($h_{00} \to 0$) the velocity, $\boldsymbol{v}$, of the freely-moving particle must be constant, and the increment of coordinate time t must be related to the proper time τ by a special-relativistic equation (2.162). Taking into account the slow-motion and weak-gravity approximation and expanding both the exponent and $(1 - \boldsymbol{v}^2/c^2)^{-1}$ in Taylor series, yield the proper time τ as a function of the coordinate time t,

$$\tau = t - \frac{1}{2}\int_{t_0}^{t}\left(\frac{\boldsymbol{v}^2}{c^2} + h_{00}\right)ds + O\left(|h^2_{\alpha\beta}|\right) \,, \tag{3.242}$$

where t_0 is the initial epoch of integration at which $\tau(t_0) = t_0$, and integration is performed along the trajectory of the particle. This equation allows us to approximate equation (3.239) for particle's three velocity as follows

$$\frac{d^2x^i}{dt^2} = \frac{c^2}{2}\frac{\partial h_{00}}{\partial x^i} \,. \tag{3.243}$$

The correspondence principle demands this equation must match exactly the corresponding Newtonian equations of motion (3.1) and (3.3). This immediately tells us that in the Newtonian approximation, h_{00} is to be proportional to the Newtonian potential U. Identification of the Newtonian potential U with h_{00} needs some care, because the potential is defined up to an additive constant. In case of the gravitational field being produced by an isolated astronomical system, one can choose both U and h_{00} vanishing at infinity, in which case the difference between the two disappears, and one arrives to

$$h_{00} = \frac{2U}{c^2} \; . \tag{3.244}$$

Moreover, since U satisfies the Poisson equation (1.40), the metric tensor component h_{00} must satisfy the field equation

$$\triangle h_{00} = -\frac{8\pi G}{c^2}\rho \; , \tag{3.245}$$

where G is the universal gravitational constant, ρ is the mass density of matter, and $\triangle$ is the Laplace differential operator (Laplacian). Substituting equation (3.244) to (3.242) brings about an important relationship

$$\tau = t - \frac{1}{c^2}\int_{t_0}^{t}\left(\frac{1}{2}\boldsymbol{v}^2 + U\right)ds \; , \tag{3.246}$$

between the coordinate time t and the proper time τ that is used in relativistic ephemeris astronomy for integration of equations of motion of major planets of the solar system (see section 9.3 for more details).

3.8.4.2 *The Newtonian gravity and the Einstein field equations*

The correspondence principle does not allow us to derive all components of the metric tensors as functions of the distribution of mass and mass current densities. Nevertheless, it provides us with a powerful hint to the structure of the field equations beyond the Newtonian approximation as demonstrated by Ehlers [1973]; Lovelock [1972]. They key idea is that the metric tensor, $g_{\alpha\beta}$, replaces the Newtonian potential, U, as the field variable, and it must satisfy the second-order differential equations in partial derivatives which are reduced in the Newtonian approximation to the Poisson equation (3.245) for the Newtonian potential. The number of the equations must not exceed the number of algebraically independent components of the metric tensor. Moreover, the equations must be tensorial in accordance with the principle of the general covariance.

In subsection 3.7, one has seen that the Riemann tensor, $R^{\alpha}{}_{\beta\gamma\delta}$ is the *only* tensor, which is assembled of the metric and its derivatives, and which is a linear function of the second derivatives. Hence the only way of obtaining second-rank tensors. The Riemann tensor has twenty algebraically-independent components, which exceeds ten algebraically-independent components of the metric tensor. However, contracting the Riemann tensor with respect to two indices, and recalling the symmetry properties of the Riemann tensor, one obtains the Ricci tensor $R_{\alpha\beta} = R^{\mu}{}_{\alpha\mu\beta}$

that has exactly ten algebraically-independent components. One can try to use it in the left side of the relativistic gravity field equations as a differential operator replacing the Laplace operator, $\triangle$, of the Newtonian theory. The field equations must have a source of the gravitational field in the right side. In the Newtonian gravity the only source of gravity field is the mass density ρ. In relativity, the mass density enters the time-time component of the energy-momentum tensor, $T^{\alpha\beta}$. As explained in section 2.8, this tensor has exactly ten algebraically-independent components. Hence, it looks reasonable to use it as a source of the gravitational field. However, in the Newtonian approximation not only T^{00} but also the trace of the energy-momentum tensor, $T = T^{\alpha}{}_{\alpha}$, approximates the mass density ρ. Hence, one arrives to the following structure of the gravity field equations on a curved spacetime manifold in general relativity

$$R^{\alpha\beta} = \kappa\left(T^{\alpha\beta} + \zeta g_{\alpha\beta} T\right) , \tag{3.247}$$

where κ and ζ are constants which should be determined [52].

Let us determine ζ first. For this, the contracted Bianchi identity (3.209) are used. They tell us that, if the field equations are valid, the following identity for the covariant derivatives of the energy-momentum tensor must hold

$$\nabla_\beta T^{\alpha\beta} - \frac{1}{2}(1 + 2\zeta)\nabla^a T \equiv 0 . \tag{3.248}$$

The energy-momentum tensor is conserved in special relativity. This law of conservation must be extended to curved manifold by making use of the general relativity principle as discussed above in section 3.8.2, leading to

$$\nabla_\beta T^{\alpha\beta} = 0 , \tag{3.249}$$

which gives four equations of motion of matter in the presence of gravitational field, that is the perturbed metric tensor $g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}$. Equation (3.249) reduces equation (3.248) to

$$(1 + 2\zeta)\nabla^a T = 0 . \tag{3.250}$$

Since in general, $\nabla^a T \neq 0$, one gets $\zeta = -1/2$, and if this value is put back into (3.247), and the Einstein tensor, $G^{\alpha\beta} = R^{\alpha\beta} - (1/2)g^{\alpha\beta}R$, with the Ricci scalar, $R = g^{\alpha\beta}R_{\alpha\beta}$, are used, there results

$$G^{\alpha\beta} = \kappa T^{\alpha\beta} , \tag{3.251}$$

where the constant κ still has to be determined. Notice that Einstein's early attempts to write down the field equations by identifying $R_{\alpha\beta}$ with $T_{\alpha\beta}$ (that is taking $\zeta = 0$) did not go far because of a conflict with equation (3.250). Indeed, as can be seen from this equation, the case of $\zeta = 0$ imposes the strong limitation $\nabla^a T = 0$ that is valid only for the matter with the trace of the energy-momentum tensor being identically equal to zero. In particularly, this is valid for electromagnetic field [53]

52) There are only two constants to find out. One might think that there could be a third constant, ς, in equation (3.247) in the form of $R^{\alpha\beta} + \varsigma g_{\alpha\beta} R = \kappa\left(T^{\alpha\beta} + \zeta g_{\alpha\beta} T\right)$. However, contraction of this equation with respect to free indices reveals that ς is an algebraic combination of κ and ζ, and is not independent.
53) See equation (2.385).

and for other massless particles but not for slowly-moving matter.

To determine the value of κ, one exploits the slow-motion and weak-field approximation to reduce the covariant field equations (3.251) to the Newtonian limit. For simplicity, one will use the tensor of energy-momentum of a perfect fluid (2.371). In the slow-motion approximation, three-velocity of matter $v \ll c$, pressure $p \ll \rho c^2$, and the energy-momentum tensor components satisfy the inequality, $T^{00} \gg T^{0i} \gg T^{ij}$. Accordingly, the same inequality is hold for the components of the Einstein tensor, $G^{00} \gg G^{0i} \gg G^{ij}$ due to the field equations (3.251). Hence, the main terms in the field equations are reduced to

$$R^{00} + \frac{1}{2}R = \kappa T^{00} \simeq \kappa c^2 \rho \,, \tag{3.252}$$

$$R^{ij} - \frac{1}{2}\delta_{ij}R = \kappa T^{ij} \simeq 0 \,, \tag{3.253}$$

$$R^{0i} = \kappa T^{0i} \simeq 0 \,. \tag{3.254}$$

where one has approximated the metric tensor $g^{\alpha\beta}$ with $g^{\alpha\beta} = \eta^{ab} = \mathrm{diag}(-1, 1, 1, 1)$, and T^{00} with ρc^2, which dominates the energy-momentum tensor in the slow-motion approximation [54)].

Since for weak gravity one has $g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}$ with $h_{\alpha\beta} \ll 1$, the scalar curvature, $R = g_{\alpha\beta}R^{\alpha\beta}$, can be approximated with

$$R = -R^{00} + R^{ii} = \frac{3}{2}R - R^{00} \,, \tag{3.255}$$

where equations (3.253) and (3.254) have been used. It gives us

$$R = 2R^{00} \,, \tag{3.256}$$

insertion whereof in equation (3.252) results in

$$R^{00} = \frac{\kappa}{2}\rho c^2 \,. \tag{3.257}$$

An expression for R^{00} ensues from the formula (3.199) for the Ricci tensor, which is in the weak-field approximation reads

$$R^{00} = R^{\mu}{}_{0\mu 0} = R^{i}{}_{0i0} = R_{i0i0} \,, \tag{3.258}$$

where the symmetry properties of the Riemann tensor has been employed. Assuming now that the gravity field is also stationary, one can neglect all time-derivatives in calculating R_{i0i0} which can be done in the approximation under consideration with the help of equation (3.193). This results in

$$R^{00} = -\frac{1}{2}\Delta h_{00} \,, \tag{3.259}$$

54) Other components of the energy-momentum tensor are much smaller that T^{00} in the slow-motion approximation. Specifically: $|T^{0i}|/T^{00} \sim v/c \ll 1$ and $|T^{ij}|/T^{00} \sim v^2/c^2 \ll 1$.

so equation (3.257) becomes

$$\triangle h_{00} = -\kappa \rho c^2 \,. \tag{3.260}$$

According to the principle of correspondence, equations (3.260) and (3.245) must be equal exactly. It tells us that the constant

$$\kappa = \frac{8\pi G}{c^4} \,, \tag{3.261}$$

and the field equations assume the final form

$$G^{\alpha\beta} = \frac{8\pi G}{c^4} T^{\alpha\beta} \,. \tag{3.262}$$

These equations have been derived by Einstein and are called the Einstein field equations. They provide a means for constructing the geometry and the gravitational field in the form of the metric tensor $g_{\alpha\beta}$ when the energy-momentum tensor $T^{\alpha\beta}$ (the source of the gravitational field) and suitable boundary conditions are given. The Einstein equations are non-linear, therefore, the procedure of finding their solutions is very hard. There are few exact solutions (like Schwarzschild, Kerr, Friedmann) which are physically relevant in astrophysics and cosmology. In relativistic celestial mechanics of N-body system one should resort to approximations to solve the Einstein equations (see Chapter 4). From the field equations one can also derive equations of motion of test particles and extended bodies which must reduce to the Newtonian equations of motion in the case of weak fields and low velocities. We discuss the derivation of the equations of motion from the field equations in Chapter 6.

3.8.5
The principle of the gauge invariance

One should stress a special aspect of the physical character of the metric in general relativity: the metric is not given a priori but is influenced or determined by the matter distribution via invariant field equations (3.262). General theory of relativity put together on the same mathematical basis the Minkowski metric, $\eta_{\alpha\beta}$, and the potentials of gravitational field, $h_{\alpha\beta}$, to form a single geometric object - the metric tensor,

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} \,, \tag{3.263}$$

which becomes the sole gravity-field variable of the Einstein equations. There are no other geometric objects made exclusively of the Minkowski metric, $\eta_{\alpha\beta}$, and its linear combination with additional vector or tensor fields besides $h_{\alpha\beta}$. Gravitation, according to Einstein, is a pure geometric phenomenon, a mode of expression of the metrical field $g_{\alpha\beta}$, nothing else [55]. This interpretation, however, should be handled

55) Alternative gravity theories relax this requirement and introduce other long-range tensor fields of various ranks that can couple to gravity.

with care as it introduces a specific freedom to the mathematical description of the gravitational field. As a consequence of the principle of general covariance, the components $g_{\alpha\beta}$ of metric tensor are allowed to be expressed in arbitrary coordinates related to each other by a smooth coordinate transformation (diffeomorphism): $x^\alpha = x^\alpha(x^{\beta'})$. Tensor transformation of the metric tensor from one coordinate chart to another is given by

$$g_{\alpha\beta}(x) = g_{\mu'\nu'}(x')\frac{\partial x^{\mu'}}{\partial x^\alpha}\frac{\partial x^{\nu'}}{\partial x^\beta} , \tag{3.264}$$

where $g_{\alpha\beta}(x)$ and $g_{\mu'\nu'}(x')$ are components of the metric tensor in coordinates x^α and $x^{\alpha'}$ respectively. However, the metric tensor transformation can not be associated with the physical change of gravitational field if it is governed by a self-consistent system of mathematical equations. In other words, if general relativity is a correct theory of gravitational field and $g_{\alpha\beta}(x)$ is a solution of the Einstein equations (3.262) with appropriate boundary and initial conditions, the metric $g_{\mu'\nu'}(x')$ must be also a solution of the same equations in coordinates $x^{\alpha'}$ with the boundary and initial conditions being diffeomorphic to those imposed in coordinates x^α. Briefly, *while the particular descriptions of gravitational field in various coordinates may look different mathematically, the gravitational interaction between the bodies must remain the same*. This is the principle of the gauge invariance of gravitational field.

The gauge invariance of the metric tensor brings about four degrees of the coordinate freedom to ten algebraically-independent components of the metric tensor. It may look like a contradiction since the number of the Einstein equations is ten and they may look over-determined. Fortunately, this is not the case because of the Bianchi identities

$$\nabla_\beta G^{\alpha\beta} \equiv 0 , \tag{3.265}$$

which are satisfied for any solution of the Einstein equations and initial and boundary conditions. The Bianchi identities reduce the number of functionally-independent Einstein equations to six that exactly coincides with the number of functionally-independent components of the metric tensor making the gravity field equations fully determined and gauge-invariant. The gauge-invariant nature of the Einstein theory of gravity to uniquely determine the metric tensor $g_{\alpha\beta}$ is similar to the gauge freedom of the Maxwell equations for electromagnetic field that was discussed in section 2.1.5. The difference is that in electrodynamics there is only one degree of freedom in choosing the solution of the Maxwell equations for electromagnetic vector potential, A^α. Moreover, the gauge transformation (2.24) in electrodynamics, changes the potential A^α without involving a coordinate transformation. One can have different mathematical descriptions of electromagnetic field in one and the same coordinate chart. In general relativity the gauge transformation of the metric and the coordinate transformation are intimately related to each other. Any coordinate transformation brings about the gauge transformation of the metric tensor. Opposite statement is, however, not true. There are physically different solutions of the Einstein equations that are not related through diffeomorphism. The gauge freedom complicates comparison of the solutions of Einstein's equations. Two metric tensors may be physically different, that is to describe dissimilar gravitational fields, or be equivalent in the sense that there is a coordinate diffeomorphism (3.264) formally transforming one mathematical form of the metric to another.

In order to facilitate solution of the Einstein equations four coordinate, also called gauge, conditions should be imposed on the components of the metric tensor. Coordinate conditions must not be generally covariant because they are supposed to pick out a certain set of coordinates, and not others. The choice of the conditions depend on a particular situation under consideration. One convenient choice, called the *harmonic gauge*, [56] is frequently used in relativistic celestial mechanics. It is represented by equation

$$\Gamma^{\alpha} \equiv g^{\beta\gamma}\Gamma^{\alpha}_{\beta\gamma} = 0 , \tag{3.266}$$

that singles out a rather large class of *harmonic coordinates*. In order to understand why the coordinates are called harmonic, let us consider a covariant D'Alambertian for a scalar field ϕ that is defined by

$$\Box_{\mathrm{g}}\phi = g^{\mu\nu}\boldsymbol{\nabla}_{\mu}\boldsymbol{\nabla}_{\nu}\phi . \tag{3.267}$$

Any function which satisfy a homogeneous D'Alamert's equation

$$\Box_{\mathrm{g}}\phi = 0 , \tag{3.268}$$

is called harmonic. Opening the covariant derivatives with the help of equation (3.91), yields

$$\Box_{\mathrm{g}}\phi = g^{\mu\nu}\partial_{\mu\nu}\phi - \Gamma^{\mu}\partial_{\mu}\phi . \tag{3.269}$$

If one chose the harmonic condition (3.266), and substitute for ϕ a coordinate x^{α}, which is a scalar, one finds

$$\Box_{\mathrm{g}}x^{\alpha} = 0 . \tag{3.270}$$

This explains why the coordinates are called harmonic if the condition $\Gamma^{\alpha} = 0$ is imposed on the metric tensor. The Cartesian inertial coordinates used in special relativity also satisfy d'Alembert's equation, so the harmonic coordinates in curved spacetime are the closest approximation available in general relativity to a global inertial frame of reference in special relativity. This point of view led Fock [1964] to believe that harmonic coordinates have certain physical advantages in building a foundation of general relativity. Infeld and Plebanski [1960] and other physicists [Gorelik, 1993] strongly criticized a privileged role of the harmonic coordinates. Indeed, the harmonic coordinates are convenient to a certain extent for doing mathematical calculations in slow-motion and weak-field approximation of an astronomical N-body system but they have bad analytic properties at infinity in the presence

56) This gauge is also known as de Donder gauge condition [Fock, 1964].

of gravitational radiation emitted by the system [Blanchet, 1987]. Moreover, Fock [1964] suggested that harmonic coordinates have a privileged inertial property in that they are defined up to the Lorentz transformation in asymptotically flat spacetime. Fock argued that this is an objective argument in favor of the Copernican system of the world with the origin taken at the Sun because the harmonic coordinates can not move along an accelerated worldline. We pointed out [Kopejkin, 1988a] that this property of the harmonic coordinates is not absolute and depends on the boundary conditions imposed on the solution of the Einstein equations. There are different boundary conditions that are more relaxed and admit construction of the local harmonic charts. The local harmonic coordinates have no privileged position with respect to "distant stars" and can move along any timelike worldline, for example, that of Earth's geocenter. Hence, Fock's "heliocentricism" of the harmonic coordinates can not be accepted.

The main question raised by any gauge theory hinges upon how one should understand the relationship between mathematics and physics. This concerns the question of which quantities represent the "physically real" properties. The answer to this question is well known in classic electrodynamics [57] where the physically real quantities are the electric, $\boldsymbol{E}$, and magnetic, $\boldsymbol{B}$ fields. It is these fields which make up the Lorentz force (2.21) exerted on electric charges. The force and the field do not depend on the choice of the gauge condition, and remain invariant under the gauge transformations (2.24).

In general relativity the non-trivial quantities that can be predicted by the theory and physically measured by an observer are, by definition, the gauge invariant quantities. They are called *observables* [Bergmann, 1961]. The observables are scalars taken at some point of spacetime manifold, therefore, their values remain the same irrespectively of the choice of a gauge condition and coordinates. If spacetime geometry possesses some symmetry (like in cosmology) the observables can be associated with the Killing vectors [58] corresponding to this symmetry. In more general cases, one can use the canonical variables of ADM formalism [Arnowitt *et al.*, 1962, 2008], Zelmanov's invariants [59] [Zel′manov, 1944; Zel′manov and Agakov, 1989] or Bergmann's "intrinsic coordinates" [Bergmann, 1961]. In relativistic celestial mechanics the observables are ordinarily associated with the proper time, an angle between two light rays, an electromagnetic wave frequency, etc. The problem, however, is that they are not the best choice to represent dynamic variables for solving Einstein's equations. More simple computational procedure is to solve the Einstein equations for the metric tensor in a particular gauge (coordinates) and

57) The gauge-invariance of electrodynamics is more subtle in quantum mechanics as the Schrödinger equation predicts that the wave function of charged particles interferes directly with the electromagnetic potential A^α [Aharonov and Bohm, 1959, 1961]. This, *Aharonov-Bohm effect*, was confirmed experimentally [Peshkin and Tonomura, 1989]. General relativity is a theory of classical gravitational field, so one does not need to discuss the gauge invariance in "quantum gravity" domain [Kleinert, 2008].

58) The concept of the Killing vectors was introduced in section 3.6.5.

59) Zel′manov and Agakov [1989] introduce three different types of invariants: chronometric, kinemetric and orthometric. Formalism of the kinemetric invariants is similar to the ADM formalism.

to use the solution in order to predict theoretical values of the observables which are compared then, with their experimental values. Usually, this procedure is considered to be sufficient in practical calculations in relativity such as construction of numerical ephemerides of major planets of the solar system [Moyer, 2003] or in data processing of binary pulsars [Lorimer and Kramer, 2004] or in experimental testing of general relativity [Will, 1993]. We emphasize, however, that the consistency of the theoretical predictions of the observables with their experimental values can not serve as a confirmation criterion of the gauge-invariance of the gravity theory. We advocate more advanced approach based on the idea that the gauge-invariance must be tested independently by doing all workable gauge transformations of the mathematical objects and elements of the data processing algorithm, and by checking its computational stability versus such transformations [Kopeikin *et al.*, 2007; Kopeikin, 2009; Xie and Kopeikin, 2010]. This procedure is to infiltrate the spurious contributions to the observables that may be generated by the computer code due to uncontrollable propagation of the gauge condition imposed on the initial data, to the future.

3.8.6
Principles of measurement of gravitational field

3.8.6.1 *Clocks and Rulers*

General relativity defines geometry of the spacetime manifold with the metric tensor, $g_{\alpha\beta}$, that is identified with the gravitational field potentials. Hence, measurement of the gravitational field and the measurement of spacetime geometry of the manifold are the equivalent concepts in general relativity. The study of spacetime geometry requires introduction of standard clocks and rules in order to measure time and length between points (events) of the spacetime manifold. In accordance with the correspondence principle, the standards in general relativity must be compatible with those introduced earlier in special relativity.

The invariant measure of time and space distance in special relativity is the interval, ds, introduced in section 2.2.4.4. Similarly to special relativity, the metric tensor in general relativity determines the interval, ds, between two neighboring spacetime points

$$ds^2 = g_{\alpha\beta}(x)dx^\alpha dx^\beta \;, \tag{3.271}$$

where x^α are the coordinates of the first point, and $x^\alpha + dx^\alpha$ are the coordinates of the second point. Contrary to special relativity, the metric tensor in general relativity can not be transformed to the Minkowski metric, $\eta_{\alpha\beta}$, globally, that is everywhere on the manifold, but only locally in a close vicinity to each point as shown in equation (3.213). The Minkowski metric describes the flat, pseudo-Euclidean geometry of tangent space at each point of the manifold but it does not characterize any property of gravitational field that is essentially a non-localizable geometric entity. The non-local geometric properties of the manifold are enciphered to its affine connection, $\Gamma^\alpha_{\beta\gamma}$, and the curvature tensor, $R^\alpha{}_{\beta\gamma\delta}$, that are expressed through the first and second partial derivatives of the metric tensor [60]. Hence, the experimental gravitational physics is to investigate the geometric properties of spacetime manifold in the domains being large enough to measure the partial derivatives of the metric tensor. This can be carried out operationally by observing motion of test particles, electromagnetic waves, and massive bodies. Relativistic celestial mechanics, astrometry, and geodesy provide scrutiny analyzes of these observations as will be explained in subsequent chapters of this book.

The local presence of the Minkowski metric separates the interval in three different types (see section 2.2.4.4):

$ds^2 < 0$ - timelike interval;
$ds^2 = 0$ - null interval;
$ds^2 > 0$ - spacelike interval.

Initially, Einstein used all three types of the interval in discussion of space and time measurements in terms of ideal rods (rulers) and clocks. The ideal rods were used for measuring the infinitesimal length, $d\ell = \sqrt{g_{\alpha\beta}(x)dx^\alpha dx^\beta}$, identified with the spacelike interval. The ideal clocks were used for measuring the infinitesimal intervals of proper time, $d\tau = c^{-1}\sqrt{-g_{\alpha\beta}(x)dx^\alpha dx^\beta}$, identified with the timelike interval along the clock's worldline. Light signals move in vacuum on the null hypersurface with invariant speed c, and they were used by Einstein for definition of the local synchronization (Einstein's synchronization) of clocks moving sufficiently close to each other along timelike worldlines. Rods and clocks are additional structures introduced to the theory of gravity "by hands" and their compatibility with the foundation of the theory must be carefully explored. In fact, the structure of physical rods and clocks is ultimately defined by the laws of atomic physics, which is not a constituent part of classical general relativity. Moreover, fundamental constants of atomic physics [61] are independent from the fundamental constants of general relativity - the universal gravitational constant, G, and the invariant speed, c. Therefore, the behavior of the "standard" rods and clocks may depend on time, place, and their worldlines if the laws of atomic physics are violated in some approximation or the constants of atomic physics evolve as the universe expands. For this reason, it would be preferable to eliminate from the process of measurement of gravitational field all measuring tools which working substance is not described by general relativity. Whether it is possible depends on the completeness of general relativity, that is its ability to describe measurements without attracting other complementary theories [62].

60) In alternative theories of gravity additional geometric structures like torsion and nonmetricity should be discussed as well [Hehl and Obukhov, 2007].

61) Primarily, it is the reduced Planck constant, $\hbar = 1.05457168 \times 10^{-34}$ J·s. However, one should include the elementary charge, $e = 1.60217653 \times 10^{-19}$ C, and mass of electron, $m_e = 9.1093826 \times 10^{-31}$ kg, and proton, $m_p = 1.67262158 \times 10^{-27}$ kg, as the physical rods and clocks are made of the elementary particles bound by atomic forces.

62) Each physical theory must have a predictive power and deals with measurements which goals is to test the predictions and to decide whether the theory is right (realistic) or wrong. In some cases, description of the measuring device requires to resort to another theory so that interpretation of the measurement includes the elements of both theories. Theory which mathematical formalism allows both to make

Marzke and Wheeler [1964] and Kundt and Hoffmann [1962] showed how to build a clock based on the propagation and reflection of light between two worldlines. Operation of such photon clock can be described exclusively in the framework of general relativity alone, thus, making measurement of time independent of the postulates of atomic physics. Manufacturing of photon clock is challenging task that has not yet been performed. At the same time, the quality of atomic clocks is constantly improving, and their practical performance as time-keepers is outstanding [Bauch, 2003; Gibble, 2007]. Therefore, the best way for experimental gravitational physics to proceed is to accept three postulates:

1) an ideal physical clock that operates on the quantum mechanical principles (atomic clock), shows the proper time, τ, that defines an affine parameter of the clock's worldline. The unit of time is the atomic SI second defined as 9, 192, 631, 770 cycles of the radiation corresponding to the ground state hyperfine transition of Cesium 133 [Göbel *et al.*, 2006];
2) the ratio of the readings of two ideal physical clocks moving together along the same timelike worldline, is constant, and does not depend on the choice of the worldline;
3) physical speed of light propagating in vacuum is constant $c_0 = 299792458$ m/s. It is equal to the invariant speed c of the Minkowski spacetime of special relativity: $c_0 = c$.

First two postulates define the operational concept of the ideal clocks in general relativity. The third postulate replaces the rigid-rod measurement of length between two points with its operational measurement of the round-trip time of light traveled between the two points, because the length in general relativity can be invariantly defined in terms of the product of the invariant speed c and the proper time measured by the ideal clocks. Notice that the third postulate is an operational definition of the invariant speed c which appears in special and general relativity as a mathematical constant in the definition of interval ds of the pseudo-Euclidean space that makes the dimension of time coordinate the same as the dimension of space. The speed of light, c_0, is a physical realization of this mathematical constant [63]. With these postulates, which are a subject of continuous experimental monitoring by the national time laboratories and other hi-tech experimental groups, one can work out theoretical relationships between the metric tensor, coordinates, and the operational definition of observables.

3.8.6.2 *Time Measurements*

One of the most important observable in gravitational physics is the proper time. Let us introduce coordinates $x^\alpha = (x^0, x^i) = (ct, \boldsymbol{x})$, and consider an observer with clocks who moves along an arbitrary timelike worldline, $x^\alpha = x^\alpha(t)$, parameterized with the

predictions and to interpret results of the physical measurements of those predictions, is called complete [Carrier, 1994].

63) It is worthwhile to remind once again that it is the invariant speed c which appears in all relativistic equations, not the physical speed of light c_0 [Ellis and Uzan, 2005; Kopeikin and Fomalont, 2006].

coordinate time t. The proper time of the observer's clock relates to the metric tensor $g_{\alpha\beta} = g_{\alpha\beta}(t, \boldsymbol{x})$, by equation

$$-c^2 d\tau^2 = g_{\alpha\beta} dx^\alpha dx^\beta \,, \tag{3.272}$$

wherefrom,

$$d\tau = \sqrt{-g_{00} - 2g_{0i}\beta^i - g_{ij}\beta^i\beta^j}\, dt \,, \tag{3.273}$$

with $\beta^i = v^i/c$, $v^i = dx^i/dt$ - the coordinate velocity of the observer and $g_{\alpha\beta} = g_{\alpha\beta}(t, \boldsymbol{x}(t))$ where both the coordinate velocity and the metric tensor are taken at the points of the observer's worldline. Equation (3.273) can be integrated along the worldline of observer to find out how the finite proper time interval depends on the gravitational field and the motion of observer:

$$\tau = \int_{t_0}^{t} \sqrt{-g_{00} - 2g_{0i}\beta^i - g_{ij}\beta^i\beta^j}\, dt \tag{3.274}$$

where t_0 is the initial epoch (constant of integration).

Equations (3.273) and (3.274) can be simplified in case of a weak field and slow motion approximation. In this case, the metric tensor is expanded with respect to the Minkowski metric in accordance with equation (3.233) with $h_{\alpha\beta} \ll 1$, and velocity $v \ll c$. It gives the following approximation of equation (3.273)

$$d\tau = \sqrt{1 - \boldsymbol{\beta}^2 - h_{\mu\nu}\beta^\mu\beta^\nu}\, dt \,, \tag{3.275}$$

where $\beta^\mu = c^{-1}dx^\alpha/dt = (1, \beta^i)$ is the coordinate four-velocity of observer referred to the coordinate time t. Integrated interval of the proper time

$$\tau = 1 - \frac{1}{2}\int_{t_0}^{t} \left(\boldsymbol{\beta}^2 + h_{\mu\nu}\beta^\mu\beta^\nu\right) dt \,. \tag{3.276}$$

Taking into account equation (3.244) for h_{00}, one can re-write equation (3.276) in the first approximation

$$\tau = t - \frac{1}{c^2}\int_{t_0}^{t} \left(\frac{1}{2}\boldsymbol{v}^2 + U\right) dt \,. \tag{3.277}$$

This coincides with equation (3.246) derived previously. Higher-order terms can be easily included to the result (3.277) after taking more terms in expansion of the right side of equation (3.273).

Notice that in general relativity the equations describing the proper time of a moving clock are expressed solely in terms of the instantaneous value of its velocity and the potentials of the gravitational field. It does not depend on acceleration and/or derivatives of the gravitational field as a consequence of the first clock postulate. Concerning this postulate one should notice that any real atomic clock has finite size and is made of materials having finite stiffness. For this reason, it is clear that the clock is to respond to a sufficiently large acceleration and/or tidal gravitational force that can affect its rate, and even destroy the clock. The first clock postulate is not about the composite mechanical system called clock, but a statement about behavior of elementary atomic processes like the frequency of quantum transition between two energy levels in atom. The fact that atomic processes are not affected by acceleration has been verified experimentally by Pound and Rebka [1960] who measured the thermal dependence of the fractional frequency shift for samples of ^{57}Fe. They confirmed an excellent agreement with the proper time delay equation up to extraordinarily high precision excluding any dependence on acceleration being as much as 10^17 m/s^2.

3.8.6.3 ***Space Measurements***

Proper time and light signals are used in order to introduce another observable quantity - radar distance ϱ. This is the most precise method of measuring astronomical distances in the solar system. It relays upon measurement of observer's proper time and the postulate on the constancy of the speed of light. There are two other methods to measure distances in astronomy - photometry and parallax [Weinberg, 1972]. They are used in cosmology and galactic astronomy and are not crucial for the ongoing discussion (see sections 7.1 and 7.6 for further details).

Let observer and particle move along worldlines $x^i_o = x^i_o(t)$ and $x^i_p = x^i_p(t)$ in a given coordinate chart $x^\alpha = (ct, x^i)$. Let observer send a light signal to the particle at the proper time instant τ_1. It reaches the particle when observer's clock shows proper time τ_2. After reaching the particle, light signal is immediately sent back, e.g. by means of a mirror retro-reflector or a radio transponder, and arrives to the observer at proper time τ_3 as depicted in Figure 3.10. The round-trip time taken by light to travel, is $\tau_3 - \tau_1$. The radar distance, ϱ, between observer and the particle is defined as a product of the invariant speed c and a half of the round-trip time:

$$\varrho = \frac{1}{2}c\,(\tau_3 - \tau_1)\ . \tag{3.278}$$

This definition is invariant and does not depend on the choice of coordinates because the invariant speed c is a fundamental constant, and the proper time of observer is a scalar integral (3.274) having a numerical value which is not affected by coordinate transformations. In practice, the invariant speed c is substituted with the speed of light c_0 in vacuum [64] The instant of the proper time of particle's clock

$$\tau_2 = \tau_1 + \frac{1}{2}c\,(\tau_3 - \tau_1) = \frac{1}{2}\,(\tau_1 + \tau_3)\ , \tag{3.279}$$

is accepted to be simultaneous with the time of reflection τ_2^* measured by observer's clock: $\tau_2^* = \tau_2$. This operational convention is called the Einstein synchronization of clock. It has the same definition as in special relativity since one has no any other means but electromagnetic signals to perform synchronization of clocks in curved spacetime.

64) In the case when light propagates in a medium, its impact on the light propagation should be taken into account and subtracted from the round-trip time of light [Lorimer and Kramer, 2004; Yakovlev, 2002].

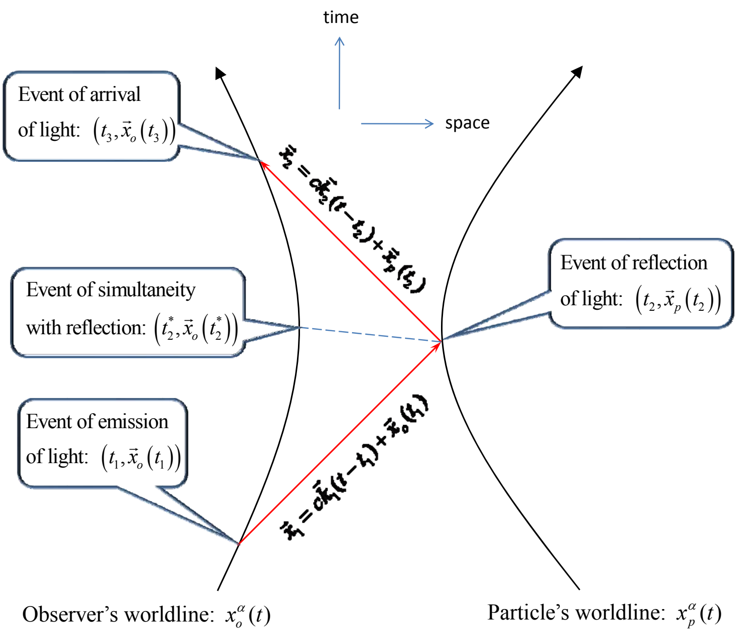

Figure 3.10 Observer sends a light pulse to a particle at instant t_1 of coordinate time. It reaches the particle at observer's coordinate time t_2^*. After reaching the particle, light is immediately sent back and arrives to the observer at time t_3. Einstein's synchronization convention postulates the reflection time, t_2, is simultaneous with time $t_2^* = (t_1 + t_3)/2$ on observer's worldline, that is, $t_2 = t_2^*$, by definition. In practice, the proper time $\tau = \tau(t)$ of observer is used instead of the coordinate time. The radar distance between observer and particle is defined as $\varrho = c(\tau_3 - \tau_1)/2$, where τ_1 and τ_3 are the times of light's emission by and arrival to observer measured in observer's proper time.

The radar distance, ϱ, is a function of the coordinate distance, $r = |\boldsymbol{x}_p(t_2) - \boldsymbol{x}_o(t_2)|$, between observer and particle taken at the coordinate time of reflection of the signal $t_2 = t_2(\tau_2)$. Relationship between ϱ and r is found by solving the equation of light-ray geodesics, which first integral is simply $ds = 0$, or

$$c^2 g_{00} + 2cg_{0i}\dot{x}^i + g_{ij}\dot{x}^i\dot{x}^j = 0 , \tag{3.280}$$

where $\dot{x}^i \equiv dx^i/dt$ is the coordinate velocity of light (photon) that is not equal to the invariant speed c in curved spacetime. Equation (3.280) must be solved with respect to dx^i/dt, and, then, integrated from the event of emission of light, $x_o^\alpha(t_1) = (ct_1, x_o^i(t_1))$, to the event of its reflection, $x_p^\alpha(t_2) = (ct_2, x_p^i(t_2))$, and back to the arrival event, $x_o^\alpha(t_3) = (ct_3, x_o^i(t_3))$, along the light-ray trajectory. In case of weak-field approximation the metric tensor $g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}$, where $h_{\alpha\beta}$ are the gravitational potentials found from Einstein's equations. In this approximation, equation (3.280) becomes

$$\dot{\boldsymbol{x}}^2 = c^2 - h_{00} - 2ch_{0i}\dot{x}^i - h_{ij}\dot{x}^i\dot{x}^j , \tag{3.281}$$

that clearly shows that gravitational field affects the coordinate speed of light making it different from c. Since the perturbation of the coordinate speed of light from its nominal value, c, is small in the weak-field approximation, equation (3.281) can be solved by iterations with the coordinate speed of light decomposed as

$$\dot{x}^i = c\left(k^i + \dot{\Xi}\right) , \tag{3.282}$$

where k^i is the constant unit vector along the unperturbed trajectory of light propagation, $\boldsymbol{k}\cdot\boldsymbol{k} = \delta_{ij}k^ik^j = 1$, and $\dot{\Xi}^i = d\Xi^i/dt$ is the perturbation of the coordinate velocity of light.

In case of the round trip of light signal, the integration will yield two equations relating the coordinate time differences:

$$c(t_2 - t_1) = r_1 + \Delta T_{12} , \tag{3.283a}$$
$$c(t_3 - t_2) = r_3 + \Delta T_{23} , \tag{3.283b}$$

where the coordinate distances are $r_1 = |\boldsymbol{x}_o(t_1) - \boldsymbol{x}_p(t_2)|$, $r_3 = |\boldsymbol{x}_o(t_3) - \boldsymbol{x}_p(t_2)|$, and functions

$$\Delta T_{12} = \frac{c}{2}\int_{t_1}^{t_2} h_{\mu\nu}\left[t, \boldsymbol{x}_1(t)\right] k_1^\mu k_1^\nu dt , \tag{3.284a}$$
$$\Delta T_{23} = \frac{c}{2}\int_{t_2}^{t_3} h_{\mu\nu}\left[t, \boldsymbol{x}_2(t)\right] k_2^\mu k_2^\nu dt , \tag{3.284b}$$

are calculated along the unperturbed light-ray paths,

$$\boldsymbol{x}_1 = c\boldsymbol{k}_1(t - t_1) + \boldsymbol{x}_o(t_1) , \tag{3.285a}$$
$$\boldsymbol{x}_2 = c\boldsymbol{k}_2(t - t_2) + \boldsymbol{x}_p(t_2) , \tag{3.285b}$$

that are defined by null vectors $k_1^\alpha = (1, k_1^i)$ and $k_2^\alpha = (1, k_2^i)$ on the background Minkowski spacetime (see Figure 3.10). Function ΔT is a complicated integral from the gravitational potentials $h_{\alpha\beta}$ calculated along the light-ray trajectory. It was introduced to general relativity by Shapiro [1964]. The Shapiro time delay was measured in various time-delay experiments in the solar system to the fractional precision 10^{-4} and better [Bertotti *et al.*, 2003; Fomalont *et al.*, 2009a; Lambert and Le Poncin-Lafitte, 2009; Shapiro *et al.*, 2004]. Precise mathematical details of calculation of the relativistic time delay and its applications in astrometry are given in section 7.

The radar distance ϱ between observer and particle is calculated by adding up equations (3.283) and converting the coordinate time difference, $t_3 - t_1$, to the difference $\tau_3 - \tau_1$ between the proper times with the help of equation (3.274) or, if the first approximation is sufficient, with the help of equation (3.277). Calculation of the radar distance by means of equation (3.278) yields

$$\varrho = r + \Delta C + \Delta T \ , \tag{3.286}$$

where

$$r = \frac{1}{2}\left(r_1 + r_3\right) \ , \tag{3.287}$$

is the coordinate distance,

$$\Delta C = -\frac{c}{4}\int_{t_1}^{t_3}\left\{\boldsymbol{\beta}^2 + h_{\mu\nu}\left[t, \boldsymbol{x}_o(t)\right]\beta^\mu\beta^\nu\right\}dt \ , \tag{3.288}$$

is the relativistic time delay of observer's clock, calculated along the observer's worldline $\boldsymbol{x}_o(t)$, and

$$\Delta T = \frac{c}{4}\left\{\int_{t_1}^{t_2} h_{\mu\nu}\left[t, \boldsymbol{x}_1(t)\right]k_1^\mu k_1^\nu dt + \int_{t_2}^{t_3} h_{\mu\nu}\left[t, \boldsymbol{x}_2(t)\right]k_2^\mu k_2^\nu dt\right\} \ , \tag{3.289}$$

is the Shapiro time delay due to the propagation of light through the gravitational field.

3.8.6.4 ***Are coordinates measurable?***

In Newtonian mechanics and also in special relativity inertial coordinates are considered as measurable quantities. General relativity and any other valid theory of gravity dismantled coordinates as observable quantities. Indeed, any self-consistent theory of gravity obeys the general principle of covariance which states that the description of physics is independent of one's choice of coordinates. Coordinates are merely bookkeeping parameters taking a series of values of mathematical functions that are used in order to numerate events and their sequences taking place in the curved spacetime manifold. As such, coordinates are intangible things bearing a subjective imprint of an objective physical reality. Two different coordinate charts are completely equivalent from an objective standpoint. Any preference in choosing particular coordinates is exclusively associated with a pure mathematical convenience for solving equations. Hence, coordinates are *unobservable variables* that *can never* be measured in any kind of spacetime experiment nor in astronomical practice. Observations and experiments can only provide us with a series of objective relationships between astronomical bodies (particles) and fields described as geometric objects residing in spacetime manifold.

Let us discuss this effacing property of coordinates in case of the measurement of the radar distance ϱ. Equation (3.286) establishes a functional relationship between the measured quantity ϱ and the objective geometric objects - observer's worldline, particle's worldline, light-ray trajectory, and the metric tensor $g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}$. One has chosen some particular coordinates $x^\alpha = (x^0, x^i) = (ct, x^i)$ to present mathematical description of these geometric objects and to establish relationship between them and the measured quantity ϱ. Let us now choose new coordinates, $x'^\alpha = (x'^0, x'^i) = (ct', x'^i)$, where t' is a new coordinate time, and x'^i are new spatial coordinates related to the old coordinates x^α by the transformation equation $x'^\alpha = x'^\alpha(x)$, where x in the argument of functions denotes four-dimensional coordinates: $f(x) \equiv f(x^\alpha) = f(t, \boldsymbol{x})$. Subsequent calculations can be done exactly but, since the above-given calculation of ϱ has been done only in the first approximation with respect to the potentials of gravitational field $h_{\alpha\beta}$ and up to the terms being quadratic with respect to velocity of observer, it is reasonable to consider the coordinate transformation in the same approximation discarding higher-order terms

$$x'^\alpha = x^\alpha + \xi^a(x) , \tag{3.290}$$

where ξ^α are the gauge functions of the same order as $h_{\alpha\beta}$ and $\boldsymbol{\beta}^2$. Metric tensor in the new coordinates, $g'_{\alpha\beta}(x') = \eta_{\alpha\beta} + h'_{\alpha\beta}(x')$, relates to the metric tensor, $g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}(x)$, in the old coordinates by transformation (3.264). In the linearized approximation this transformation reads

$$h_{\alpha\beta}(x) = h'_{\alpha\beta}(x') + \xi_{\alpha,\beta}(x) + \xi_{\beta,\alpha}(x) , \tag{3.291}$$

where indices are raised and lowered with the help of the Minkowski metric $\xi_\alpha = \eta_{\alpha\beta}\xi^\beta$, and comma denotes a partial derivative with respect to the old coordinates, $\xi_{\alpha,\beta} \equiv \partial_\beta \xi_\alpha$, etc.

The coordinate transformations (3.290) lead to the corresponding point-like transformation of time and spatial coordinates of observer and the particle

$$t'_1 = t_1 + \xi^0_1 , \qquad x'^i_o(t'_1) = x^i_o(t_1) + \xi^i_1 , \tag{3.292a}$$

$$t'_2 = t_2 + \xi^0_2 , \qquad x'^i_p(t'_2) = x^i_p(t_2) + \xi^i_2 , \tag{3.292b}$$

$$t'_3 = t_3 + \xi^0_3 , \qquad x'^i_o(t'_3) = x^i_o(t_3) + \xi^i_3 , \tag{3.292c}$$

where $\xi^\alpha_1 = \xi^\alpha\left[t_1, \boldsymbol{x}_o(t_1)\right]$, $\xi^\alpha_2 = \xi^\alpha\left[t_2, \boldsymbol{x}_p(t_2)\right]$, $\xi^\alpha_3 = \xi^\alpha\left[t_3, \boldsymbol{x}_o(t_3)\right]$, are the gauge functions taken at the point of emission of light, its reflection, and arrival to observer, respectively.

Other terms entering relationship (3.286), are transformed as follows:

$$r_1 = r_1' + k_1^i\left(\xi_1^i - \xi_2^i\right) , \tag{3.293a}$$
$$r_3 = r_3' + k_2^i\left(\xi_2^i - \xi_3^i\right) , \tag{3.293b}$$
$$\beta^i = \beta'^i - \beta^\alpha \partial_\alpha \xi^i , \tag{3.293c}$$
$$\boldsymbol{\beta}^2 = \boldsymbol{\beta}'^2 - 2\beta^\mu\beta^\nu\xi_{\nu,\mu} - 2\beta^\mu\partial_\mu\xi^0 , \tag{3.293d}$$
$$h_{\mu\nu}\beta^\mu\beta^\nu = h'_{\mu\nu}\beta'^\mu\beta'^\nu + 2\beta^\mu\beta^\nu\xi_{\nu,\mu} , \tag{3.293e}$$
$$h_{\mu\nu}k^\mu k^\nu = h'_{\mu\nu}k'^\mu k'^\nu + 2k^\mu k^\nu\xi_{\nu,\mu} , \tag{3.293f}$$

where the quantities with sub-indices 1,2,3 are taken at the events of emission of light, the point of its reflection, and the point of its arrival respectively; $r_1' = |\boldsymbol{r}'_1|$, $r_3' = |\boldsymbol{r}'_3|$, $\boldsymbol{r}_1' = \boldsymbol{x}_p'(t_2') - \boldsymbol{x}_o'(t_1')$, $\boldsymbol{r}_3' = \boldsymbol{x}_o'(t_3') - \boldsymbol{x}_p'(t_2')$ are coordinate distances between the corresponding events expressed in the new coordinates; $k'^i_1 = r'^i_1/r'_1$ and $k'^i_2 = r'^i_3/r'_3$ are the unit vectors defining propagation of light in the new coordinates; the gauge functions ξ^α in equations (3.293c)-(3.293e) are taken along the worldline of observer; the gauge functions ξ^α in equations (3.293f) are taken along the light-ray path.

Substituting these equations to the right side of relationship (3.286), and calculating integrals

$$\int_{t_1}^{t_3}\beta^\mu\partial_\mu\xi^0 dt = \int_{t_1}^{t_3}\frac{d\xi^0}{dt}dt = \xi_3^0 - \xi_1^0 , \tag{3.294a}$$
$$\int_{t_1}^{t_2}k_1^\mu k_1^\nu\xi_{\nu,\mu}dt = \int_{t_1}^{t_2}\frac{d}{dt}\left(k_1^\mu\xi_\mu\right)dt = \xi_1^0 - \xi_2^0 + k_1^i\left(\xi_2^i - \xi_1^i\right) , \tag{3.294b}$$
$$\int_{t_2}^{t_3}k_2^\mu k_2^\nu\xi_{\nu,\mu}dt = \int_{t_2}^{t_3}\frac{d}{dt}\left(k_2^\mu\xi_\mu\right)dt = \xi_2^0 - \xi_3^0 + k_2^i\left(\xi_3^i - \xi_2^i\right) , \tag{3.294c}$$

one finds out that in the approximation under consideration, the radar distance, ϱ, calculated in the new (primed) coordinates, is given by the relationship

$$\varrho = r' + \Delta' C + \Delta' T , \tag{3.295}$$

where

$$r' = \frac{1}{2}\left(r_1' + r_3'\right) , \tag{3.296a}$$
$$\Delta' C = -\frac{c}{4}\int_{t_1'}^{t_3'}\left\{\boldsymbol{\beta}'^2 + h'_{\mu\nu}\left[t', \boldsymbol{x}_o'(t')\right]\beta'^\mu\beta'^\nu\right\}dt' , \tag{3.296b}$$
$$\Delta' T = \frac{c}{4}\int_{t_1'}^{t_2'}h'_{\mu\nu}\left[t', \boldsymbol{x}_1'(t')\right]k'^\mu_1 k'^\nu_1 dt' \tag{3.296c}$$
$$+\frac{c}{4}\int_{t_2'}^{t_3'}h'_{\mu\nu}\left[t', \boldsymbol{x}_2'(t')\right]k'^\mu_2 k'^\nu_2 dt' . \tag{3.296d}$$

Relationship (3.295) in the new coordinates has exactly the same functional form as its counterpart (3.286) expressed in the old coordinates. It proves that the measured quantity ϱ does not provide by itself any information about the coordinates used for its theoretical calculation. Only the relationship between the separate terms entering equation (3.295) remains the same in any coordinate chart. An observer can only *derive* coordinates of events that occur in the spacetime manifold. This is the main function of coordinates and any coordinates accomplish this task. Whether or not these coordinates conform to any particular criteria is immaterial from a principal standpoint. In practice, however, it is found that, when one actually tries to analyze astronomical measurements made by an arbitrary observer, the mathematics become progressively difficult to handle in arbitrary coordinates. What kind of difficulties one mets and how to handle them, is a subject of a dedicated theoretical study having been undertaken by Synge [1964] and called him *chronogeometry*. Synge's relativistic theory of astronomical measurements is based on the world function, Ω, that is a measure of distance between two arbitrary points on a curved spacetime manifold connected by a geodesic. For two events separated by null geodesic the world function Ω is similar to the concept of the radar distance ϱ. Description of measurements in terms of the world function is invariant but involves many technicalities which are alien to the procedures and protocols of the data reduction algorithms used by astronomers. Although modern astronomical observations can measure various relativistic effects with an unparallel precision, the practice is to keep the data processing algorithms as close to the classic conceptions of space and time as possible. Thus, coordinate-based description of astronomical phenomena is maintained in the form of astronomical catalogues of stars and quasars, ephemerides of planets and their satellites, the Earth orientation parameters representing coordinates of the Earth rotational axis, etc. Such an approach simplifies mathematical formalism and allows us to record astronomical data continuously. This introduction of coordinates to the algorithms of astronomical data reduction should not be misinterpreted. The coordinates and coordinate-dependent parameters like semi-major axis and eccentricity of osculating orbit, can be determined from observations only after one fixes the coordinate conditions and remove the residual gauge freedom in the metric tensor that is solution of the gravitational field equations. As soon as the gauge or coordinate parametrization of the orbit change, the "determined" values of the coordinates and coordinate parameters change accordingly (see, for example, the paper by Klioner and Kopeikin [1994] demonstrating the dependence of the "observed" coordinate parameters on the choice of parametrization in the case of binary pulsar dynamics).

3.8.7
Experimental testing of general relativity

Each physical theory must have a predictive power and deals with measurements which goals is to test the predictions and to decide whether the theory is right that is, if it corresponds to reality, or wrong. The concept of realism of a physical theory was introduced by Einstein *et al.* [1935]. They define the *element of physical reality* as a physical quantity associated with a certain mathematical object within the theory whose observed value can be predicted by the theory with certainty before measuring it or disturbing it in any way. Einstein *et al.* [1935] also defined a complete physical theory as one in which every element of physical reality is accounted for by the mathematical formalism of the theory. Given a complete physical theory, one can predict the outcome of physical measurements made on any material system described by the theory. In addition, such a theory should also be capable of describing the behavior of the measuring devices used to observe a system without recourse to additional assumptions taking us outside of the realm of the theory.

General relativity is a theory of gravitational field that is an intrinsic element of the geometric structure of a pseudo-Riemannian spacetime manifold. This structure is the metric tensor, $g_{\alpha\beta}$, that plays a double role in general relativity as it defines both the metrical relationships between the elements of physical reality and determines evolution of gravitational field from the initial data. Hence, measurement of gravitational field in general relativity is the same thing as measurement of the metrical relationship between events taking place on the spacetime manifold.

Let us consider testing general relativity by taking measurement of the radar distance ϱ defined by equation (3.286) as an example. For concreteness, one can think about the radar distance between Earth and Moon measured by lunar laser ranging (LLR) technique [Battat *et al.*, 2007; Dickey *et al.*, 1994; Kopeikin *et al.*, 2008]. As time passes on, one can make sufficiently large number of measurements of the radar distance ϱ which will allow us to form a system of equations for determination of the functional relationship between elements of the physical reality presented in the right side of equation (3.286). These elements are the worldlines of observer (Earth), particle (Moon), photon (laser pulse) as well as the metric tensor. Metric tensor is determined by solving the Einstein equations. The worldlines are determined from equations of motion, which are consequences of the Einstein equations due to the Bianchi identity. Their solution establishes a functional correspondence between gravitational field and the worldlines which is unique for each gravity theory. Substituting solutions for the worldlines and for the metric tensor to the right side of equation (3.286) allows us to make a theoretical prediction of the radar distance, $\varrho_{\rm c}(\vartheta,\tau)$, as a function of the proper time τ of observer and parameters, $\vartheta = \{\vartheta_1, \vartheta_2, ..., \vartheta_M\}$, of the theoretical model. Comparing this prediction with observed value of the radar distance, $\varrho_{\rm o}(\tau)$, at each observational point, determines parameters of the model. As a rule, this is done by minimizing functional [65]

$$\chi^2 = \sum_{i=1}^{N} \left[\varrho_{\rm c}(\vartheta,\tau_i) - \varrho_{\rm o}(\tau_i)\right]^2 , \tag{3.297}$$

by the least square method [Bevington and Robinson, 1992; Gubanov, 1997] for a series of observations taken at times $\tau_i = \{\tau_1, \tau_2, ..., \tau_N\}$.

Number of the parameters ϑ describing the theoretical model is usually pretty large and can vary from several hundred to thousands, and even more. Many of these parameters originate from the Newtonian physics (orbital parameters, multipole moments of gravitational field, parameters of elasticity of the planets, tides, etc.) and

65) In practice, there are different approaches to determination of parameters but among them, the least squares method is the most common.

technical requirements for observing system (clocks, lasers, mounting system, retro-reflectors, antennas, etc.). Only few parameters remain to test the consistency of general relativity. Standard tests within the solar system are usually based on the parameterized post-Newtonian (PPN) formalism [Will, 1993] which introduces to the metric tensor a set of ten parameters, $\vartheta_{\rm PPN} = \{\gamma,\beta,\xi,\alpha_1,\alpha_2,\alpha_3,\zeta_1,\zeta_2,\zeta_3,\zeta_4\}$ in such a way that competing gravitational theories can be accommodated [66]. PPN formalism fixes the gauge freedom of the metric by postulating a special form of some of the metric tensor coefficients. In this standard PPN gauge (also known as the Chandrasekhar-Nutku gauge [Chandrasekhar and Nutku, 1969; Schäfer, 1982]) the PPN parameters are given the following significance [Will, 1993]:

- γ describes how much space-curvature is produced by unit rest mass;
- β describes how much nonlinearity is there in the superposition law for gravity;
- ξ describes if there are preferred-location effects associated with gravity field;
- $\alpha_1,\alpha_2,\alpha_3$ describe if there are preferred frame effects associated with motion of the solar system with respect to a hypothetical privileged reference frame of the universe violating the Lorentz invariance;
- $\zeta_1,\zeta_2,\zeta_3,\zeta_4$ describe if there exists violation of conservation of total momentum of the solar system.

In fact, there are three other PPN parameters which are the components of the velocity of motion of the solar system with respect to the privileged frame of the universe: $\boldsymbol{w} = (w^i) = (w^1,w^2,w^3)$. These parameters are never fit but are always fixed by assuming that the privileged frame of the universe is that in which the cosmic microwave background radiation (CMBR) is isotropic. The reason for fixing the preferred-frame velocity $\boldsymbol{w}$ is that preferred-frame parameters are always enter PPN formalism in the form of products: $\alpha_1\boldsymbol{w}$, $\alpha_2\boldsymbol{w}$, $\alpha_3\boldsymbol{w}$. Hence, the tight experimental limitations currently given on the preferred-frame parameters (see [Will, 2006]) strongly depend on one's guess about the velocity $\boldsymbol{w}$ of the solar system relative to a "preferred frame" which was assumed to be the CMBR. But, one does not know if it defines the preferred coordinate system in the sense forbidden by special relativity; it is just a convenient frame to describe the global isotropy of the cosmological black-body radiation. Furthermore, future observations of the relic gravitational wave background (GWB), for example, may lead to the frame somehow moving with respect to the CMBR frame. Since the GWB was formed in the very early universe, long before the CMBR decoupled from matter, it would be more likely to associate it with the preferred frame. However, neither the global topological isotropy of CMBR nor GWB may be related to the local Lorentz-invariant isotropy of the spacetime itself. Modern multi-connected cosmological models of the universe [Barrow and Levin, 2001, 2003] also indicate that various possibilities for the global preferred frame may not coincide with the CMBR and GWB frames. Recently, Bailey and Kostelecky [2008] have proposed the gravitational Standard-Model Extension (SME) as a covariant alternative to the PPN formalism in using gravitational experi-

66) The best-known competing theory is the Brans-Dicke theory of gravity that will be discussed in section 4.

ments to test the spontaneous Lorentz violation in the gravity sector of relativity.

PPN formalism is valid only in the standard PPN gauge and should not be interpreted as covariant statements. Misunderstandings can arise if one takes no heed of this caution. For example, there was some discussion about physical significance of β parameter as a true marker of the non-linear gravitational effects. Specifically, Deser and Laurent [1973] pointed out that the non-linear gravity effects interfere with the choice of coordinates. This led Karlhede [1982] to claim that gravitational nonlinearities are not observable in the post-Newtonian approximation. This claim is erroneous since the coordinate transformation do removes the PPN parameter β from g_{00} component of the metric tensor but re-introduces it to the space-space, g_{ij}, components so that β remains among the PPN parameters entering the observable effects like the shift of Mercury's perihelion.

A distinctive feature of PPN formalism is that it postulates a complete compatibility of the parameterized metric tensor with the affine connection. Hence, PPN formalism is not capable to handle alternative theories of gravity which are not fully metric-based and include additional pieces of the connection - torsion and nonmetricity. Several affine-metric theories have been proposed to fill up this gap in respond to various yet unexplained phenomena within the solar system, in galactic astronomy, and in cosmology [Altschul *et al.*, 2010; Bekenstein, 2004; Blome *et al.*, 2010; Heinicke *et al.*, 2005; Moffat, 1995]. To answer the question on how the metric tensor is coupled to the affine connection, one need to introduce additional PPN parameters tracing the presence of torsion [Mao *et al.*, 2007] and nonmetricity in the affine connection [Kopeikin, 2004; Kopeikin and Fomalont, 2007].

3.9
Variational Principle in General Relativity

3.9.1
The action functional

In physics, action is an attribute of the dynamics of a physical system. It is a functional, S, that takes a function of time (for particles) and space (for fields) as input and returns a scalar. The action can be found for any physical field. However, in many cases this is not a straightforward task and its solution requires elaboration on certain assumptions and justifications. General relativity is a theory of gravitational field described by the metric tensor, $g_{\alpha\beta}$, which is considered as a dynamic variable in gravity sector of the theory. Therefore, the action of the gravitational field, S_H, is to be build out of the metric $g_{\alpha\beta}$ and its partial derivatives $g_{\alpha\beta,\gamma}$ [67]. No other variables are supposed to appear in the gravitational action.

Matter is the source of gravitational field, and one has to find out the action for matter, S_M, that should depend on a set of matter variables, ψ, and their partial derivatives, $\nabla_\alpha\psi$. Each of the set of matter variables is a function on spacetime manifold and possesses indices to indicate its tensorial or spinorial character but they have been omitted to simplify notations. As the action is invariant, it suggests that the derivatives of the dynamic variables must be covariant in the most general case. Moreover, the action of matter is usually a scalar quadratic function build out of the dynamic variables and their first derivatives [68] It implies that the metric tensor and its first derivatives should participate in the definition of the action of matter.

In general relativity, the overall action S for a physical system consisting of matter interacting with gravitational field, is a linear superposition of two terms [69]:

$$S = S_H + S_M\,, \tag{3.298}$$

which is a consequence of the *principle of minimal coupling* of gravity with matter. This principle establishes the simplest form of the interaction term between gravity and matter governed by the principle of equivalence, according to which any physical equation of special relativity can be turned into its general-relativistic counterpart by replacing the Minkowski metric, $\eta_{\alpha\beta}$, with the relevant metric of spacetime, $g_{\alpha\beta}$, and by replacing any partial derivative, ∂_α, with a covariant one, ∇. The minimal coupling of matter to gravity leads to a natural absorption of the gravity field variable (metric) and the affine connection to the structure of the Lagrangian S_M which original definition usually comes from special relativity.

67) Covariant derivatives of the metric tensor are identically nil according to the basic principles of general relativity discussed in section 3.8.2. Thus, covariant derivatives of the metric do not appear in the action.

68) Higher derivatives are also allowed, for example, in relativistic dynamics of N-body system emitting gravitational waves [Damour, 1987; Damour *et al.*, 1989; Grishchuk and Kopeikin, 1986].

69) Matter fields are localizable, and can be measured at one point while gravitational field is not localizable due to the principle of equivalence. Hence, one can not have cross-coupling terms in the action consisting of matter variables and the Christoffel symbols, which are not a part of the covariant derivative.

The action is defined as an integral from some scalar function depending on gravitational field and matter variables, taken over the entire evolutionary history of the physical system. The scalar function is usually denoted, L, and is called *Lagrangian* that is defined on the entire spacetime manifold. Because the integration must give us a scalar - the action S, the volume of integration over the manifold must be a scalar that is invariant with respect to coordinate transformations. To establish the invariant measure of integration in curved spacetime, let us consider two coordinate charts, x^α and x'^α, connected by an invertible coordinate transformation, $x^\alpha = x^\alpha(x')$. The coordinate four-volumes, $d^4x = dx^0dx^1dx^2dx^3$ and $d^4x' = dx'^0dx'^1dx'^2dx'^3$, are related to each other by the determinant of the matrix of the infinitesimal transformation, $J = \det\left[\partial x'^a/\partial x^\beta\right]$, that is called the Jacobian

$$d^4x' = Jd^4x \,. \tag{3.299}$$

The Jacobian, J, can be expressed in terms of the determinant of the metric tensor after calculating determinants from both sides of the metric-tensor transformation equation

$$g_{\alpha\beta}(x) = g'_{\mu\nu}(x')\frac{\partial x'^\mu}{\partial x^\alpha}\frac{\partial x'^\nu}{\partial x^\beta} \,. \tag{3.300}$$

Applying the rule of calculation of determinants from the product of matrices, yields

$$g = J^2 g' \,, \tag{3.301}$$

where $g = \det\left[g_{\alpha\beta}(x)\right]$ and $g' = \det\left[g'_{\alpha\beta}(x')\right]$. One has to take care about the sign of the metric tensor determinant in extracting the root square from both sides of the above equation. Assuming that the primed coordinates, x'^α, are the normal Riemann coordinates, the determinant, $g' = \det\left[\eta_{\alpha\beta}\right] = -1$, at the origin of the normal coordinates. Because the signature of the matrix determinant is invariant under diffeomorphic coordinate transformations, one comes to the conclusion that determinant of the metric tensor is always negative, $g < 0$ and $g' < 0$. Accounting for this fact, extracting the root square from equation (3.301), and substituting the result to equation (3.299), brings about the equivalence

$$\sqrt{-g'}d^4x' = \sqrt{-g}d^4x \,. \tag{3.302}$$

It tells us that the invariant measure of a volume on spacetime manifold is, $\sqrt{-g}d^4x$.

In most physical theories the Lagrangian does not include derivatives of the dynamic variables of the order higher than first. However, it is impossible to find the Lagrangian of gravitational field, that is a scalar and is built only from the metric and its first derivatives - the Christoffel symbols. This is because the Christoffel symbols can be locally eliminated by an appropriate choice of a coordinate chart. Hence, if such an invariant Lagrangian existed it could be only a trivial constant. Thus, the second derivatives of the metric tensor must be included to the gravitational Lagrangian to build a scalar invariant from the Riemann tensor. It is remarkable that in general relativity the second derivatives of the metric tensor entering the gravitational Lagrangian, L_H, do not lead to gravitational field equations of the order higher than second. This is because all terms with the second derivatives of the metric tensor are grouped to a single term which is a four-dimensional divergence of a vector field. Due to the Stock's theorem it reduces to a surface term on the boundary of integration that does not contribute to the variational equations of the gravity field. It does not mean, however, that the term with the second derivatives from the metric tensor should be discarded. Though it does not contribute to the field equations, it remains essential in definition of the invariant gravitational Lagrangian, and appears later in the theory of the (pseudo)tensor of energy-momentum of gravitational field and its conserved quantities [Babak and Grishchuk, 2000; Grishchuk *et al.*, 1984; Popova and Petrov, 1988; Szabados, 2009].

All in all, one has the functionals of action of matter and gravitational field written down schematically as

$$S_M = \int L_M\left(\psi; \nabla_\alpha\psi; g_{\alpha\beta}\right)\sqrt{-g}d^4x\,, \tag{3.303}$$

$$S_H = \int L_H\left(g_{\alpha\beta}; g_{\alpha\beta,\gamma}; g_{\alpha\beta,\gamma\delta}\right)\sqrt{-g}d^4x\,, \tag{3.304}$$

where both, L_H and L_M, are invariant scalars, ψ denotes a multiplet of matter variables of various tensor ranks, comma denotes partial derivatives with respect to a corresponding coordinate, for example, $g_{\alpha\beta,\gamma} = \partial_\gamma g_{\alpha\beta}$, $g_{\alpha\beta,\gamma\delta} = \partial_\gamma\partial_\delta g_{\alpha\beta} = \partial_{\gamma\delta} g_{\alpha\beta}$, and $\nabla_\alpha\psi$ is a covariant derivative from the matter variable. It will be convenient to denote it as follows

$$\nabla_\alpha\psi \equiv \psi_{,\alpha} + \Gamma_\alpha\psi\,, \tag{3.305}$$

where comma denotes a partial derivatives from the component of ψ, and Γ_α consists of the Christoffel symbols. If the matter field is a tensor of a (p, q) type, $\psi \equiv \psi^{\beta_1\ldots\beta_p}_{\gamma_1\ldots\gamma_q}$, then,

$$\begin{aligned}\Gamma_\alpha\psi \quad &\equiv \quad \Gamma^{\beta_1}_{\alpha\mu}\psi^{\mu\beta_2\ldots\beta_p}_{\gamma_1\gamma_2\ldots\gamma_q} + \ldots + \Gamma^{\beta_p}_{\alpha\mu}\psi^{\beta_1\ldots\beta_{p-1}\mu}_{\gamma_1\gamma_2\ldots\gamma_q}\\ &- \quad \Gamma^{\mu}_{\alpha\gamma_1}\psi^{\beta_1\beta_2\ldots\beta_p}_{\mu\gamma_2\ldots\gamma_q} - \ldots - \Gamma^{\mu}_{\alpha\gamma_q}\psi^{\beta_1\beta_2\ldots\beta_p}_{\gamma_1\ldots\gamma_{q-1}\mu}\,.\end{aligned} \tag{3.306}$$

In case of a scalar field - tensor of the type $(0, 0)$, $\Gamma_\alpha = 0$. Definition (3.306) is a direct consequence of equation (3.91) for covariant derivative of a tensor field.

Notice that the covariant derivative enters explicitly only the matter Lagrangian L_M. Metric tensor plays a role of the only gravitational variable in general relativity, and its covariant derivative equals to zero due to the principle of equivalence. For this reason, the dynamic gravitational variables in general relativity are the metric tensor and its *partial* derivatives. We shall specify explicit form of the Lagrangians in next section. Here, one would like to bring attention to the fact that any Lagrangian is defined up to a total covariant divergence from an arbitrary vector field l^α that vanishes on the boundary of integration. Indeed, the replacement

$$L \to L + \nabla_\alpha l^\alpha\,, \tag{3.307}$$

changes the action

$$S \to S + \int \nabla_\alpha l^\alpha \sqrt{-g} d^4 x \,. \tag{3.308}$$

However, the total divergence from a vector field can be always integrated and reduced to a surface integral due to the Gauss theorem

$$\int \nabla_\alpha l^\alpha \sqrt{-g} d^4 x = \oint \sqrt{-g} l^\alpha d\Sigma_\alpha \,, \tag{3.309}$$

where $d\Sigma_\alpha$ is the element of integration on the boundary of the four-dimensional domain. The surface integral in equation (3.309) is nil, as one has assumed the field $l^\alpha = 0$ on the boundary. This simple observation tells us that calculations of variational derivatives are always valid up to a total divergence of a vector field which can be discarded in the most cases.

3.9.2 Variational equations

A variational principle, also called the *principle of least action*, is used in mechanics to obtain the equations of motion for particles and continuous media. It leads to the development of the Lagrangian and Hamiltonian formulations of classical mechanics [Arnold, 1995]. The principle can be also applied to continuous distribution of matter - fields, in order to derive the field equations. The main idea of the principle is that among all configurations of a physical system under consideration, the physically realizable one corresponds to a minimal value of the *action* S. Application of this principle defines the field equations and the equations of motion of matter [70].

3.9.2.1 *Variational equations for matter*

Action, S, given in equations (3.298), depends on mutually-independent dynamic variables, ψ, $g_{\alpha\beta}$, and their derivatives. For each choice of these functions, the action takes a specific numerical value. Smoothly varying the variables one change the numerical value of S. The virtual variation is denoted by $\delta\psi$ for matter variables, and by $\delta g_{\alpha\beta}$ for gravitational field (metric). The variation is defined as an infinitesimal difference between two values of the variable taken on spacetime manifold at the same value of coordinates:

$$\delta\psi \equiv \psi'(x) - \psi(x) \,, \tag{3.310}$$
$$\delta g_{\alpha\beta} \equiv g'_{\alpha\beta}(x) - g_{\alpha\beta}(x) \,, \tag{3.311}$$

where $x \equiv x^\alpha$ denotes four-dimensional coordinates of the point. Primed functions, ψ' and $g'_{\alpha\beta}$, are not related to ψ and $g_{\alpha\beta}$ in the most general case; they are independent functions satisfying some necessary conditions of differentiability. The variations

70) In order to derive the equations of motion and field equations it is sufficient to request that the action is an extremal among all possible variations of the system configuration.

(3.310) and (3.311) may or may not be accompanied by coordinate transformations that do not appear explicitly in these definitions. These variations should be distinguished from the variations which explicitly entail the change of coordinates besides taking the change in functional values,

$$\hat{\delta}\psi \equiv \psi'(x') - \psi(x) , \tag{3.312}$$
$$\hat{\delta}g_{\alpha\beta} \equiv g'_{\alpha\beta}(x') - g_{\alpha\beta}(x) , \tag{3.313}$$

where $x' = x'(x)$ is the transformation from one coordinate chart to another. It should be emphasized that both forms of the variations can be used in variational calculus [Mitskevich, 1969] but those, defined by equations (3.310), (3.311) are more mathematically convenient as they commute with other mathematical operations of tensor calculus like derivatives, contraction of indices, etc. Since the variations, δ and $\hat{\delta}$, are infinitesimally small the link between them can be easily established in the form of the Taylor expansion. We do not use variations (3.312), (3.313) in what follows.

Virtual variation of the dynamic variables may be a real physical perturbation of the physical system under consideration but in some cases they are not. The variations can be also subject to some constrains but one does not consider this case. The variation of variables lead to variation of the action

$$\delta S = S' - S , \tag{3.314}$$

where S' has the same functional form as S in equations (3.303), (3.304), but depends on the primed values of the variables $\psi'(x) = \psi(x)+\delta\psi$, $g'_{\alpha\beta}(x) = g_{\alpha\beta}(x)+\delta g_{\alpha\beta}$, etc. The principle of the least action demands that the physical system evolves along those trajectories of the variables, which conform to equation

$$\delta S = 0 . \tag{3.315}$$

This equation implies that an infinitesimal change of the total action S should vanish if variations of physical fields, ψ, and the metric, $g_{\alpha\beta}$, take the system out of stationary equilibrium. Whether this equilibrium is minimum, or maximum, or a saddle point depends on the behavior of the second variation of action [Arnold, 1995; Dubrovin *et al.*, 1984]. We shall assume that variations of all variables and their derivatives giving rise to variation δS, vanish on the boundary of the domain of integration in definition of action:

$$(\delta\psi)_{\text{boundary}} = (\delta\psi_{,\alpha})_{\text{boundary}} = 0 , \tag{3.316a}$$
$$\left(\delta g_{\alpha\beta}\right)_{\text{boundary}} = \left(\delta g_{\alpha\beta,\gamma}\right)_{\text{boundary}} = 0 . \tag{3.316b}$$

When only the matter field variation is performed, one gets $\delta S = \delta S_M$. Introducing a new notation,

$$\mathcal{L} \equiv \sqrt{-g}L , \tag{3.317}$$

for the Lagrangian density and expanding action S'_M around the unprimed value of the matter variable ψ, yields

$$\delta S_M = \int \left[\frac{\partial \mathcal{L}_M}{\partial \psi} \delta\psi + \frac{\partial \mathcal{L}_M}{\partial \psi_{,\alpha}} \delta\psi_{,\alpha} \right] d^4x \, , \tag{3.318}$$

where $\delta\psi_{,\alpha} \equiv \delta(\partial_a \psi)$ is a variation of the partial derivative of the matter field ψ. According to definition (3.310) of variation, the two mathematical operations - partial derivative and variation - commute, that is [71]

$$\delta(\partial_a \psi) = \partial_\alpha (\delta\psi) \, . \tag{3.319}$$

It allows us to integrate the second term in right side of equation (3.318) by parts, and to discard the surface term as it vanishes on the boundary of the domain of integration. Finally, one gets

$$\delta S_M = \int \frac{\delta \mathcal{L}_M}{\delta \psi} \delta\psi d^4x \, , \tag{3.320}$$

where the expression

$$\frac{\delta \mathcal{L}_M}{\delta \psi} \equiv \frac{\partial \mathcal{L}_M}{\partial \psi} - \frac{\partial}{\partial x^\alpha} \left(\frac{\partial \mathcal{L}_M}{\partial \psi_{,\alpha}} \right) , \tag{3.321}$$

is called the variational derivative and, here and everywhere else, the comma before the sub-index denotes a partial derivative $\psi_{,\alpha} \equiv \partial_\alpha \psi$. The principle of the least action demands $\delta S_M = 0$ for arbitrary variation $\delta\psi$. The only possible way to satisfy this principle, is to demand vanishing of the variational derivative of $\mathcal{L}_M$,

$$\frac{\delta \mathcal{L}_M}{\delta \psi} = 0 \, . \tag{3.322}$$

Clearly, equations (3.321), (3.322) entails the Euler-Lagrange equations for the matter fields.

It is worth mentioning, that although one did calculations in this section with the partial derivatives from matter variables, the variational equations of motion for matter are tensorial, so that partial derivatives in equation (3.321) can be replaced with covariant ones. Indeed, $\mathcal{L}_M \equiv \mathcal{L}_M \left(\psi; \boldsymbol{\nabla}_\alpha \psi; g_{\alpha\beta} \right) = \mathcal{L}_M \left(\psi; \partial_\alpha \psi + \Gamma_\alpha \psi; g_{\alpha\beta} \right)$. Therefore,

$$\frac{\partial \mathcal{L}_M}{\partial \psi} = \left(\frac{\partial \mathcal{L}_M}{\partial \psi} \right)_{\boldsymbol{\nabla}\psi = \text{fixed}} + \frac{\partial \mathcal{L}_M}{\partial \psi_{,\alpha}} \Gamma_a \, , \tag{3.323}$$

where the partial derivative in the first term is taken for a fixed value of the covariant derivative of the field ψ [72]. Accounting for a rather obvious equivalence relationship

$$\frac{\partial \mathcal{L}_M}{\partial \psi_{,\alpha}} = \frac{\partial \mathcal{L}_M}{\partial \boldsymbol{\nabla}_\alpha \psi} \, , \tag{3.324}$$

71) Notice that variation does *not* commute with a covariant derivative, since according to equation (3.305) one has $\delta(\boldsymbol{\nabla}_a \psi) = \boldsymbol{\nabla}_\alpha (\delta\psi) + \psi \delta\Gamma_\alpha$.

72) This is an important remark since $\boldsymbol{\nabla}_\alpha \psi = \partial_\alpha \psi + \Gamma_\alpha \psi$ depends on the field ψ. This dependence is ignored when the first term in the right side of equation (3.323) is calculated.

one can re-write the variational derivative (3.318) in terms of the covariant derivative

$$\delta S_M = \int \left[\frac{\partial \mathcal{L}_M}{\partial \psi} \delta\psi + \frac{\partial \mathcal{L}_M}{\partial \nabla_\alpha \psi} \nabla_\alpha(\delta\psi) \right] d^4x\,, \tag{3.325}$$

where the first term in the integrand must be interpreted now in the sense of equation (3.323). Integrating second term by parts yields

$$\int \frac{\partial \mathcal{L}_M}{\partial \nabla_\alpha \psi} \nabla_\alpha(\delta\psi) d^4x = \int \nabla_\alpha \left(\frac{\partial \mathcal{L}_M}{\partial \nabla_\alpha \psi} \delta\psi \right) d^4x - \int \nabla_\alpha \left(\frac{\partial \mathcal{L}_M}{\partial \nabla_\alpha \psi} \right) \delta\psi d^4x\,, \tag{3.326}$$

The first integral in the right side is effectively a covariant divergence from a vector density $W^\alpha \equiv (\partial \mathcal{L}_M / \partial \nabla_\alpha \psi)\, \delta\psi$, where $\mathcal{L} = \sqrt{-g} L_M$ is a scalar density. Covariant divergence from a vector density is simplified to a partial derivative, which can be integrated by virtue of the Gauss theorem to a surface integral

$$\int \nabla_\alpha W^\alpha d^4x = \int \frac{\partial W^\alpha}{\partial x^\alpha} d^4x = \oint W^\alpha d\Sigma_\alpha\,, \tag{3.327}$$

where $d\Sigma_\alpha$ is the element of integration on the three-dimensional surface surrounding the domain of integration in the four-dimensional integral of the action. The surface integral in equation (3.327) vanishes because of the boundary conditions (3.316) imposed on variations. It allows us to recast equation (3.321) to explicitly covariant form

$$\frac{\delta \mathcal{L}_M}{\delta \psi} \equiv \frac{\partial \mathcal{L}_M}{\partial \psi} - \nabla_\alpha \left(\frac{\partial \mathcal{L}_M}{\partial \nabla_\alpha \psi} \right)\,, \tag{3.328}$$

quod erat demonstrandum. It is worth emphasizing that all equations of variational analysis of matter variables can be written down in terms of covariant derivatives instead of partial derivatives. The proof can be accomplished by direct calculations given, for example, in [Mitskevich, 1969; Szabados, 2009].

3.9.2.2 *Variational equations for gravitational field*

Our next goal is to vary the action with respect to the metric $g_{\alpha\beta}$, in a fashion similar to variation with respect to the matter field ψ. One reminds that the variation of the metric and its first derivatives are chosen to be nil at the boundary of the domain of integration. Calculating variation of the action δS with respect to the metric variation in the same manner as in previous section, one obtains

$$\delta S = \delta S_H + \delta S_M\,, \tag{3.329}$$

where variations

$$\delta S_H = \int \frac{\delta \mathcal{L}_H}{\delta g_{\alpha\beta}} \delta g_{\alpha\beta} d^4x\,, \tag{3.330}$$

$$\delta S_M = \int \frac{\delta \mathcal{L}_M}{\delta g_{\alpha\beta}} \delta g_{\alpha\beta} d^4x\,, \tag{3.331}$$

are given in terms of the variational derivatives taken with respect to the metric

$$\frac{\delta \mathcal{L}_H}{\delta g_{\alpha\beta}} \equiv \frac{\partial \mathcal{L}_H}{\partial g_{\alpha\beta}} - \frac{\partial}{\partial x^\mu}\left(\frac{\partial \mathcal{L}_H}{\partial g_{\alpha\beta,\mu}}\right) + \frac{\partial^2}{\partial x^\mu \partial x^\nu}\left(\frac{\partial \mathcal{L}_H}{\partial g_{\alpha\beta,\mu\nu}}\right) , \tag{3.332}$$

$$\frac{\delta \mathcal{L}_M}{\delta g_{\alpha\beta}} \equiv \frac{\partial \mathcal{L}_M}{\partial g_{\alpha\beta}} - \frac{\partial}{\partial x^\mu}\left(\frac{\partial \mathcal{L}_M}{\partial g_{\alpha\beta,\mu}}\right) , \tag{3.333}$$

where the comma next to sub-indices denote a partial derivative, $g_{\alpha\beta,\mu} \equiv \partial_\mu g_{\alpha\beta}$ and $g_{\alpha\beta,\mu\nu} \equiv \partial_{\mu\nu} g_{\alpha\beta}$. The principle of the least action demands $\delta S = 0$, which yields the variational equations for gravitational field

$$-\frac{\delta \mathcal{L}_H}{\delta g_{\alpha\beta}} = \frac{\delta \mathcal{L}_M}{\delta g_{\alpha\beta}} . \tag{3.334}$$

Left side of this equation depends only on the metric tensor and its derivatives while the right side depends on the matter fields being the source of gravity. The right side of equation (3.334) also depends on the metric and its first derivatives. Hence, equation (3.334) represents a complicated coupled system of equations for matter and gravitational field. If the Lagrangian of gravitational field is chosen in the form provided by Hilbert, the variational equations (3.334) take on the form of the Einstein field equations 3.251).

3.9.3 The Hilbert action and the Einstein equations

3.9.3.1 *The Hilbert Lagrangian*

In late June - early July of 1915, Einstein spent several days in Göttingen. In a series of six lectures, he explained his ideas on gravity. Among other things, he suggested that the Riemann metric's components, $g_{\alpha\beta}$, be employed as potentials of the gravitational field. Hilbert got interested in the approach outlined, and joined Einstein in the quest for the fundamental equations governing gravity. These equations were discovered independently by Hilbert and Einstein, in the months following Einstein's visit. Working on this problem, the two great minds were not competing but were actively exchanging information, as can be seen from their correspondence of November that year. Hilbert was tackling at the problem as a mathematician should, using variational methods. Einstein's way was more physically intuitive [73].

Hilbert's idea was that the Lagrangian, L_H, of gravitational field must be proportional to the Ricci scalar R which is the only scalar built of the metric and its first and second derivatives. It may look like that the appearance of the second derivatives of the metric in the gravitational Lagrangian gives rise to the field equations of the order higher than second. However, the Ricci scalar depends on the second derivatives of the metric linearly, and taking the variational derivative (3.332) from such a Lagrangian can not give higher-order derivatives of the metric tensor in the field equations, which will become equations of the second order. The principle of

73) A representative selection of excerpts from Einstein's papers of that period is provided in a review article by Logunov *et al.* [2004].

correspondence with the Newtonian theory (see section 3.8.4.2) fixes the constant of proportionality between L_H and R. Traditional definition of Hilbert's Lagrangian is

$$\mathcal{L}_H = -\frac{1}{2\kappa}\sqrt{-g}R\,, \tag{3.335}$$

where $\kappa = 8\pi G/c^4$. Functional derivative from $\mathcal{L}_H$ with respect to the metric, defines the Einstein tensor

$$G^{\alpha\beta} \equiv +\frac{2\kappa}{\sqrt{-g}}\frac{\delta\mathcal{L}_H}{\delta g_{\alpha\beta}}\,. \tag{3.336}$$

Functional derivative from the Lagrangian density of matter, $\mathcal{L}_M$, with respect to the metric yields the *metrical* energy-momentum tensor of matter

$$T^{\alpha\beta} \equiv -\frac{2}{\sqrt{-g}}\frac{\delta\mathcal{L}_M}{\delta g_{\alpha\beta}}\,, \tag{3.337}$$

that is also called the Hilbert tensor of energy-momentum. Substituting these definitions to variational equation (3.334) brings forth the Einstein field equations

$$G^{\alpha\beta} = \kappa T^{\alpha\beta}\,, \tag{3.338}$$

which were derived by Hilbert from the variational principle [74].

3.9.3.2 *The Einstein Lagrangian*

Einstein had also proposed a Lagrangian for gravitational field that does not depend on the second derivatives of the metric tensor. The procedure is based on splitting the Hilbert Lagrangian in two parts, one of each is a four-dimensional divergence from a pseudo-vector [75]

$$W^\alpha \equiv \sqrt{-g}\left(g^{\beta\gamma}\Gamma^\alpha_{\beta\gamma} - g^{\alpha\beta}\Gamma^\gamma_{\beta\gamma}\right) = \sqrt{-g}g^{\alpha\beta}g^{\mu\nu}\left(g_{\beta\mu,\nu} - g_{\mu\nu,\beta}\right);. \tag{3.340}$$

Straightforward calculation reveals [Landau and Lifshitz, 1975, §93] that

$$\sqrt{-g}R - \partial_\alpha W^\alpha = \sqrt{-g}g^{\mu\nu}\left(\Gamma^\alpha_{\mu\beta}\Gamma^\beta_{\nu\alpha} - \Gamma^\alpha_{\alpha\beta}\Gamma^\beta_{\mu\nu}\right)\,, \tag{3.341}$$

where the right side does not contain the second derivatives from the metric. The Einstein Lagrangian of gravitational field is defined as

$$\mathcal{L}_E = -\frac{1}{2\kappa}\sqrt{-g}g^{\mu\nu}\left(\Gamma^\alpha_{\mu\beta}\Gamma^\beta_{\nu\alpha} - \Gamma^\alpha_{\alpha\beta}\Gamma^\beta_{\mu\nu}\right)\,. \tag{3.342}$$

74) Notice that if one has used variational derivatives with respect to the contravariant metric tensor, the signs in definitions (3.336) and (3.337) had to be changed

$$G_{\alpha\beta} \equiv -\frac{2\kappa}{\sqrt{-g}}\frac{\delta\mathcal{L}_H}{\delta g^{\alpha\beta}}\,, \qquad T_{\alpha\beta} \equiv +\frac{2}{\sqrt{-g}}\frac{\delta\mathcal{L}_M}{\delta g^{\alpha\beta}}\,. \tag{3.339}$$

Signs in Einstein's equations remain the same: $G_{\alpha\beta} = \kappa T_{\alpha\beta}$. Comprehensive discussion of various relationships of variational calculus on curved manifolds is given in textbook by Mitskevich [1969].

75) We call W^α pseudo-vector as it has one index like a vector but it is made of a non-tensorial combination of the metric tensor and the Christoffel symbols that does not transform as a vector.

It is not a scalar and can be nullify at each point of spacetime manifold by choosing the normal Riemannian coordinates. For this reason, the Einstein Lagrangian makes sense only in the expression for the action. Indeed, the action,

$$S_E = \int \mathcal{L}_E d^4x \,, \tag{3.343}$$

for the Einstein Lagrangian (3.342) and that,

$$S_H = \int \mathcal{L}_H d^4x \,, \tag{3.344}$$

for the Hilbert Lagrangian (3.335), differ only by an integral from a total divergence, $\partial_\alpha W^\alpha$, which is reduced to a surface integral due to the Gauss theorem. The boundary conditions (3.316) imposed on the action variables and their derivatives tell us that the surface term does not play any role in derivation of the Einstein equations, and can be discarded. Hence, from the point of view of variational calculus, both types of the action are equivalent for the purpose of derivation of the Einstein field equations.

3.9.3.3 *The Einstein Tensor*

We still need to find out explicit expression for the Einstein tensor. The most direct way would be, of course, the calculation of this tensor from its definition (3.336) given in the form of the variational derivative. The main equations for partial derivatives with respect to the metric tensor and its partial derivatives are

$$\frac{\partial g_{\alpha\beta}}{\partial g_{\mu\nu}} = \frac{1}{2}\left(\delta^\mu_\alpha\delta^\nu_\beta + \delta^\nu_\alpha\delta^\mu_\beta\right) \,, \tag{3.345a}$$

$$\frac{\partial g_{\alpha\beta,\gamma}}{\partial g_{\mu\nu,\sigma}} = \frac{1}{2}\left(\delta^\mu_\alpha\delta^\nu_\beta + \delta^\nu_\alpha\delta^\mu_\beta\right)\delta^\sigma_\gamma \,, \tag{3.345b}$$

$$\frac{\partial g_{\alpha\beta,\gamma\pi}}{\partial g_{\mu\nu,\sigma\rho}} = \frac{1}{4}\left(\delta^\mu_\alpha\delta^\nu_\beta + \delta^\nu_\alpha\delta^\mu_\beta\right)\left(\delta^\sigma_\gamma\delta^\rho_\pi + \delta^\sigma_\pi\delta^\rho_\gamma\right) \,. \tag{3.345c}$$

Calculation of variational derivatives (3.332), (3.333) from any Lagrangian with the help of the above-given equations is straightforward but entails an enormous amount of tedious tensor algebra[76]. It can be more conveniently handled by a computer program for doing calculations with indexed objects [Klioner, 2000].

It turns out more simple to calculate the variational derivative from the variation of the action for gravitational field. Substituting Lagrangian (3.335) to definition (3.304) and taking variation one obtains

$$\delta S_H = -\frac{1}{2\kappa}\int \delta\mathcal{R} d^4x \,, \tag{3.346}$$

where the variation of the Ricci scalar density is

$$\delta\mathcal{R} = \delta\left(\sqrt{-g}g^{\alpha\beta}R_{\alpha\beta}\right) = R\delta\sqrt{-g} + \sqrt{-g}R_{\alpha\beta}\delta g^{\alpha\beta} + \sqrt{-g}g^{\alpha\beta}\delta R_{\alpha\beta} \,. \tag{3.347}$$

76) We thank Dr. Alexander Petrov from Sternberg Astronomical Institute (Moscow) for demonstrating the technique of variational derivatives on curved manifolds to us.

Variation

$$\delta\sqrt{-g} = -\frac{\delta g}{2\sqrt{-g}}\,, \tag{3.348}$$

is reduced to calculation of variation from determinant of the metric tensor, $g = \det\left[g_{\alpha\beta}\right]$. To derive this variation, let us recall the explicit expression for determinant of the metric tensor [Schutz, 1995, §4.12]

$$g = \frac{1}{4!}\varepsilon^{\alpha\beta\gamma\delta}\varepsilon^{\mu\nu\rho\sigma}g_{\alpha\mu}g_{\beta\nu}g_{\gamma\rho}g_{\delta\sigma}\,, \tag{3.349}$$

where $\varepsilon^{\alpha\beta\gamma\delta}$ is the fully antisymmetric Levi-Civita symbol defined in equation (3.229), and definition of the elements of the inverse matrix of the metric tensor,

$$gg^{\alpha\mu} = \frac{1}{3!}\varepsilon^{\alpha\beta\gamma\delta}\varepsilon^{\mu\nu\rho\sigma}g_{\beta\nu}g_{\gamma\rho}g_{\delta\sigma}\,. \tag{3.350}$$

One can easily check that equations (3.349) and (3.350) are consistent by using them for calculation of transvection of $g_{\alpha\beta}$ and $g^{\alpha\beta}$, which amounts to 4 as expected. After taking variation from both sides of equation (3.349) one gets,

$$\delta g = gg^{\alpha\beta}\delta g_{\alpha\beta}\,, \tag{3.351}$$

which yields

$$\delta\sqrt{-g} = \frac{1}{2}\sqrt{-g}g^{\alpha\beta}\delta g_{\alpha\beta}\,. \tag{3.352}$$

Second term in the right side of equation (3.347) can be re-written in terms of variation $\delta g_{\alpha\beta}$ with the help of relationship

$$g_{\alpha\mu}\delta g^{\mu\beta} = -g^{\mu\beta}\delta g_{\alpha\mu}\,, \tag{3.353}$$

which is a consequence of taking a variation from the identity $g_{\alpha\mu}g^{\mu\beta} = \delta^{\beta}_{\alpha}$. It results in relationship [77)]

$$\delta g^{\mu\nu} = -\frac{1}{2}\left(g^{\mu\alpha}g^{\nu\beta} + g^{\mu\beta}g^{\nu\alpha}\right)\delta g_{\alpha\beta}\,, \tag{3.354}$$

that allows us to recast the second term in equation (3.347) as follows

$$R_{\alpha\beta}\delta g^{\alpha\beta} = -R^{\alpha\beta}\delta g_{\alpha\beta}\,. \tag{3.355}$$

In order to calculate the third term in the right side of equation (3.347) one uses definition (3.170) to obtain the Ricci tensor

$$R_{\alpha\beta} = \partial_\mu\Gamma^\mu_{\alpha\beta} - \partial_\beta\Gamma^\mu_{\alpha\mu} + \Gamma^\mu_{\alpha\beta}\Gamma^\nu_{\mu\nu} - \Gamma^\mu_{\alpha\nu}\Gamma^\nu_{\beta\mu}\,. \tag{3.356}$$

77) Compare this variation with equation (3.345a). Notice the minus sign in front of the brackets as opposed to equation (3.345a).

Since the mathematical operation of taking a partial derivative and the operation of variation commutes one can write

$$\delta R_{\alpha\beta} = \partial_\mu \delta\Gamma^\mu_{\alpha\beta} - \partial_\beta \delta\Gamma^\mu_{\alpha\mu} + \delta\Gamma^\mu_{\alpha\beta}\Gamma^\nu_{\mu\nu} + \Gamma^\mu_{\alpha\beta}\delta\Gamma^\nu_{\mu\nu} - \delta\Gamma^\mu_{\alpha\nu}\Gamma^\nu_{\beta\mu} - \Gamma^\mu_{\alpha\nu}\delta\Gamma^\nu_{\beta\mu}\ . \tag{3.357}$$

Next step is to observe that the variation, $\delta\Gamma^\alpha_{\beta\gamma}$, of the Christoffel symbol is a tensor of third rank. Indeed, this variation consists of a difference between two different connections, $\Gamma'^\alpha_{\beta\gamma}$ and $\Gamma^\alpha_{\beta\gamma}$, taken at the same value of coordinates. This difference is transformed as a tensor in accordance with the reasonings given at the end of section 3.4.5. Thus, covariant derivative of the variation of the Christoffel symbols is

$$\nabla_\alpha \delta\Gamma^\mu_{\beta\gamma} = \partial_\alpha \delta\Gamma^\mu_{\beta\gamma} + \Gamma^\mu_{\alpha\nu}\delta\Gamma^\nu_{\beta\gamma} - \Gamma^\nu_{\alpha\gamma}\delta\Gamma^\mu_{\nu\beta} - \Gamma^\nu_{\alpha\beta}\delta\Gamma^\mu_{\nu\gamma}\ . \tag{3.358}$$

This expression allows us to recast equation (3.357) in simpler form

$$\delta R_{\alpha\beta} = \nabla_\mu \delta\Gamma^\mu_{\alpha\beta} - \nabla_\beta \delta\Gamma^\mu_{\alpha\mu}\ . \tag{3.359}$$

Contraction of this equation with $g^{\alpha\beta}$ and taking into account that covariant derivative from the metric tensor is zero, brings about the following expression

$$g^{\alpha\beta}\delta R_{\alpha\beta} = \nabla_\mu V^\mu\ , \tag{3.360}$$

where one has used a shorthand notation for a vector field

$$V^\mu \equiv g^{\alpha\beta}\delta\Gamma^\mu_{\alpha\beta} - g^{\alpha\mu}\delta\Gamma^\nu_{\alpha\nu}\ . \tag{3.361}$$

Now, one can apply a known expression for a covariant divergence of a vector [Misner *et al.*, 1973, Equation 8.51c]

$$\nabla_\mu V^\mu = \frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}V^\mu\right)\ , \tag{3.362}$$

in order to present equation (3.359) in the form

$$g^{\alpha\beta}\delta R_{\alpha\beta} = \frac{1}{\sqrt{-g}}\partial_\mu\left[\sqrt{-g}\left(g^{\alpha\beta}\delta\Gamma^\mu_{\alpha\beta} - g^{\alpha\mu}\delta\Gamma^\nu_{\alpha\nu}\right)\right]\ . \tag{3.363}$$

Replacing this expression in the third term in the right side of equation (3.347) and integrating, yields

$$\begin{aligned}\int \sqrt{-g}g^{\alpha\beta}\delta R_{\alpha\beta}d^4x &= \int \partial_\mu\left[\sqrt{-g}\left(g^{\alpha\beta}\delta\Gamma^\mu_{\alpha\beta} - g^{\alpha\mu}\delta\Gamma^\nu_{\alpha\nu}\right)\right]d^4x \\ &= \oint \sqrt{-g}\left(g^{\alpha\beta}\delta\Gamma^\mu_{\alpha\beta} - g^{\alpha\mu}\delta\Gamma^\nu_{\alpha\nu}\right)d\Sigma_\mu\ ,\end{aligned} \tag{3.364}$$

where one has used the four-dimensional version of the divergence (Gauss) theorem [Schutz, 1995, §4.23] in order to reduce the volume integral from a divergence of a vector field to a hypersurface integral from the vector field itself, with the unit element of integration on the surface [78] being $d\Sigma_\mu$. Since one assumes that all variations of fields and their derivatives vanish on the boundary of integration, the surface integral in equation (3.364) is nil, and does not contribute to the result of calculation of variation δS_H.

Hence, equation (3.346) becomes

$$\delta S_H = \frac{1}{2\kappa}\int \sqrt{-g}\left(R^{\alpha\beta} - \frac{1}{2}g^{\alpha\beta}R\right)\delta g_{\alpha\beta}d^4x\ , \tag{3.365}$$

which should be compared with definition (3.330) of the variational derivative and that (3.336) of the Einstein tensor. It unveils the structure of the Einstein tensor obtained from the principle of the least action,

$$G^{\alpha\beta} \equiv R^{\alpha\beta} - \frac{1}{2}g^{\alpha\beta}R\ . \tag{3.366}$$

This tensor was introduced to general relativity by Einstein, independently of Hilbert. The main motivation for Einstein to use this tensor in the field equations was its conservation, $\nabla_\alpha G^{\alpha\beta} \equiv 0$, known under the name of the Bianchi identity that has been discussed in section 3.7.6.3.

3.9.3.4 *The Generalizations of the Hilbert Lagrangian*

Progress in understanding the basic fundamental principles of gravitational physics led scientists to search for generalizations of the Einstein theory of gravity. Cosmology suggests existence of a special epoch in the history of the early universe called inflation [Mukhanov, 2005]. Inflation is associated with a period of an exponentially-fast expansion of the universe driven by a hypothetical scalar field or a multiplet of the fields which still may affect the orbital motion of celestial bodies [Damour and Esposito-Farese, 1992, 1993; Damour and Esposito-Farèse, 1998a] and other astronomical phenomena [Damour and Nordtvedt, 1993a,b]. It suggests to use the scalar-tensor theory of gravity for discussing relativistic frames and equations of motion of massive bodies for extension of the existing IAU resolutions on reference frames, time scales and equations of motion [Kopeikin, 2010a].

Scalar fields admit the equation of state being formally equivalent to the equation of state of a perfect fluid with negative pressure, $p < 0$. The most extremal case is the equation of state $p = -\epsilon$ corresponding to a Lorentz-invariant tensor of energy-momentum of vacuum

$$T^{\alpha\beta}_{\Lambda} = -\frac{\Lambda}{\kappa}g^{\alpha\beta}\ , \tag{3.367}$$

where Λ is, so-called, *cosmological constant* having dimension of $1/length^2$. Formally, substitution of this tensor of energy-momentum to the Einstein equations

78) In particular case, when normal to the hypersurface coincides locally with direction of time axis, the hypersurface element of integration is merely an element dV of spatial volume, $d\Sigma_\mu = (d\Sigma_0, 0, 0, 0) = (dV, 0, 0, 0)$.

would correspond to adding a new term $L_\Lambda = \Lambda$ to the overall Lagrangian of general relativity,

$$L = L_\Lambda + L_H + L_M \,. \tag{3.368}$$

Einstein himself introduced the (negatively-valued) cosmological term, $\Lambda < 0$ in vain hope to build a static cosmological model of universe filled up with matter and gravitational field obeying the Einstein equations. Tensor of energy-momentum with a negative Λ-term counteracts gravity and prevents the model of universe from expansion or collapse if the numerical value of Λ is chosen correspondingly. Through the fine-tuning of Λ to the average cosmological density of matter, such a model was assembled, but it turned out to be unstable under small perturbations, so the cosmological term appeared to be of no use, and Einstein dropped it out from the gravity field equations. After the expansion of the Universe was discovered by Hubble, the cosmological models admitting existence of a stationary epoch (like the Lemaître model [Lemaître, 1931]) fell into disuse for quite a long time. However, the cosmological constant remained a subject of theoretical and empirical interest. It was born to a new life and returned to cosmology after the theory of inflation was proposed by Guth [1981] and further developed by Linde [1990].

The situation became even more favorable to restoration of the cosmological constant, after several observational groups had claimed that the most remote Type-1a supernovas are moving radially away from the solar system much faster than the closer ones, a surprise indicating that the Hubble expansion is effectively accelerating, not decelerating. While critics suggested that the distance to cosmological supernovas was overevaluated (and that their light gets dimmed by inter-galactic dust, not distance), the belief in accelerated expansion of the universe has become mainstream. This result promulgated support to the concept of dark energy which, if exists, yields a nonvanishing cosmological constant and, thereby, the acceleration. Minimalist's way of implementing the dark energy is simply to postulate the existence of a cosmological constant as an *a priori* given entity. A more likely implementation of the dark-energy concept is a scalar field called *quintessence* with energy density and pressure obeying the equation of state, $p = w\epsilon$ with a constant parameter, $w < -1/3$, causing the acceleration of cosmic expansion. Quintessence is hypothesized to be dynamic and to have density and pressure varying over time and space [Gromov *et al.*, 2004; Ostriker and Steinhardt, 2001; Teerikorpi *et al.*, 2003].

Quantum mechanics associates the energy of vacuum with the existence of quantum fluctuations of various physical fields virtually present in vacuum - the idea pioneered by Paul Dirac and later on, in application to gravity, by Andrei Sakharov [Sakharov, 2000, 1967]. The problem with the idea of the cosmological constant induced by the fluctuations of quantized fields is that the vacuum energy diverges at high frequencies. It has a natural cut-off associated with the Planck scale

$$\ell_{Pl} \equiv \sqrt{\frac{G\hbar}{c^3}} = 1.6 \times 10^{-35}\mathrm{m} \,, \tag{3.369}$$

where G and $\hbar$ are the universal gravitational and Planck's constants, respectively.

Direct employment of this cut-off produces a value of

$$\Lambda \sim \frac{1}{\ell_{Pl}^2} \simeq 0.4 \times 10^{70}\mathrm{m}^{-2} , \tag{3.370}$$

which exceeds current limitation on it, coming from measurement of fluctuations of cosmic microwave background radiation [Tegmark *et al.*, 2004], by about 120 orders of magnitude. There are other problems discussed in article by Dyson *et al.* [2002].

The problem of divergency of vacuum energy indicates that classical general relativity can not be directly extended to the quantum world. It must be generalized to account for the quantum fluctuations of spacetime itself. There are several approaches to tackle the problem among which string theory, M-theory, and loop quantum gravity look the most promising. Smolin [2001] suggests that these approaches may be approximations of a single, underlying theory. From a phenomenological point of view all these approaches assume that the Hilbert action, $S_H = \int R\sqrt{-g}d^4x$, of general theory of relativity must be replaced with a new one,

$$S_H \to S_G = \int d^D x\sqrt{-g}L_G\left(R, R^{ab}R_{ab}, R^{abcd}R_{abcd}, ...\right) , \tag{3.371}$$

where the new Lagrangian is a function of various scalars, and perhaps their derivatives, made of the Riemann tensor of D-dimensional space (the indices $a, b, ...$ run through values $(0, 1, 2, ..., D)$ that incorporates higher-order dimensions besides the known four-dimensional spacetime. These ideas allows an enormous freedom for theoretical speculations. Nevertheless, the correspondence principle demands that all such theories must match with general relativity in the low-energy limit. It means that there must exist a reduction procedure from higher-dimensional space integration to four-dimensional spacetime where the Lagrangian, L_G, appears as a Taylor expansion with respect to the scalars made of the four-dimensional Riemann tensor,

$$S_G = \int d^4x\sqrt{-g}\left(\Lambda + R + \gimel_1 R^2 + \gimel_2 R^{\mu\nu}R_{\mu\nu} + \gimel_3 R^{\alpha\beta\gamma\delta}R_{\alpha\beta\gamma\delta} + ...\right), \tag{3.372}$$

with Λ being the cosmological constant, coefficients $\gimel_i$ $(i = 1, 2, 3, ...)$ being the coupling constants coming from the reduction of extra-dimensions, and ellipses denoting all the possible scalars one can assemble out of the metric tensor, the Levi-Civita symbol, and curvature tensor through multiplication, differentiation, and transvection. It is remarkable that a specifically chosen combinations of the curvature terms in this expansion allow us to keep the modified differential equations for the modified gravity field not exceeding the second order. These combinations are known as the Lovelock scalars [Lovelock, 1971, 1972] among which the Gauss-Bonnet term is the most famous one,

$$\mathcal{G} \equiv R^2 - 4R^{\alpha\beta}R_{\alpha\beta} + R^{\alpha\beta\gamma\delta}R_{\alpha\beta\gamma\delta} , \tag{3.373}$$

This term is only nontrivial in five-dimensional space or greater as it reduces to a surface term in four-dimensional spacetime,

$$\int \mathcal{G}\sqrt{-g}d^4x = \int \partial_\alpha \mathcal{D}^\alpha d^4x = \oint \mathcal{D}^\alpha d\Sigma_\alpha , \tag{3.374}$$

where vector $\mathcal{D}^\alpha$ is defined as follows [de Felice and Tsujikawa, 2010],

$$\mathcal{D}^\alpha = \frac{1}{2}\sqrt{-g}\varepsilon^{\alpha\beta\gamma\delta}\varepsilon^{\rho\sigma\mu\nu}\Gamma_{\rho\mu\beta}\left(R_{\sigma\nu\gamma\delta} + \frac{2}{3}\Gamma_{\sigma\lambda\gamma}\Gamma^\lambda_{\nu\sigma}\right) , \tag{3.375}$$

where $\varepsilon^{\alpha\beta\gamma\delta}$ is the anti-symmetric Levi-Civita symbol defined in equation (3.229), and $\Gamma_{\alpha\beta\gamma} = g_{\alpha\mu}\Gamma^\mu_{\beta\gamma}$ is the Christoffel symbol. Hence, the combination of the coupling constants giving rise to the Gauss-Bonnet term in equation (3.372) should be either carefully excluded or the Gauss-Bonnet term must be multiplied with some other functions of the curvature scalars in order to provide a physically meaningful extension of the Hilbert Lagrangian for gravitational field.

In order to keep the dimension of all terms in the expansion the same as that of Hilbert's term $L = R$, one should account for that the coupling constants have corresponding dimensions. Since they are supposedly reflect the presence of quantum fluctuations of gravity field, their corresponding values are to be, $\daleth_1 \sim \daleth_2 \sim \daleth_3 \sim \ell^2_{Pl}$, and so on. These values are by many orders of magnitude lower than the values of spacetime curvature achievable in any thinkable gravitational experiment. Hence, if this scenario for the Lagrangian of gravitational field is selected by Nature one may safely neglect the nonlinear terms in the action, should they exist. Nonetheless, things may be not so straightforward and the coupling constants in the Taylor expansion (3.372) of the gravitational Lagrangian may be generated by different mechanisms that will take their values to the range of observational capabilities in cosmology and/or in high-energy particle physics [Carroll *et al.*, 2004; de Felice and Tsujikawa, 2010].

3.9.4
The Noether theorem and conserved currents

Noether's theorem states that any differentiable symmetry of the action, S, of a physical system has a corresponding conservation law. The theorem was proved by Emmy Noether in 1915 and published in 1918 [Noether, 1918] [79] In order to analyze the conservation laws in gravity, one needs to perceive the nature of the variations of dynamic variables defining the evolution of the action of the physical system.

The Noether theorem will be derived in this section for a general case of a Lagrangian depending on a set of dynamic variable that are tensor fields, ϕ^A, where the cumulative index A labels the fields and their components. Spacetime indices of the fields are suppressed to avoid complex notations. These indices can be introduced later, when necessary. The fields are functions defined on a four-dimensional spacetime manifold whose points are parameterized by coordinates $x \equiv x^\alpha = (x^0, x^i)$.

3.9.4.1 *The anatomy of the infinitesimal variation*

Variation of an independent dynamic field variable consists, in general, of three parts:

79) English translations of the original article are available [Noether, 1971, 2005].

1) *Actual* variation that changes the variable from one function to another in functional space without any constrain. Variation of the action is not nil, if the actual variation is applied.
2) *Intrinsic gauge* variation that changes the variable from one function to another in functional space with the constrain that keeps variation of the action equal to zero.
3) *Extrinsic gauge* variation that is caused by an infinitesimally-small coordinate transformation. Due to the covariant nature of physical laws, the extrinsic gauge variation leaves the action unchanged.

The intrinsic and extrinsic gauge variations are essentially disparate as they are associated with different symmetries of the Lagrangian of the action of the physical system.

Let us make an infinitely small deformation of both fields and coordinates. Under the infinitesimal transformation, the change in the coordinates (diffeomorphism) is written

$$x'^{\mu} = x^{\mu} + \xi^{\mu}(x) \,, \tag{3.376}$$

where ξ^{α} is a vector field defining a displacement of the coordinate grid, and vanishing on the boundary of the integration of the action. It would be right to view this displacement as a Lie transform of each point, x, of the spacetime manifold $\mathcal{M}$ to a point, x', of the deformed manifold $\mathcal{M}'$, caused by the flow of vector field ξ^{α}, and executed along a congruence of its integral curves (see Figure **??**). The overall variation of the field variables is expressed as

$$\phi'^{A}(x) = \phi^{A}(x) + \delta\phi^{A}(x) \,, \tag{3.377}$$

and it results from the three factors: *actual* change in the fields themselves, the *intrinsic* change associated with the internal (Lie group) symmetry of the field equations, and the extrinsic change due to the Lie transform of the manifold. The intrinsic and extrinsic variations are called gauge transformations of the fields.

To isolate the extrinsic gauge variation, let us look at the transformed fields produced by the diffeomorphism (3.376). It deforms the fields as follows [80)]

$$\phi'^{A}(x') = \Lambda^{A}{}_{B}(x', x)\phi^{B}(x) \,, \tag{3.378}$$

where $\Lambda^{A}{}_{B}(x', x)$ is a symbolic notation for the operator of transformation of the field ϕ^{A} that is produced by the diffeomorphism (3.376). Exact structure of $\Lambda^{A}{}_{B}(x', x)$ is dictated by the nature of the field ϕ^{A} and it may be not tensorial, like in case of the Christoffel symbols which are allowed to be considered as an independent dynamic variable in variational calculus on curved manifolds [Misner *et al.*, 1973]. From the point of view of the Lie transform, the transformed field $\phi'^{A}(x')$ resides on the deformed manifold $\mathcal{M}'$ but one can pull it back to the undistorted manifold $\mathcal{M}$ by

80) We recommend to compare the explanation of Lie transform given in this section with that done in terms of the vector flow in section 3.6.

making use of the Lie dragging along the integral curves of the field ξ^α. It is achieved by doing a Taylor expansion of the field about the point x,

$$\phi'^A(x') = \phi'^A(x) + \xi^\alpha \partial_\alpha \phi^A(x) + O\left(\xi^2\right) , \tag{3.379}$$

It is then substituted to the left side of equation (3.378) where the matrix $\Lambda^A{}_B(x', x)$ has to be expanded in Taylor's series with respect to the magnitude of components of vector field ξ^α. Calculating variation (3.377) of the fields reveal that their extrinsic gauge variation induced by the infinitesimal diffeomorphism, is a Lie derivative (see section 3.6) of the field along a vector flow, ξ^α:

$$\delta\phi^A = -\pounds_{\boldsymbol{\xi}}\phi^A . \tag{3.380}$$

We recall that irrespectively of the nature of the field shown in equation (3.380), its Lie derivative is a tensor [Mitskevich, 1969; Popova and Petrov, 1988]. In case, when $\phi^A \equiv \phi^{\alpha_1\alpha_2\ldots\alpha_p}_{\beta_1\beta_2\ldots\beta_q}$ is a tensor field of a (p, q) type, the Lie derivative of the tensor is given by equation

$$\pounds_{\boldsymbol{\xi}}\phi^A \equiv \xi^\mu \boldsymbol{\nabla}_\mu \phi^A - \boldsymbol{\nabla}_\mu \xi^\nu \mathfrak{D}^\mu{}_\nu \phi^A , \tag{3.381}$$

where $\boldsymbol{\nabla}_\mu$ is a covariant derivative defined in terms of the Christoffel symbols of the metric $g_{\alpha\beta}$, and symbols $\mathfrak{D}^\mu{}_\nu\phi^A$ denotes a tensor product

$$\mathfrak{D}^\mu{}_\nu\phi^A \equiv (\mathfrak{D}^\mu{}_\nu)^{\alpha_1\ldots\alpha_p\rho_1\ldots\rho_q}_{\beta_1\ldots\beta_q\gamma_1\ldots\gamma_p}\,\phi^{\gamma_1\ldots\gamma_p}_{\rho_1\ldots\rho_q} . \tag{3.382}$$

Here, the tensor $\mathfrak{D}^\mu{}_\nu = 0$, if $p = q = 0$, and in all other cases it is just the constant tensor operator made of the Kronecker symbols [Szabados, 2009],

$$\begin{aligned}(\mathfrak{D}^\mu{}_\nu)^{\alpha_1\ldots\alpha_p\rho_1\ldots\rho_q}_{\beta_1\ldots\beta_q\gamma_1\ldots\gamma_p} \quad &\equiv \quad \left(\delta^{\alpha_1}_\nu \delta^\mu_{\gamma_1}\ldots\delta^{\alpha_p}_{\gamma_p} + \ldots + \delta^{\alpha_1}_{\gamma_1}\ldots\delta^{\alpha_p}_\nu \delta^\mu_{\gamma_p}\right)\delta^{\rho_1}_{\beta_1}\ldots\delta^{\rho_q}_{\beta_q} \\ &- \quad \left(\delta^{\rho_1}_\nu \delta^\mu_{\beta_1}\ldots\delta^{\rho_q}_{\beta_q} + \ldots + \delta^{\rho_1}_{\beta_1}\ldots\delta^{\rho_q}_\nu \delta^\mu_{\beta_q}\right)\delta^{\alpha_1}_{\gamma_1}\ldots\delta^{\alpha_p}_{\gamma_p} .\end{aligned} \tag{3.383}$$

It can be checked by inspection that definition (3.381) of the Lie derivative is identical to that given in equation (3.156). Covariant derivative, $\boldsymbol{\nabla}_\mu$, in equation (3.381) can be replaced with a partial derivative, ∂_μ. This equivalent definition of the Lie derivative has been given above in equation (3.153).

Finally, one can write decomposition of variation of the physical fields in the Lagrangian of the action as follows,

$$\delta\phi^A(x) = \delta^\dagger\phi^A(x) + \delta^\star\phi^A(x) + \delta^*\phi^A(x) , \tag{3.384}$$

where $\delta^\dagger\phi^A$ is the *actual* variation, $\delta^\star\phi^A$ is the *intrinsic gauge* variation, and

$$\delta^*\phi^A(x) = -\pounds_{\boldsymbol{\xi}}\phi^A , \tag{3.385}$$

is the *extrinsic* gauge variation due to the coordinate diffeomorphism. The intrinsic gauge transformations correspond to the intrinsic change of the field variables that are associated with internal gauge symmetries of the physical system, and they have nothing to do with the diffeomorphisms at all. An example of such intrinsic gauge transformation is delivered by the gradient transform (2.24) of electromagnetic vector potential, $A^{\to}A'^{\alpha} = A^{\alpha} + \partial^{\alpha}\chi$, where χ is arbitrary scalar function. This transform is totally separate from any change in coordinates but leaves the Maxwell equations (2.16) invariant. Other examples of the intrinsic gauge symmetries may be found in modern gauge theories dealing with amazingly rich world of elementary particles and their interactions [O'Raifeartaigh and Straumann, 2000].

3.9.4.2 ***Examples of the gauge transformations***

To see more clearly how equation (3.380) works, let us dwell upon few examples. First, one takes a scalar field, $\phi^A \equiv \phi$. In this case, equation (3.378) is merely

$$\phi'(x') = \phi(x) , \tag{3.386}$$

so that $\Lambda^A{}_B = 1$ for a scalar field. Applying equation (3.379) and calculating the variation $\delta\phi(x) = \phi'(x) - \phi(x)$ yields

$$\delta\phi = -\xi^{\alpha}\phi_{,\alpha} \equiv -£_{\boldsymbol{\xi}}\phi , \tag{3.387}$$

confirming equation (3.380).

Second example covers a vector field, $\phi^A \equiv A^{\alpha}$. Equations (3.378) and (3.379) read now as follows

$$A'^{\alpha}(x') = \frac{\partial x'^{\alpha}}{\partial x^{\beta}} A^{\beta}(x) , \qquad A'^{\alpha}(x') = A'^{\alpha}(x) + \xi^{\beta}\partial_{\beta}A^{\alpha}(x) , \tag{3.388}$$

and the matrix of the field transformation is

$$\Lambda^A{}_B \equiv \frac{\partial x'^{\alpha}}{\partial x^{\beta}} = \delta^{\alpha}_{\beta} + \partial_{\beta}\xi^{\alpha} . \tag{3.389}$$

These equations give rise to the field variation

$$\delta A^{\alpha} = -\xi^{\beta}\partial_{\beta}A^{\alpha} + \partial_{\beta}\xi^{\alpha}A^{\beta} \equiv -£_{\boldsymbol{\xi}}A^{\alpha} , \tag{3.390}$$

again in accordance with equation (3.380). Had the dynamic variable been a covector, A_{α}, the gauge variation due to the change of coordinates would be given by

$$\delta A_{\alpha} = -\xi^{\beta}\partial_{\beta}A_{\alpha} - \partial_{\alpha}\xi^{\beta}A_{\beta} = -£_{\boldsymbol{\xi}}A_{\alpha} . \tag{3.391}$$

The last example addresses the gauge variation of the metric tensor, $g_{\alpha\beta}$. In this case equations (3.378) and (3.379) are equivalent to

$$g'_{\alpha\beta}(x') = \frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\beta}} g_{\mu\nu}(x) , \qquad g'_{\alpha\beta}(x') = g'_{\alpha\beta}(x) + \xi^{\gamma}g_{\alpha\beta,\gamma}(x) . \tag{3.392}$$

The field transformation matrix is

$$\Lambda^A{}_B \equiv \frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\beta}} = \delta^{\mu}_{\alpha}\delta^{\nu}_{\beta} - \delta^{\mu}_{\alpha}\partial_{\beta}\xi^{\nu} - \delta^{\nu}_{\beta}\partial_{\alpha}\xi^{\mu} , \tag{3.393}$$

so that the previous equations result in the metric variation

$$\delta g_{\alpha\beta} = -\pounds_{\boldsymbol{\xi}} g_{\alpha\beta} = -\boldsymbol{\nabla}_\alpha \xi_\beta - \boldsymbol{\nabla}_\beta \xi_\alpha \ , \tag{3.394}$$

where $\boldsymbol{\nabla}_\alpha$ is a covariant derivative defined in terms of the Christoffel symbols of the metric tensor as shown in equation (3.34). In case of a contravariant components of the metric tensor, the gauge variation of these components is

$$\delta g^{\alpha\beta} = -\pounds_{\boldsymbol{\xi}} g^{\alpha\beta} = \boldsymbol{\nabla}^\alpha \xi^\beta + \boldsymbol{\nabla}^\beta \xi^\alpha \ , \tag{3.395}$$

Notice that the extrinsic gauge variation of the metric depends on the covariant derivative of vector field ξ^α only if the metric is fully compatible with the affine connection. If the manifold has torsion and/or nonmetricity they must appear explicitly in addition to the covariant derivatives of the field ξ^α. Irrespectively of the relationship between the metric tensor and the affine connection, the extrinsic gauge variation of the metric tensor given in the form of the Lie derivative is always valid.

3.9.4.3 ***Proof of the Noether Theorem***

We shall assume in this section that the Lagrangian, $\mathcal{L} = L\sqrt{-g}$, of the physical system depends explicitly on the dynamic field variables, ϕ^A, their first and second derivatives, as well as on coordinates x^α,

$$\mathcal{L} \equiv \mathcal{L}\left(\phi^A; \partial_\alpha\phi^A; \partial_{\alpha\beta}\phi^A; x^\alpha\right) \ . \tag{3.396}$$

Noether's theorem begins with the assumption that a physical system evolves in accordance with the Euler-Lagrange equations derived from the principle of the least action for this system,

$$\frac{\partial \mathcal{L}}{\partial \phi^A} - \frac{\partial}{\partial x^\alpha}\left(\frac{\partial \mathcal{L}}{\partial \phi^A_\alpha}\right) + \frac{\partial^2}{\partial x^\alpha \partial x^\beta}\left(\frac{\partial \mathcal{L}}{\partial \phi^A_{\alpha\beta}}\right) = 0 \ , \tag{3.397}$$

where notations $\phi^A_\alpha \equiv \partial_\alpha\phi^A$ and $\phi^A_{\alpha\beta} \equiv \partial_{\alpha\beta}\phi^A$ for the partial derivatives of ϕ^A have been introduced.

The system that obeys these equations is called on-shell. Noether's theorem states that on-shell system may admit a specific transformation of the coordinates and/or the field variables which does not change the Lagrangian. In other words, the on-shell Lagrangian may admit a gauge transformation such that its variation

$$\delta\mathcal{L} = \mathcal{L}' - \mathcal{L} = 0 \ , \tag{3.398}$$

provided that

$$\mathcal{L}' = \mathcal{L}\left(\phi^A + \delta\phi^A; \phi^A_\alpha + \delta\phi^A_\alpha; \phi^A_{\alpha\beta} + \delta\phi^A_{\alpha\beta}; x^\alpha + \xi^\alpha\right) \ , \tag{3.399}$$

where again, $\phi^A_\alpha \equiv \partial_\alpha\phi^A$ and $\phi^A_{\alpha\beta} \equiv \partial_{\alpha\beta}\phi^A$. On shell variation of the Lagrangian assumes that only the intrinsic and extrinsic gauge variations of the fields are allowed

$$\delta\phi^A(x) = \delta^\star\phi^A(x) + \delta^*\phi^A \ . \tag{3.400}$$

because the actual variations of the fields are nullified, $\delta^{\dagger}\phi^A(x) = 0$, as the fields obey the Euler-Lagrange equations (3.397).

Expanding the difference in Lagrangians in a Taylor series to first-order in the infinitesimal variations and making use of the Euler-Lagrange equations (3.397), allow us to write the Lagrangian variation as follows [Bak *et al.*, 1994],

$$\delta\mathcal{L} = \frac{\partial}{\partial x^\alpha}\left[\xi^\alpha\mathcal{L} + \frac{\delta\mathcal{L}}{\delta\phi^A_\alpha}\delta\phi^A + \frac{\partial\mathcal{L}}{\partial\phi^A_{\alpha\beta}}\delta\phi^A_\alpha\right] , \tag{3.401}$$

where $\phi^A_\alpha \equiv \partial_\alpha\phi^A$, $\phi^A_{\alpha\beta} \equiv \partial_{\alpha\beta}\phi^A$, and one has introduced a new notation for variational derivative of the Lagrangian with respect to the field derivatives [Popova and Petrov, 1988]

$$\frac{\delta\mathcal{L}}{\delta\phi^A_\alpha} \equiv \frac{\partial\mathcal{L}}{\partial\phi^A_\alpha} - \frac{\partial}{\partial x^\beta}\left(\frac{\partial\mathcal{L}}{\partial\phi^A_{\alpha\beta}}\right) . \tag{3.402}$$

Since the gauge variation of the Lagrangian, $\delta\mathcal{L} = 0$, equation (3.401) can be rewritten as the law of conservation for a vector field

$$\frac{\partial J^\alpha_A}{\partial x^\alpha} = 0 , \tag{3.403}$$

where

$$J^\alpha_A = \xi^\alpha\mathcal{L} + \frac{\delta\mathcal{L}}{\delta\phi^A_\alpha}\delta\phi^A + \frac{\partial\mathcal{L}}{\partial\phi^A_{\alpha\beta}}\delta\phi^A_\alpha . \tag{3.404}$$

Vector field J^α_A, is called Noether's current, and equation (3.403) finalizes the proof of the Noether theorem.

Since one has used the Lagrangian density, $\mathcal{L} = L\sqrt{-g}$, in derivation of Noether's current conservation law, one can integrate equation (3.403) directly over entire spacetime manifold, $\int \partial_\alpha J^\alpha_A d^4x = 0$. Applying the divergence theorem, like one did it in equation (3.327), yields a conserved Noether's charge,

$$Q_A = \oint J^\alpha_A d\Sigma_\alpha; , \tag{3.405}$$

where the integration is over any spacelike three-dimensional hypersurface. One may have as many Noether' charges as the number of the gauge degrees of freedom of the physical system under consideration. As a rule, the hypersurface of integration is chosen to be an infinite three-dimensional space with the volume element $dV = d\Sigma_0$. Then, the Noether charge is expressed as a volume integral from the time component of the Noether current

$$Q_A = \int J^0_A dV; , \tag{3.406}$$

which does not depend on time, $dQ_A/dt = 0$, provided that the surface integral from spatial components, J^i_A, of the Noether current vanishes at infinity.

In fact, the Noether current J^α_A is comprised of two, linearly-independent parts corresponding to the extrinsic and intrinsic gauge variations of the fields

$$J^\alpha_A = \overset{\star}{J}{}^\alpha_A + \overset{*}{J}{}^\alpha_A \,, \tag{3.407}$$

where

$$\overset{\star}{J}{}^\alpha_A = \frac{\delta \mathcal{L}}{\delta \phi^A_\alpha}\delta^\star \phi^A + \frac{\partial \mathcal{L}}{\partial \phi^A_{\alpha\beta}}\delta^\star \phi^A_\alpha \,, \tag{3.408}$$

$$\overset{*}{J}{}^\alpha_A = \xi^\alpha \mathcal{L} + \frac{\delta \mathcal{L}}{\delta \phi^A_\alpha}\delta^* \phi^A + \frac{\partial \mathcal{L}}{\partial \phi^A_{\alpha\beta}}\delta^* \phi^A_\alpha \,. \tag{3.409}$$

The Noether current, J^α_A, is defined not uniquely but up to a term being a divergence of an antisymmetric tensor of second rank that vanishes on the boundary of integration of the action. Indeed, let us take the current in the form

$$J^\alpha_A = \mathfrak{J}^\alpha_A + \partial_\beta \mathfrak{X}^{[\alpha\beta]}_A \,, \tag{3.410}$$

where, $\mathfrak{X}^{[\alpha\beta]}_A = (1/2)(\mathfrak{X}^{\alpha\beta}_A - \mathfrak{X}^{\beta\alpha}_A)$, is a fully antisymmetric object. Substituting the current J^α_A into the law of conservation (3.403), testifies that the new current, $\mathfrak{J}^\alpha_A$, is conserved, because $\partial_\alpha \mathfrak{J}^\alpha_A$ due to identity $\partial_{\alpha\beta}\mathfrak{X}^{\alpha\beta} \equiv 0$.

This device of adding the divergence of an antisymmetric tensor, $\theta^{[\alpha\beta]}_A$, can be used to improve the *canonical Noether current* so as to attain some other property, such as the symmetry of the energy-momentum tensor while maintaining current conservation and the same value of the Noether charge. Such antisymmetric objects that are allowed to appear in definition of the Noether current are called *super-potentials* [Petrov, 2009, 2010; Petrov and Katz, 2002; Szabados, 2009]. They have been introduced to Noether's theory by Belinfante [1939, 1940] and Rosenfeld [1940].

3.9.5
The metrical energy-momentum tensor

3.9.5.1 *Hardcore of the Metrical Energy-Momentum Tensor*

The metrical energy momentum tensor was introduced by Hilbert and is given by equation (3.337) in terms of the variational derivative from the scalar density of the Lagrangian of matter, $\mathcal{L}_M = \sqrt{-g}L_M$, where L_M is the scalar Lagrangian itself. Following the *principal of minimal coupling* of matter and gravitational field, one assumes that the Lagrangian $L_M = L_M(\psi; \nabla_\alpha\psi; g_{\alpha\beta}) = L_M\left(\psi; \psi_{,\alpha} + \Gamma_\alpha\psi; g_{\alpha\beta}\right)$ depends on matter variables ψ, the metric tensor $g_{\alpha\beta}$, and its first derivatives $g_{\alpha\beta,\gamma}$ that enters the covariant derivatives of the matter variables in the form of Γ_α comprised of the Christoffel symbols in accordance with equation (3.305).

After taking a variational derivative from $\mathcal{L}_M$, equation (3.337) assumes the following form:

$$T^{\alpha\beta} = -g^{\alpha\beta}L_M - 2\frac{\delta L_M}{\delta g_{\alpha\beta}} + 2\Gamma^\mu_{\gamma\mu}\frac{\partial L_M}{\partial g_{\alpha\beta,\gamma}} \,. \tag{3.411}$$

where the variational derivative from the Lagrangian L_M is

$$\frac{\delta L_M}{\delta g_{\alpha\beta}} = \frac{\partial L_M}{\partial g_{\alpha\beta}} - \frac{\partial}{\partial x^\gamma}\left(\frac{\partial L_M}{\partial g_{\alpha\beta,\gamma}}\right) , \tag{3.412}$$

and one has taken into account that

$$\Gamma^\mu_{\gamma\mu} = \frac{1}{\sqrt{-g}}\frac{\partial\sqrt{-g}}{\partial x^\gamma} . \tag{3.413}$$

Because the metric tensor and its first partial derivatives enter the symbols Γ_α in the covariant derivatives for the matter variables, the separate terms entering equations (3.411) and (3.412) should be treated as follows

$$\frac{\partial L_M}{\partial g_{\alpha\beta}} = \left(\frac{\partial L_M}{\partial g_{\alpha\beta}}\right)_{\nabla\psi=\text{fixed}} + \frac{\partial L_M}{\partial\psi_{,\rho}}\frac{\partial\Gamma_\rho}{\partial g_{\alpha\beta}}\psi , \tag{3.414}$$

$$\frac{\partial L_M}{\partial g_{\alpha\beta,\gamma}} = \frac{\partial L_M}{\partial\psi_{,\rho}}\frac{\partial\Gamma_\rho}{\partial g_{\alpha\beta,\gamma}}\psi , \tag{3.415}$$

where the first term in equation (3.414) is calculated for fixed value of the covariant derivative of the filed ψ. Calculation of the derivative from Γ_ρ can be done if one notices that for a material field ψ of any tensorial rank, the following relationship holds

$$\Gamma_\rho\psi = \Gamma^\nu_{\rho\mu}\mathfrak{D}^\mu{}_\nu\psi , \tag{3.416}$$

where $\Gamma^\nu_{\rho\mu}$ is the Christoffel symbol and the constant tensor $\mathfrak{D}^\mu{}_\nu$ has been defined in equation (3.383). Equation (3.416) reduces calculation of the partial derivatives from Γ_ρ to that from the Christoffel symbols. Making use of equation (3.345a), (3.345b) and (3.354) results in

$$\frac{\partial\Gamma_\rho}{\partial g_{\alpha\beta}} = -\frac{1}{2}\left(\Gamma^\alpha_{\rho\mu}\mathfrak{D}^{\mu\beta} + \Gamma^\beta_{\rho\mu}\mathfrak{D}^{\mu\alpha}\right) , \tag{3.417}$$

$$\frac{\partial\Gamma_\rho}{\partial g_{\alpha\beta,\gamma}} = \frac{1}{2}\left[\delta^\alpha_\rho\mathfrak{D}^{[\gamma\beta]} + \delta^\beta_\rho\mathfrak{D}^{[\gamma\alpha]} + \delta^\gamma_\rho\mathfrak{D}^{(\alpha\beta)}\right] , \tag{3.418}$$

where indices are raised and lowered with the metric tensor $g_{\alpha\beta}$, while the square/round brackets around a pair of indices denote a full anti-symmetry/symmetry with respect to the exchange of the indices.

After accounting for these results in the expression (3.414) and replacing partial derivatives $\partial L_M/\psi_{,\alpha}$ with their covariant counterpart according to equation (3.324), the metrical tensor of energy-momentum of matter becomes:

$$T^{\alpha\beta} = -g^{\alpha\beta}L_M - 2\frac{\partial L_M}{\partial g_{\alpha\beta}} + \nabla_\gamma\left\{\sigma^{\alpha[\gamma\beta]} + \sigma^{\beta[\gamma\alpha]} + \sigma^{\gamma(\alpha\beta)}\right\} , \tag{3.419}$$

where the partial derivative with respect to the metric tensor must be taken for fixed value of the Christoffel symbols, and one has introduced a shorthand notation for a new tensor

$$\sigma^{\alpha\beta}{}_\gamma = \frac{\partial L_M}{\partial\nabla_\alpha\psi}\mathfrak{D}^\beta{}_\gamma\psi , \qquad \sigma^{\alpha\beta\gamma} = g^{\gamma\mu}\sigma^{\alpha\beta}{}_\mu . \tag{3.420}$$

Expression (3.419) can be compared with (3.411) by making use of an exact relationship

$$\begin{aligned}\nabla_\gamma\left\{\sigma^{\alpha[\gamma\beta]}+\sigma^{\beta[\gamma\alpha]}+\sigma^{\gamma(\alpha\beta)}\right\} &= \partial_\gamma\left\{\sigma^{\alpha[\gamma\beta]}+\sigma^{\beta[\gamma\alpha]}+\sigma^{\gamma(\alpha\beta)}\right\}\\ &+ \Gamma^\mu_{\gamma\mu}\left\{\sigma^{\alpha[\gamma\beta]}+\sigma^{\beta[\gamma\alpha]}+\sigma^{\gamma(\alpha\beta)}\right\}\\ &+ \Gamma^\alpha_{\gamma\mu}\sigma^{\gamma\mu\beta}+\Gamma^\beta_{\gamma\mu}\sigma^{\gamma\mu\alpha}\,, \end{aligned}\tag{3.421}$$

Formula (3.419) offers a straightforward way of calculation of the metrical energy-momentum tensor from matter's Lagrangian L_M. Nevertheless, some subtleties are still involved. Fact of the matter is that calculation of the partial derivative $\partial L_M/\partial g_{\alpha\beta}$ assumes that the field variables ψ do not depend on the metric tensor implicitly. Therefore, each time prior to performing the partial differentiation of L_M, one has to make sure this requirement is fulfilled. For example, if the independent dynamic variable is a co-vector, A_α, the corresponding vector variable, A^α, depends on metric that participates in rising the index, $A^\alpha = g^{\alpha\beta}A_\beta$ but in most cases this equation is not used explicitly. If this implicit dependence of matter' variables on metric is overlooked, a confusion may arise. In some cases the matter variable may include a functional dependence of the matter variable on metric, $\psi = \psi(g_{\alpha\beta})$. It happens, for example, in case of the rest-mass density of matter which is referred to a three-dimensional volume that depends on metric.

3.9.5.2 ***Gauge Invariance of the Metrical Energy Momentum Tensor***

Einstein's theory of gravity admits only extrinsic gauge symmetry associated with diffeomorphisms. For this reason, the only gauge variation of the metric tensor, $\delta g_{\alpha\beta}$, is that given in equations (3.394) and (3.395) in terms of the Lie derivative along vanishing on the boundary of integration but otherwise *arbitrary* vector field ξ^α. Let us derive equations of motion of matter from the gauge invariance of the total action (3.298). The gauge variation of the action with respect to the metric tensor vanishes, yielding

$$\delta S = \frac{1}{2\kappa}\int\left(G^{\alpha\beta}-\kappa T^{\alpha\beta}\right)\nabla_\alpha\xi_\beta\sqrt{-g}d^4x = 0\,, \tag{3.422}$$

where one has used equation (3.394) for the gauge variation of the metric tensor. Integrating by parts the Einstein tensor in equation (3.422) and taking into account that on the boundary of integration the diffeomorphism $\xi^\alpha = 0$, result in

$$\int G^{\alpha\beta}\nabla_\alpha\xi_\beta\sqrt{-g}d^4x = \oint G^{\alpha\beta}\xi_\alpha d\Sigma_\beta - \int\xi_\alpha\nabla_\beta G^{\alpha\beta}\sqrt{-g}d^4x = 0\,, \tag{3.423}$$

because the surface integral vanishes due to the boundary conditions, and $\nabla_\beta G^{\alpha\beta}\equiv 0$ due to the Bianchi identity. The equation (3.422) is reduced then to

$$\int T^{\alpha\beta}\nabla_\alpha\xi_\beta\sqrt{-g}d^4x = \oint T^{\alpha\beta}\xi_\alpha d\Sigma_\beta - \int\xi_\alpha\nabla_\beta T^{\alpha\beta}\sqrt{-g}d^4x = 0\,. \tag{3.424}$$

The surface integral disappears because of the boundary conditions imposed on ξ^α. Since vector field ξ^α is arbitrary inside the domain of integration the integral equation

(3.423) has the only solution

$$\nabla_\beta T^{\alpha\beta} = 0 \,, \tag{3.425}$$

that is the covariant equation of motion of matter. These equations are valid, if and only if, the metric tensor satisfies the Einstein field equations. Hence, matter tells the spacetime manifold how to curve, and the geometry of the manifold tells the matter how to move. This universal reciprocity of matter and gravity makes investigation of the problem of motion of extended bodies in general relativity be a rather challenging exercise which is discussed in next sections. Covariant form of equation (3.425) fetches out that the metrical energy-momentum tensor is gauge-invariant.

3.9.5.3 *Electromagnetic Energy-Momentum Tensor*

The independent dynamic variable of electromagnetic field is a co-vector, $A_\alpha = (A_0, A_i) = (\phi, \boldsymbol{A})$, that defines electromagnetic potentials and fields in the form of the electromagnetic tensor

$$F_{\alpha\beta} = \nabla_\alpha A_\beta - \nabla_\beta A_\alpha \,, \tag{3.426}$$

Physical meaning of the components of the electromagnetic tensor have been explained in section 2.7.4. Covariant derivatives in $F_{\alpha\beta}$ can be replaced with partial derivatives due to antisymmetry but only if the spacetime has no additional contribution to the affine connection coming from torsion and/or nonmetricity. In metric-based theories of gravity, including general relativity, the terms with the Christoffel symbols in equation (3.426) are completely canceled out. Hence, the covariant derivatives in the expression for the electromagnetic tensor are irrelevant but they will be formally kept in definition (3.426).

Lagrangian for free electromagnetic field is

$$L_M = \frac{1}{16\pi} g^{\rho\sigma} g^{\mu\nu} F_{\rho\mu} F_{\sigma\nu} \,, \tag{3.427}$$

and it disentangles the metric tensor from the field variables. One is allowed, of course, to write the Lagrangian in the form $L_M \sim g_{\rho\sigma} g_{\mu\nu} F^{\rho\mu} F^{\sigma\nu}$ but this is inconvenient since tensor $F^{\sigma\nu}$ has two indices raised with the metric tensor which, thus, enters implicitly the latter expression and complicates partial differentiation with respect to metric.

The electromagnetic Lagrangian is substituted to definition of the metrical energy-momentum tensor (3.419). Tensor $\sigma^{\alpha\beta\gamma}$ is calculated by taking partial derivatives from the electromagnetic tensor (3.426) and accounting that $\mathfrak{D}^\mu{}_\nu A_\beta = -\delta^\mu_\beta A_\nu$. It yields

$$\sigma^{\alpha\beta\gamma} = -\frac{1}{4\pi} F^{\alpha\beta} A^\gamma \,, \tag{3.428}$$

that is antisymmetric with respect to the first two indices, $\sigma^{\alpha\beta\gamma} = -\sigma^{\beta\alpha\gamma}$. It results in mutual algebraic cancelation of all terms with $\sigma^{\alpha\beta\gamma}$ entering equation (3.419) in agreement with one's observation that the electromagnetic tensor does not entail the

Christoffel symbols in general relativity. The remaining terms in definition (3.419) bring about the metrical energy-momentum tensor of electromagnetic field

$$T^{\alpha\beta} = \frac{1}{4\pi}\left(F^{\alpha\mu}F^{\beta}{}_{\mu} - \frac{1}{4}g^{\alpha\beta}F^{\mu\nu}F_{\mu\nu}\right) , \tag{3.429}$$

which should be compared with equation (2.381) for this tensor in the Minkowski spacetime. As expected, the only difference between the two expressions is in the replacement of the Minkowski metric $\eta^{\alpha\beta}$ with the full metric $g^{\alpha\beta}$.

The metrical energy-momentum tensor of free electromagnetic field is identically traceless

$$g_{\alpha\beta}T^{\alpha\beta} = 0 . \tag{3.430}$$

The reason is that free electromagnetic field is represented by an electromagnetic wave propagating in vacuum with fundamental speed c on a hypersurface of null cone. Therefore, the energy-momentum tensor is proportional to a direct tensor product of the null vectors, k^α, corresponding to the direction of propagation of the wave, $T^{\alpha\beta} = (a^2/8\pi)k^\alpha k^\beta$, where a is the scalar amplitude of the wave. Trace of such tensor is always zero since $k_\alpha k^\alpha = 0$ (for more detail see [Frolov, 1979; Misner *et al.*, 1973]). Any field propagating in vacuum with the speed c has the same property of vanishing trace of the energy-momentum tensor [Buchdahl, 1959].

3.9.5.4 *Energy-Momentum Tensor of a Perfect Fluid*

Consider a *perfect fluid*, i.e., one lacking viscosity or heat conduction. Its energy density ϵ consists of two components,

$$\epsilon = \rho\left(c^2 + \Pi\right) , \tag{3.431}$$

where ρ is the rest-mass density defined as rest mass per three-dimensional volume taken in the frame being comoving with the fluid element, while $\Pi = \Pi(\rho)$ is the compression energy density *per unit rest mass*. Hence, one can chose the density ρ as a dynamic variable.

The standard choice for the fluid Lagrangian is [Brumberg, 1972; Fock, 1964; Hawking and Ellis, 1975]

$$L_M = \frac{\epsilon}{c^2} = \rho\left(1 + \frac{\Pi}{c^2}\right) , \tag{3.432}$$

which does not depend explicitly on derivatives of the dynamic variable ρ. Hence, terms with $\sigma^{\alpha\beta\gamma}$ do not appear in the metrical energy-momentum tensor of the fluid,

$$\begin{aligned} c^2T^{\alpha\beta} &= -g^{\alpha\beta}\epsilon - 2\frac{\partial\epsilon}{\partial g_{\alpha\beta}} \\ &= -2\left(c^2 + \Pi + \rho\frac{d\Pi}{d\rho}\right)\frac{\partial\rho}{\partial g_{\alpha\beta}} - \rho(c^2 + \Pi)g^{\alpha\beta} . \end{aligned} \tag{3.433}$$

One is challenged with taking a partial derivative from the rest-mass density, ρ, of the fluid with respect to the metric tensor. The rest-mass density does not depend on the metric tensor explicitly but it does not mean that there is no such dependence at all. The density is defined as the ratio of the rest mass of baryons to a volume element of the comoving frame. The volume element depends on the metric tensor as discussed in section 3.9.1 and the derivative $\partial\rho/\partial g_{\alpha\beta}$ is not zero.

In order to calculate this derivative, let us use the fundamental law of conservation of baryons that is equivalent to the covariant equation of continuity of the perfect fluid written as [Misner *et al.*, 1973, Equation 22.1]

$$\nabla_\alpha\left(\rho u^\alpha\right) = \frac{1}{\sqrt{-g}}\partial_\alpha\left(\sqrt{-g}\rho u^\alpha\right) = 0\,, \tag{3.434}$$

where $u^\alpha = dx^a/d\tau$ is four-velocity of fluid's element. This equation suggests that the infinitesimal variation of density is related to variations of the volume and four-velocity by equation

$$\delta\left(\sqrt{-g}\rho u^\alpha\right) = 0\,, \tag{3.435}$$

wherefrom one obtains

$$\rho u^\alpha\delta\sqrt{-g} + u^\alpha\sqrt{-g}\delta\rho + \delta u^\alpha\rho\sqrt{-g} = 0\,. \tag{3.436}$$

Making transvection of this equation with u_α, and taking into account that $u_\alpha u^\alpha = -c^2$, one gets

$$\sqrt{-g}\delta\rho + \rho\delta\sqrt{-g} - \frac{\sqrt{-g}}{c^2}\rho u_\alpha\delta u^\alpha = 0\,, \tag{3.437}$$

which entails

$$\delta\rho = \frac{1}{c^2}\rho u_\alpha\delta u^\alpha - \rho\frac{\delta\sqrt{-g}}{\sqrt{-g}}\,. \tag{3.438}$$

Varying the normalization condition for four-velocity, $u_\alpha u^\alpha = g_{\alpha\beta}u^\alpha u^\beta = -c^2$, one gets

$$u_\alpha\delta u^\alpha = -\frac{1}{2}u^\alpha u^\beta\delta g_{\alpha\beta}\,. \tag{3.439}$$

Insertion of equations (3.439) and (3.352) in equation (3.438) yield

$$\frac{\partial\rho}{\partial g_{\alpha\beta}} = -\frac{\rho}{2c^2}\left(u^\alpha u^\beta + c^2 g^{\alpha\beta}\right)\,. \tag{3.440}$$

This result should be substituted to equation (3.433) where a thermodynamic relationship (2.377) should be also used

$$\rho\frac{d\Pi}{d\rho} = \frac{p}{\rho}\,. \tag{3.441}$$

All together, the metrical energy-momentum tensor of the perfect fluid is obtained

$$c^2T^{\alpha\beta} = (\epsilon + p)\,u^\alpha u^\beta + c^2 pg^{\alpha\beta}\,, \tag{3.442}$$

or, more explicitly,

$$T^{\alpha\beta} = \rho\left(1 + \frac{\Pi}{c^2}\right)u^\alpha u^\beta + \frac{p}{c^2}\left(u^\alpha u^\beta + c^2 g^{\alpha\beta}\right) . \tag{3.443}$$

This equation generalizes special-relativistic expression (2.371) of the energy-momentum tensor of the perfect fluid to the case of a curved spacetime.

Notice that projection of the energy-momentum tensor of the perfect fluid on its four-velocity is positively-defined

$$T^{\alpha\beta}u_\alpha u_\beta \geq 0 , \tag{3.444}$$

This inequality is often imposed on the energy-momentum tensor, especially in cases when the form of the tensor is not known a priory. It is called the *weak energy condition*. One can demonstrate that it is satisfied also by the Maxwell field and by other realistic *classical* fields encountered in physics. In the quantum field theory, the weak energy condition often gets violated. This happens because the vacuum energy density can assume negative values [Mukhanov, 2005]. It is worth mentioning that condition (3.444) is imposed pointwise, i.e., in tangent space. Being helpful in proving local theorems, the weak energy condition is unrelated to global (topological) features of the manifold.

The trace of the energy-momentum tensor

$$T^{\alpha\beta}g_{\alpha\beta} = 3p - \epsilon , \tag{3.445}$$

and is negative for macroscopic bodies, $p < \epsilon/3$. In case of ultra-relativistic perfect fluid, velocity of the fluid is approaching to the fundamental speed c, and the trace of the energy-momentum tensor approaches zero. Relationship between pressure and energy density for ultra-relativistic fluid is

$$p = \frac{\epsilon}{3} , \tag{3.446}$$

which is used in cosmology as an equation of state of the radiation-dominating epoch in the history of the expanding universe [Mukhanov, 2005]

3.9.5.5 ***Energy-Momentum Tensor of a Scalar Field***

Lagrangian of a scalar field ϕ is given by equation

$$L_M = \frac{1}{2}g^{\alpha\beta}\nabla_\alpha\phi\nabla_\beta\phi - V(\phi) , \tag{3.447}$$

where $V(\phi)$ is the potential energy of the field, and one has taken ϕ and $\nabla_\alpha\phi$ as independent dynamic variables which have no implicit dependence on the metric tensor. The metrical energy-momentum tensor of the scalar field is

$$T^{\alpha\beta} = -g^{\alpha\beta}L_M - 2\frac{\partial L_M}{\partial g_{\alpha\beta}} , \tag{3.448}$$

as the covariant derivative of a scalar field is simply a partial derivative, $\nabla_\alpha\phi = \partial_\alpha\phi$, and there is no contribution from tensor $\sigma^{\alpha\beta\gamma}$. Calculation of partial derivative with respect to the metric tensor is straightforward. The energy-momentum tensor is obtained in the following form

$$T^{\alpha\beta} = \partial^\alpha\phi\partial^\beta\phi - \frac{1}{2}g^{\alpha\beta}\partial^\mu\phi\partial_\mu\phi + g^{\alpha\beta}V(\phi) \tag{3.449}$$

which agrees with its special-relativistic expression (2.389).

One notices that the Lagrangian for a scalar field looks similar to the Lagrangian for a particle in classic mechanics: the first term in equation (3.447) is the kinetic energy of the field, and the second term is the potential energy of the field. The trace of the energy-momentum tensor is

$$T^{\alpha\beta}g_{\alpha\beta} = -g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi + 4V(\phi)\ . \tag{3.450}$$

The trace is not zero if the field potential energy $V(\phi) \neq 0$. Only in the case, when the field has no potential and propagates freely, its trace is zero as a consequence of the field equations [Mukhanov, 2005].

3.9.6
The canonical energy-momentum tensor

3.9.6.1 *Definition*

One has introduced definition (3.337) of the *metrical* tensor of energy-momentum for matter. The Noether theorem proposes another definition of the matter tensor of energy-momentum which is called *canonical*. We shall derive the canonical energy-momentum tensor, $mathfrakT^{\alpha\beta}$, in this section and establish its relationship to the metrical tensor of energy-momentum, $T^{\alpha\beta}$.

Let us again assume for simplicity that the Lagrangian of matter contains only first derivatives of matter variables: $L_M = L_M\left(\psi; \psi_{,\alpha} + \Gamma_a\psi; g_{\alpha\beta}\right)$. Second and higher-order derivatives can be also included [Szabados, 2009] but they perplex calculations without bringing about essentially new physical information.

Let ξ^α be an *arbitrary* vector field on curved spacetime manifold and $\mathcal{L}_M = \sqrt{-g}L_M$ is a scalar density of the Lagrangian. Lie derivative of the Lagrangian density along the vector field is

$$\pounds_{\boldsymbol{\xi}}\mathcal{L}_M = \frac{\partial\mathcal{L}_M}{\partial\psi}\pounds_{\boldsymbol{\xi}}\psi + \frac{\partial\mathcal{L}_M}{\partial\psi_{,\alpha}}\pounds_{\boldsymbol{\xi}}\psi_{,\alpha} + \frac{\partial\mathcal{L}_M}{\partial g_{\alpha\beta}}\pounds_{\boldsymbol{\xi}}g_{\alpha\beta} + \frac{\partial\mathcal{L}_M}{\partial g_{\alpha\beta,\gamma}}\pounds_{\boldsymbol{\xi}}g_{\alpha\beta,\gamma}\ . \tag{3.451}$$

Since the Lie derivative commutes with partial derivative

$$\pounds_{\boldsymbol{\xi}}\psi_{,\alpha} = (\pounds_{\boldsymbol{\xi}}\psi)_{,\alpha}\ , \qquad \pounds_{\boldsymbol{\xi}}g_{\alpha\beta,\gamma} = (\pounds_{\boldsymbol{\xi}}g_{\alpha\beta})_{,\gamma}\ , \tag{3.452}$$

one is allowed to re-write equation (3.451) in terms of the variational derivatives and a total divergence,

$$\pounds_{\boldsymbol{\xi}}\mathcal{L}_M = \frac{\delta\mathcal{L}_M}{\delta\psi}\pounds_{\boldsymbol{\xi}}\psi + \frac{\delta\mathcal{L}_M}{\delta g_{\alpha\beta}}\pounds_{\boldsymbol{\xi}}g_{\alpha\beta} + \frac{\partial}{\partial x^\gamma}\left(\frac{\partial\mathcal{L}_M}{\partial\psi_{,\gamma}}\pounds_{\boldsymbol{\xi}}\psi + \frac{\partial\mathcal{L}_M}{\partial g_{\alpha\beta,\gamma}}\pounds_{\boldsymbol{\xi}}g_{\alpha\beta}\right)\ . \tag{3.453}$$

Next step is to use equations of motion of matter (3.322), definition of the metrical tensor of energy-momentum (3.337), and expression (3.415), which is also valid for the Lagrangian density

$$\frac{\partial \mathcal{L}_M}{\partial g_{\alpha\beta,\gamma}} = \frac{\partial \mathcal{L}_M}{\partial \psi_{,\rho}} \frac{\partial \Gamma_\rho}{\partial g_{\alpha\beta,\gamma}} \psi \ . \tag{3.454}$$

Then, equation (3.453) takes on the following form,

$$\pounds_{\boldsymbol{\xi}} \mathcal{L}_M = -\frac{\sqrt{-g}}{2} T^{\alpha\beta} \pounds_{\boldsymbol{\xi}} g_{\alpha\beta} + \frac{\partial}{\partial x^\gamma} \left[\frac{\partial \mathcal{L}_M}{\partial \psi_{,\rho}} \left(\delta^\gamma_\rho \pounds_{\boldsymbol{\xi}} \psi + \frac{\partial \Gamma_\rho}{\partial g_{\alpha\beta,\gamma}} \psi \pounds_{\boldsymbol{\xi}} g_{\alpha\beta} \right) \right] . \tag{3.455}$$

This equation can be further transformed after observing that

$$\pounds_{\boldsymbol{\xi}} \mathcal{L}_M = \frac{\partial}{\partial x^\gamma} \left(\xi^\gamma \mathcal{L}_M \right) , \tag{3.456}$$

$$\frac{\sqrt{-g}}{2} T^{\alpha\beta} \pounds_{\boldsymbol{\xi}} g_{\alpha\beta} = \frac{\partial}{\partial x^\gamma} \left(\sqrt{-g} T^{\beta\gamma} \xi_\beta \right) , \tag{3.457}$$

where one has used definition of the Lie derivative from the metric tensor (3.394), the symmetry and the law of conservation of the metrical tensor of energy-momentum (3.425) as well as equation (3.362). Now equation (3.455) can be written as a law of conservation

$$\frac{\partial}{\partial x^\gamma} \left[\sqrt{-g} T^{\gamma\beta} \xi_\beta + \xi^\gamma \mathcal{L}_M - \frac{\partial \mathcal{L}_M}{\partial \psi_{,\rho}} \left(\delta^\gamma_\rho \pounds_{\boldsymbol{\xi}} \psi + \frac{\partial \Gamma_\rho}{\partial g_{\alpha\beta,\gamma}} \psi \pounds_{\boldsymbol{\xi}} g_{\alpha\beta} \right) \right] = 0 , \tag{3.458}$$

where the term in the round parentheses yet has to be simplified.

Calculation of the derivative from Γ_ρ is performed with the help of equation (3.418) Making use of that equation along with equation (3.381) for the Lie derivative of the field ψ, and doing some algebra, one obtains

$$\frac{\partial}{\partial x^\gamma} \left\{ \sqrt{-g} \left[\left(T^\gamma{}_\beta - \mathfrak{T}^\gamma{}_\beta \right) \xi^\beta - \left(\sigma^{\alpha[\gamma\beta]} + \sigma^{\beta[\gamma\alpha]} + \sigma^{\gamma[\alpha\beta]} \right) \boldsymbol{\nabla}_\beta \xi_\alpha \right] \right\} , \tag{3.459}$$

where

$$\mathfrak{T}^\gamma{}_\beta = -\delta^\gamma_\beta L_M + \frac{\partial L_M}{\partial \boldsymbol{\nabla}_\gamma \psi} \boldsymbol{\nabla}_\beta \psi , \tag{3.460}$$

$$\sigma^{\gamma[\alpha\beta]} = -\frac{\partial L_M}{\partial \boldsymbol{\nabla}_\gamma \psi} g^{\mu[\alpha} \mathfrak{D}^{\beta]}{}_\mu \psi \ . \tag{3.461}$$

Tensor $\mathfrak{T}^\gamma{}_\beta$ is called the *canonical energy-momentum tensor*, while $\sigma^{\gamma[\alpha\beta]}$ is the *tensor of spin* also called *tensor of helicity*. Tensor of helicity has already appeared in definition (3.419) of the metrical energy-momentum tensor of matter.

3.9.6.2 ***Relationship to the Metrical Energy-Momentum Tensor***

Relationship between the canonical and metrical tensors of energy-momentum of matter are derived from equation (3.459). Let us introduce a new notation

$$S^{\alpha\beta\gamma} \equiv \sigma^{\alpha[\gamma\beta]} + \sigma^{\beta[\gamma\alpha]} + \sigma^{\gamma[\alpha\beta]} \ . \tag{3.462}$$

Then, the helicity terms in equation (3.459) can be reshuffled as

$$\partial_\gamma\left(\sqrt{-g}S^{\alpha\beta\gamma}\nabla_\beta\xi_\alpha\right)=\partial_\gamma\left[\sqrt{-g}\nabla_\beta\left(S^{\alpha\beta\gamma}\xi_\alpha\right)\right]-\partial_\gamma\left(\sqrt{-g}\xi_\alpha\nabla_\beta S^{\alpha\beta\gamma}\right)\;. \tag{3.463}$$

Transvection $S^{\alpha\beta\gamma}\xi_\alpha$ is a second rank tensor that is fully anti-symmetric with respect to its two free indices, $S^{\alpha\beta\gamma}\xi_\alpha=-S^{\alpha\gamma\beta}\xi_\alpha$. Therefore,

$$\sqrt{-g}\nabla_\beta\left(S^{\alpha\beta\gamma}\xi_\alpha\right)=\partial_\beta\left(S^{\alpha\beta\gamma}\xi_\alpha\right)\;, \tag{3.464}$$

and

$$\partial_\gamma\left[\sqrt{-g}\nabla_\beta\left(S^{\alpha\beta\gamma}\xi_\alpha\right)\right]=\frac{\partial^2}{\partial x^\beta\partial x^\gamma}\left(S^{\alpha\beta\gamma}\xi_\alpha\right)\equiv 0\;. \tag{3.465}$$

It cuts equation (3.463) down to

$$\partial_\gamma\left[\sqrt{-g}\left(S^{\alpha\beta\gamma}\right)\nabla_\beta\xi_\alpha\right]=-\partial_\gamma\left[\sqrt{-g}\xi_\alpha\nabla_\beta S^{\alpha\beta\gamma}\right]\;, \tag{3.466}$$

which allows us to simplify the law of conservation (3.459) to

$$\frac{\partial}{\partial x^\gamma}\left\{\sqrt{-g}\xi_\beta\left[T^{\gamma\beta}-\mathfrak{T}^{\gamma\beta}-\nabla_\alpha\left(\sigma^{\gamma[\alpha\beta]}+\sigma^{\beta[\alpha\gamma]}+\sigma^{\alpha[\beta\gamma]}\right)\right]\right\}=0\;. \tag{3.467}$$

The terms enclosed to the curl brackets is a vector density so that the above equation can be re-written in terms of a covariant law of conservation,

$$\nabla_\gamma\left\{\xi_\beta\left[T^{\gamma\beta}-\mathfrak{T}^{\gamma\beta}-\nabla_\alpha\left(\sigma^{\gamma[\alpha\beta]}+\sigma^{\beta[\alpha\gamma]}+\sigma^{\alpha[\beta\gamma]}\right)\right]\right\}=0\;, \tag{3.468}$$

that is another expression of the Noether theorem pointing out to the existence of a conserved current

$$J^\gamma=\left[T^{\gamma\beta}-\mathfrak{T}^{\gamma\beta}-\nabla_\alpha\left(\sigma^{\gamma[\alpha\beta]}+\sigma^{\beta[\alpha\gamma]}+\sigma^{\alpha[\beta\gamma]}\right)\right]\xi_\beta\;, \tag{3.469}$$

where vector field ξ^β is *arbitrary*. If spacetime has no symmetries, the Noether current $J^\gamma=0$, and equation (3.469) has a non-trivial solution

$$T^{\gamma\beta}=\mathfrak{T}^{\gamma\beta}+\nabla_\alpha\left(\sigma^{\gamma[\alpha\beta]}+\sigma^{\beta[\alpha\gamma]}+\sigma^{\alpha[\beta\gamma]}\right)\;, \tag{3.470}$$

that relates the metrical and canonical tensors of energy-momentum with the helicity tensor of matter [Belinfante, 1939, 1940; Rosenfeld, 1940].

The canonical energy-momentum tensor $\mathfrak{T}^{\alpha\beta}$ is not symmetric in general case. On the other hand, the metrical energy-momentum tensor $T^{\alpha\beta}$ is always symmetric. It imposes some limitations on the components of the canonical tensor known as the Belinfante-Rosenfeld identities. Specifically, subtracting from equation (3.470) its fully-symmetric part, one obtains an algebraic constrain

$$\mathfrak{T}^{[\beta\gamma]}=\nabla_\alpha\sigma^{\alpha[\beta\gamma]}\;. \tag{3.471}$$

This equation tells us that the antisymmetric part of the canonical energy-momentum tensor is intimately related to the helicity tensor. Taking a covariant derivative from previous equation yields a differential relationship

$$\nabla_\nu\mathfrak{T}^\nu{}_\mu=R_{\mu\alpha\beta\gamma}\sigma^{\alpha\beta\gamma}\;. \tag{3.472}$$

One can notice that if the helicity tensor vanishes the metrical and canonical tensors coincide. In the presence of a non-trivial matter's helicity, the metrical and canonical tensors are different. Classical concept of the helicity tensor is closely related to the concept of particle's spin in quantum theory. Scalar particles have spin zero, while photons are bosons with spin 1. Gravitational field in a linearized approximation is described by bosons with spin 2, which are termed *gravitons*. Due to the analogy between spin and helicity, one has to expect that helicity of the classical scalar field is zero, but it does not vanish for electromagnetic and gravitational field. We discuss this issue below.

3.9.6.3 *Killing Vectors and the Global Laws of Conservation*

Let us assume that spacetime has a global symmetry associated with existence of a Killing vector, ξ^β, of the metric tensor. Killing vector satisfies equation (3.159), that is $\nabla_{(\alpha}\xi_{\beta)} = 0$. Since the metrical energy-momentum tensor is conserved independently, $\nabla_\beta T^{\beta\gamma} = 0$, the Noether current, corresponding to the Killing vector ξ^β is reduced to

$$J^\gamma = \left[\mathfrak{T}^{\gamma\beta} + \nabla_\alpha\left(\sigma^{\gamma[\alpha\beta]} + \sigma^{\beta[\alpha\gamma]} + \sigma^{\alpha[\beta\gamma]}\right)\right]\xi_\beta\,. \tag{3.473}$$

It is more convenient to reassemble this equation by noticing that for any antisymmetric tensor of a second rank, $A^{\alpha\beta} = A^{[\alpha\beta]}$, the following identity is fulfilled, $\nabla_\alpha\nabla_\beta A^{\alpha\beta} = R_{\alpha\beta}A^{[\alpha\beta]} \equiv 0$. It brings the Noether current (3.473) to the following form

$$J^\gamma = \mathfrak{T}^{\gamma\beta}\xi_\beta - \sigma^{\gamma[\alpha\beta]}\nabla_\alpha\xi_\beta\,. \tag{3.474}$$

Physical interpretation of the Noether currents in spacetimes with isometries is a non-trivial problem which continue to attract peer attention of various researchers (see, for example, [Deruelle *et al.*, 2004; Fatibene *et al.*, 2010; Petrov and Katz, 2002; Uzan *et al.*, 1998], and references therein).

The case of Killing vectors corresponding to isometries of flat Minkowski geometry admits unambiguous and almost obvious interpretation. Four Killing vectors ξ_β, corresponding to global translations, have components $\xi^\alpha_\beta = \delta^\alpha_\beta$ [81]. In this case, the Noether current (3.474) is reduced to $J^\alpha_\beta = \mathfrak{T}^\alpha{}_\beta$ and represents a linear momentum of the matter field. Lorentz boosts corresponding to six Killing vectors $\xi^\alpha_{[\beta\gamma]} = x^\nu\left(g_{\nu\beta}\delta^\alpha_\gamma - g_{\nu\gamma}\delta^\alpha_\beta\right)$. The corresponding Noether current (3.474) is the total angular momentum of matter:

$$J^\alpha_{[\beta\gamma]} = g_{\beta\mu}g_{\gamma\nu}\left(\mathfrak{T}^{\alpha[\mu}x^{\nu]} + \sigma^{\alpha[\mu\nu]}\right)\,. \tag{3.475}$$

The total angular momentum consists of two parts - the *orbital angular momentum* of $\mathfrak{T}^{\alpha\beta}$ and the helicity (3.461) also known as the *spin angular momentum*.

81) Subscript index numerates vectors, while superscript index numerates the components of the vector.

3.9.6.4 *The Canonical Energy-Momentum Tensor for Electromagnetic Field*

Lagrangian for electromagnetic field was defined in equation (3.427). The dynamic matter variable is the electromagnetic co-vector A_α and its derivative $A_{\alpha,\beta}$. Making use of equation (3.460) one obtains the canonical energy-momentum tensor of electromagnetic field

$$\mathfrak{T}^{\alpha\beta} = \frac{1}{4\pi}\left(F^{\alpha\gamma}\boldsymbol{\nabla}^\beta A_\gamma - \frac{1}{4}g^{\alpha\beta}F^{\mu\nu}F_{\mu\nu}\right) , \tag{3.476}$$

that is apparently non-symmetric because of a non-vanishing helicity tensor

$$\sigma^{\alpha[\beta\gamma]} = -\frac{1}{4\pi}F^{\alpha[\beta}A^{\gamma]} . \tag{3.477}$$

Calculating covariant derivative from the helicity tensor, one finds out that

$$\boldsymbol{\nabla}_\gamma\left(\sigma^{\alpha[\gamma\beta]} + \sigma^{\beta[\gamma\alpha]} + \sigma^{\gamma[\beta\alpha]}\right) = -\frac{1}{4\pi}F^{\alpha\gamma}\boldsymbol{\nabla}_\gamma A^\beta . \tag{3.478}$$

Adding this result to the canonical energy-momentum tensor (3.476) brings back the metrical energy-momentum tensor of electromagnetic field (3.429). The Belinfante-Rosenfeld equation (3.471) can be checked easily if one bears in mind that transvection $F^{\mu\alpha}F^\beta{}_\mu = F^{\mu(\alpha}F^{\beta)}{}_\mu$ is fully-symmetric. Antisymmetric part of this transvection vanishes identically leading to equality

$$\mathfrak{T}^{[\alpha\beta]} = -\frac{1}{4\pi}F^{\gamma[\alpha}\boldsymbol{\nabla}^{\beta]}A_\gamma = -\frac{1}{4\pi}F^{\gamma[\alpha}\boldsymbol{\nabla}_\gamma A^{\beta]} = \boldsymbol{\nabla}_\gamma\sigma^{\gamma[\alpha\beta]} , \tag{3.479}$$

quod erat demonstrandum.

3.9.6.5 *The Canonical Energy-Momentum Tensor for Perfect Fluid*

Derivation of the canonical energy-momentum tensor of a perfect fluid by making use of the Lagrangian (3.431) is not obvious since it does not depend explicitly on the partial derivative of the dynamic variable ρ - the rest-mass density of the fluid. The Lagrangian should be modified to reveal the presence of a dynamic variable and its partial derivative, which are used in the definition of the canonical energy-momentum tensor. This requires introduction of a new variable - a scalar potential, ϕ, of the perfect fluid that is a gradient of four-velocity of the fluid [Landau and Lifshitz, 1959, Equation 134.12]

$$\mu u_\alpha = -\boldsymbol{\nabla}_\alpha\phi , \tag{3.480}$$

where μ is the chemical potential of the fluid in a comoving frame.

The chemical potential, μ, is defined as the total energy required, per baryon, to inject a small additional amount of fluid into a given sample, without changing its entropy or volume. Entropy of a perfect fluid is constant and does not change as the fluid flows. Thus, the chemical potential μ is a unique function of the rest-mass density ρ like the compression energy, Π, and pressure, p. According to the definition [Misner *et al.*, 1973, Equation 22.8]

$$\mu = \frac{d\epsilon}{d\rho} , \tag{3.481}$$

where ϵ is the rest energy of the fluid element defined in equation (3.431). Taking the derivative and accounting for the second law of thermodynamics in the form of equation (3.441), one obtains an algebraic relationship between the thermodynamic variables

$$\epsilon = \rho\mu - p \, . \tag{3.482}$$

Let us consider the product $\rho\mu$. The chemical potential can be expressed in terms of the derivative from the scalar potential after contracting equation (3.480) with four-velocity of the fluid, and taking into account that $u_\alpha u^\alpha = -c^2$, where c is the fundamental speed of the Minkowski spacetime. It yields

$$\mu c^2 = u^\alpha \nabla_\alpha \phi \, , \tag{3.483}$$

which allows us to write the following sequence of equations

$$c^2 \rho\mu = \rho u^\alpha \nabla_\alpha \phi = \nabla_\alpha \left(\rho\phi u^\alpha\right) - \phi \nabla_\alpha \left(\rho u^\alpha\right) = \nabla_\alpha \left(\rho\phi u^\alpha\right) \, , \tag{3.484}$$

where one has used equation of continuity (3.434). Equation (3.484) demonstrates that the product $\rho\mu$ is a covariant divergence of a four-vector $\rho\phi u^\alpha$. Hence, the integral of action of a perfect fluid can be transformed as follows

$$\int \epsilon \sqrt{-g} d^4 x = \int (\rho\mu - p) \sqrt{-g} d^4 x = -\int p \sqrt{-g} d^4 x + \oint \rho\phi u^\alpha d\Sigma_\alpha \, , \tag{3.485}$$

where the surface integral can be discarded. We conclude that the Lagrangian of a perfect fluid can be also written in the form of pressure

$$L_M = -p = \epsilon - \rho\mu \, . \tag{3.486}$$

Both forms of the Lagrangians given in equations (3.432) and (3.486), are equivalent for variational analysis of motion of a perfect fluid.

Let us work at this time with the Lagrangian (3.486). One chooses the scalar potential ϕ as an independent dynamic variable. Its partial derivative $\nabla_\alpha \phi$ appears explicitly in the Lagrangian through the chemical potential. Indeed, contracting equation (3.480) with itself, one obtains

$$\mu = \frac{1}{c} \sqrt{-g^{\mu\nu} \nabla_\mu \phi \nabla_\nu \phi} \, , \tag{3.487}$$

which can be plugged to expression (3.486) bringing it down to the following form

$$L_M = -\frac{\rho(\phi)}{c} \sqrt{-g^{\mu\nu} \nabla_\mu \phi \nabla_\nu \phi} + \epsilon(\phi) \, , \tag{3.488}$$

where all thermodynamic functions depend on the scalar potential ϕ implicitly while the derivatives of the potential appear in the Lagrangian explicitly. We calculate a partial derivative of the Lagrangian

$$\frac{\partial L_M}{\partial \nabla_\alpha \phi} = -\frac{\rho u^\alpha}{c^2} \, , \tag{3.489}$$

where equations (3.480) and (3.487) have been used for making corresponding replacements of the variables. Making use of this result in definition (3.460) of the canonical energy-momentum tensor, one obtains

$$c^2\mathfrak{T}^{\alpha\beta} = (\epsilon + p)\, u^\alpha u^\beta + c^2 p g^{\alpha\beta}\ , \tag{3.490}$$

which agrees with the metrical tensor of energy-momentum (3.442). It was expected as the helicity tensor, $\sigma^{\alpha[\beta\gamma]} = 0$, for a perfect fluid.

The Lagrangian defined by equation (3.488) is more convenient for derivation of the metrical tensor of energy-momentum because the scalar potential ϕ does not depend on the metric tensor $g_{\alpha\beta}$ at all. Therefore, the only place where the metric tensor appear is the chemical potential (3.487). It is rather simple to prove that

$$\frac{\partial L_M}{\partial g_{\alpha\beta}} = -\rho\frac{\partial \mu}{\partial g_{\alpha\beta}} = -\frac{\rho\mu}{2c^2}u^\alpha u^\beta\ , \tag{3.491}$$

which reproduces the metrical tensor after substitution of this result to its definition (3.419).

It is also instructive to see what kind of equations of motion follow from the definition of the Lagrangian of a perfect fluid taken in form (3.488. Equations of motion of the fluid with the scalar potential ϕ as a dynamic variable, are given by variational equation

$$\frac{\delta L_M}{\delta\phi} = \frac{\partial L_M}{\partial\phi} - \boldsymbol{\nabla}_\alpha\left(\frac{\partial L_M}{\boldsymbol{\nabla}_\alpha\phi}\right) = 0\ . \tag{3.492}$$

In this equation

$$\frac{\partial L_M}{\partial\phi} = \frac{\partial\epsilon}{\partial\phi} - \frac{\partial\rho}{\partial\phi}\mu = \left(\frac{d\epsilon}{d\rho} - \mu\right)\frac{\partial\rho}{\partial\phi} = 0\ , \tag{3.493}$$

because of definition (3.481) of the chemical potential. On the other hand, making use of equation (3.489) elucidates that the partial derivative

$$\boldsymbol{\nabla}_\alpha\left(\frac{\partial L_M}{\partial\boldsymbol{\nabla}_\alpha\phi}\right) = -\frac{1}{c^2}\boldsymbol{\nabla}_\alpha\left(\rho u^\alpha\right) = 0\ , \tag{3.494}$$

because of the continuity equation (3.434).

3.9.7
Pseudotensor of Landau and Lifshitz

Formalism of the canonical energy-momentum tensor can be also applied for gravitational field. However, there is a drastic distinction between the case of matter and gravity. The reason is that the action of matter, S_M, depends on two different kind of dynamic variables - material fields ψ and gravitational field characterized by the metric tensor $g_{\alpha\beta}$. Gravitational action, S_H, depends only on the metric tensor and its derivatives making the canonical energy-momentum tensor of gravity effectively coinciding with the Einstein tensor $G_{\alpha\beta}$ which is zero in vacuum due to the Einstein equations. It means that the values of the energy and momentum of gravitational field taken at each point of spacetime manifold are nil - gravitational field is not localizable at one point. This is a consequence of the principle of equivalence. However, gravitational field has energy in finite volumes of space as well-known from the Newtonian gravity.

Several approaches exist in order to quantize gravitational energy in general relativity . One idea to build the, so-called, *pseudotensor* of gravitational field. This approach was originated by Einstein and further developed to perfection by Landau and Lifshitz [1975]. Another line of research is to chose a non-dynamic background metric [Babak and Grishchuk, 2000; Grishchuk, 2009; Grishchuk *et al.*, 1984; Popova and Petrov, 1988] or affine connection [Ferraris *et al.*, 2008] on the manifold, and to build a covariant formalism on this ground. We shall focus here on a more formal approach of pseudotensor.

We shall assume that the Lagrangian of gravitational field, L_H, depends on the metric tensor and its first and second derivatives. We define a *canonical* pseudotensor, $t^{\alpha}{}_{\beta}$, of gravitational field by equation

$$\sqrt{-g}t^{\alpha}{}_{\beta} \equiv -\mathcal{L}_H\delta^{\alpha}_{\beta} + \left[\frac{\partial\mathcal{L}_H}{\partial g_{\rho\sigma,\alpha}} - \frac{\partial}{\partial x^{\nu}}\left(\frac{\partial\mathcal{L}_H}{\partial g_{\rho\sigma,\alpha\nu}}\right)\right]\frac{\partial g_{\rho\sigma}}{\partial x^{\beta}} + \frac{\partial\mathcal{L}_H}{\partial g_{\rho\sigma,\alpha\nu}}\frac{\partial^2 g_{\rho\sigma}}{\partial x^{\beta}\partial x^{\nu}}\,, \tag{3.495}$$

where the Lagrangian density $\mathcal{L}_H = \sqrt{-g}L_H$. Taking divergence from both sides of this equation and using equation (3.332) for variational derivative from the Lagrangian, yield

$$\partial_\alpha\left(\sqrt{-g}t^{\alpha}{}_{\beta}\right) = -\frac{\delta\mathcal{L}_H}{\delta g_{\rho\sigma}}\frac{\partial g_{\rho\sigma}}{\partial x^{\beta}}\,. \tag{3.496}$$

Substituting equations of gravitational field (3.334) and using definition (3.337) of the metrical energy-momentum tensor of matter, $T^{\alpha\beta}$, result in

$$\partial_\alpha\left(\sqrt{-g}t^{\alpha}{}_{\beta}\right) = -\frac{1}{2}\sqrt{-g}T^{\rho\sigma}\frac{\partial g_{\rho\sigma}}{\partial x^{\beta}}\,. \tag{3.497}$$

Accepting that the covariant equations of motion for matter are hold: $\nabla_\alpha T^{\alpha}{}_{\beta} = 0$, the right side of the above equation is

$$\frac{1}{2}\sqrt{-g}T^{\rho\sigma}\frac{\partial g_{\rho\sigma}}{\partial x^{\beta}} = \partial_\alpha\left(\sqrt{-g}T^{\alpha}{}_{\beta}\right)\,. \tag{3.498}$$

Thus, equation (3.497) takes on the form of the law of conservation

$$\partial_\alpha\left[\sqrt{-g}\left(t^{\alpha}{}_{\beta} + T^{\alpha}{}_{\beta}\right)\right] = 0\,. \tag{3.499}$$

This is gravitational analogue of the Noether theorem that implies that the sum of two terms is equal to a superpotential $U^{\alpha\gamma}{}_{\beta} = U^{[\alpha\gamma]}{}_{\beta}$ such that

$$\sqrt{-g}\left(t^{\alpha}{}_{\beta} + T^{\alpha}{}_{\beta}\right) = \partial_\gamma U^{[\alpha\gamma]}{}_{\beta}\,. \tag{3.500}$$

Pseudotensors, defined by this equation, are not symmetric in the most general case. Moreover, they are not defined uniquely because the superpotential is defined only up to a class of equivalence being determined by equation $U^{[\alpha\gamma]}{}_\beta = \partial_\delta Z^{\alpha\gamma\delta}{}_\beta$, where the indexed object, $Z^{\alpha\gamma\delta}{}_\beta$, possesses the following (anti)symmetry properties: $Z^{\alpha\gamma\delta}{}_\beta = Z^{[\alpha\gamma]\delta}{}_\beta = Z^{\alpha[\gamma\delta]}{}_\beta$. It is impossible to chose on a physical ground the "best" superpotential. However, one may be guided by pure mathematical reasons of convenience in practical calculations. Perhaps, the most reasonable option was proposed in the textbook by [Landau and Lifshitz, 1975], which is discussed next.

Landau-Lifshitz pseudotensor, $t^{\alpha\beta}_{LL}$ is defined by equation [Landau and Lifshitz, 1975]

$$(-g)\left(t^{\alpha\beta}_{LL} + T^{\alpha\beta}\right) = \frac{1}{2\kappa}\partial^2_{\mu\nu}\left[(-g)\left(g^{\alpha\beta}g^{\mu\nu} - g^{\alpha\mu}g^{\beta\nu}\right)\right] , \tag{3.501}$$

where $\partial^2_{\mu\nu} = \partial^2/\partial x^\mu \partial x^\nu$ - a second partial derivative, and $\kappa = 8\pi G/c^4$. Landau-Lifshitz pseudotensor is not, strictly speaking, a canonical pseudotensor that is defined by equation (3.500) having a factor of $\sqrt{-g}$ in its left side. However, it has a nice property of symmetry, $t^{\alpha\beta}_{LL} = t^{\beta\alpha}_{LL}$, and satisfies the law of conservation

$$\partial_\alpha\left[(-g)\left(t^{\alpha\beta}_{LL} + T^{\alpha\beta}\right)\right] = 0 . \tag{3.502}$$

In addition, Landau-Lifshitz pseudotensor is made only of the metric tensor and its first derivatives.

Calculation of the pseudotensor is performed under assumption that the Einstein field equations (3.338) are valid. Replacing the metrical energy-momentum tensor of matter, $T^{\alpha\beta}$, with the Einstein tensor $G^{\alpha\beta}$ gives us

$$t^{\alpha\beta}_{LL} = \frac{1}{2\kappa}\partial^2_{\mu\nu}\left[(-g)\left(g^{\alpha\beta}g^{\mu\nu} - g^{\alpha\mu}g^{\beta\nu}\right)\right] - \frac{1}{\kappa}G^{\alpha\beta} , \tag{3.503}$$

Taking derivatives and reducing similar terms yield expression for the pseudotensor in terms of the Christoffel symbols

$$\begin{aligned} 2\kappa t^{\alpha\beta}_{LL} &= \left(2\Gamma^\sigma_{\mu\nu}\Gamma^\rho_{\sigma\rho} - \Gamma^\sigma_{\mu\rho}\Gamma^\rho_{\nu\sigma} - \Gamma^\sigma_{\mu\sigma}\Gamma^\rho_{\nu\rho}\right)\left(g^{\alpha\mu}g^{\beta\nu} - g^{\alpha\beta}g^{\mu\nu}\right) \\ &+ g^{\alpha\mu}g^{\nu\sigma}\left(\Gamma^\beta_{\mu\rho}\Gamma^\rho_{\nu\sigma} + \Gamma^\beta_{\nu\sigma}\Gamma^\rho_{\mu\rho} - \Gamma^\beta_{\sigma\rho}\Gamma^\rho_{\mu\nu} - \Gamma^\beta_{\mu\nu}\Gamma^\rho_{\sigma\rho}\right) \\ &+ g^{\beta\mu}g^{\nu\sigma}\left(\Gamma^\alpha_{\mu\rho}\Gamma^\rho_{\nu\sigma} + \Gamma^\alpha_{\nu\sigma}\Gamma^\rho_{\mu\rho} - \Gamma^\alpha_{\sigma\rho}\Gamma^\rho_{\mu\nu} - \Gamma^\alpha_{\mu\nu}\Gamma^\rho_{\sigma\rho}\right) \\ &+ g^{\mu\nu}g^{\sigma\rho}\left(\Gamma^\alpha_{\mu\sigma}\Gamma^\beta_{\nu\rho} - \Gamma^\alpha_{\mu\nu}\Gamma^\beta_{\sigma\rho}\right) . \end{aligned} \tag{3.504}$$

There is another equivalent expression for the pseudotensor, given in terms of the contravariant Gothic-style metric, $\mathfrak{g}^{\alpha\beta} \equiv \sqrt{-g}g^{\alpha\beta}$. If one imposes condition, $\mathfrak{g}^{\alpha\beta}\mathfrak{g}_{\beta\gamma} = \delta^\alpha_\gamma$, the covariant Gothic-style metric will be, $\mathfrak{g}_{\alpha\beta} \equiv g_{\alpha\beta}/\sqrt{-g}$. Landau-Lifshitz pseudotensor, expressed in terms of the Gothic-style metric reads

$$\begin{aligned} 2\kappa(-g)t^{\alpha\beta}_{LL} &= \partial_\mu\mathfrak{g}^{\alpha\beta}\partial_\nu\mathfrak{g}^{\mu\nu} - \partial_\mu\mathfrak{g}^{\alpha\mu}\partial_\nu\mathfrak{g}^{\beta\nu} + \frac{1}{2}\mathfrak{g}^{\alpha\beta}\mathfrak{g}_{\mu\nu}\partial_\rho\mathfrak{g}^{\mu\sigma}\partial_\sigma\mathfrak{g}^{\rho\nu} \\ &- \mathfrak{g}^{\alpha\mu}\mathfrak{g}_{\nu\sigma}\partial_r ho\mathfrak{g}^{\beta\sigma}\partial_\mu\mathfrak{g}^{\nu\rho} + \mathfrak{g}^{\beta\mu}\mathfrak{g}_{\nu\sigma}\partial_\rho\mathfrak{g}^{\alpha\sigma}\partial_\mu\mathfrak{g}^{\nu\rho} + \mathfrak{g}_{\mu\nu}\mathfrak{g}^{\sigma\rho}\partial_\sigma\mathfrak{g}^{\alpha\mu}\partial_\rho\mathfrak{g}^{\beta\nu} \\ &+ \frac{1}{8}\left(2\mathfrak{g}^{\alpha\mu}\mathfrak{g}^{\beta\nu} - \mathfrak{g}^{\alpha\beta}\mathfrak{g}^{\mu\nu}\right)\left(2\mathfrak{g}_{\sigma\rho}\mathfrak{g}_{\lambda\omega} - \mathfrak{g}_{\rho\lambda}\mathfrak{g}_{\sigma\omega}\right)\partial_\mu\mathfrak{g}^{\sigma\omega}\partial_\nu\mathfrak{g}^{\rho\lambda} . \end{aligned} \tag{3.505}$$

Each term in pseudotensor contains a first derivative from the metric tensor. Consequently, pseudotensor vanishes at any point of spacetime manifold when the coordinate system is chosen to make the first derivatives of the metric equal to zero. This property of pseudotensor is a natural consequence of the equivalence principle telling us that gravitational energy is not localizable. Therefore, treatment of conserved quantities in general relativity requires more refined approach [Babak and Grishchuk, 2000; Petrov and Katz, 2002]. In what follows, the pseudotensor will be used as a mathematical tool for solving Einstein's field equations by iterations.

3.10 Gravitational Waves

Einstein's theory of general relativity predicts existence of gravitational waves which is treated as a phenomenon resulting from the fluctuating curvature of spacetime propagating freely in space. In accordance with Einstein's equations the curvature is caused by the presence of material objects like planets, stars, galaxies or material fields - scalar, electromagnetic, etc. More massive and accelerating object produces stronger curvature. If massive objects are stars forming a binary system, they move around each other in spacetime, thus, producing temporal changes in the curvature that reflects the change in the location of the stars. Under certain circumstances, the disturbance in curvature tears away from the source of gravity and begins propagating in space independently. This disturbance is known as *gravitational wave* and the process of emission of gravitational waves is called *gravitational radiation*. General relativity predicts that gravitational waves travel with the fundamental speed c - the same as the speed of light in vacuum. Amplitude of gravitational wave is characterized by the dimensionless deviation of the metric tensor from the Minkowski metric, $h_{\alpha\beta} = g_{\alpha\beta} - \eta_{\alpha\beta}$. It diminishes, as the wave gets further away from its source, inversely proportional to the distance from the source.

As waves of gravitational radiation pass a distant observer, that observer will find the local curvature of spacetime distorted in a way that resembles lunisolar tides in the Newtonian gravity. It means that free test particles will feel the presence of the gravitational wave in the form of a tiny tidal force that will increase and decrease rhythmically the distances between particles as the wave passes. The magnitude of this effect is proportional to the amplitude of the gravitational wave passing through the particles. Binary neutron stars/black holes are considered to be strong sources of gravitational waves as they orbit each other at small separation and have enormous orbital acceleration. Nevertheless, since astronomical distances to such objects are large, the resulting amplitude of gravitational waves from these sources that observers are trying to measure, is very small $|h_{\alpha\beta}| \leq 10^{-21}$. Scientists are attempting to find gravitational waves with sensitive detectors of various types [Bertotti *et al.*, 1999; Cerdonio, 2002; LIGO, 2010; Verbiest *et al.*, 2009] The current best upper limit on a wave amplitude is 2.3×10^{-26} from the fifth science run of the LIGO detectors [Abbott *et al.*, 2010]. No signal detection from any of the targets was reported so far.

This section gives a mathematical introduction to the theory of gravitational radiation in the asymptotically-flat spacetime. If spacetime is not asymptotically-flat, like that in cosmology, more profound mathematical treatment is required [Bardeen, 1980; Grishchuk *et al.*, 1984; Mukhanov, 2005; Ramírez and Kopeikin, 2002].

3.10.1 The Post-Minkowskian approximations

Landau-Lifshitz pseudotensor is used to build a powerful approximation scheme of solving the Einstein equations in the most general cases when spacetime has no symmetries [Anderson and Decanio, 1975; Blanchet and Damour, 1986; Thorne, 1980]. Such approximations starts from some background geometry that has no dynamical degrees of freedom and is used as a reference for finding gravitational field. Asymptotically flat spacetime is the most simple case of the background geometry with the background value of the Riemann tensor $R^{\alpha}{}_{\beta\gamma\delta} = 0$. Therefore, there are coordinates, $x^{\alpha} = (x^0, x^i) = (ct, \boldsymbol{x})$, in which the background value of the metric tensor is given by the Minkowski metric, $\eta_{\alpha\beta} = \mathrm{diag}(-1, 1, 1, 1)$. Hence, the gravitational field potentials can be defined in the form of perturbation of the flat background

$$\mathfrak{g}^{\alpha\beta} = \eta^{\alpha\beta} + \gamma^{\alpha\beta} \,, \tag{3.506}$$

where $\mathfrak{g}^{\alpha\beta} = \sqrt{-g}g^{\alpha\beta}$, $\mathfrak{g}_{\alpha\beta} = g_{\alpha\beta}/\sqrt{-g}$, and it is the flat metric, $\eta_{\alpha\beta}$, which is used in the operation of lowering and rising indices of geometric objects in all subsequent calculations, for example $\gamma^{\alpha}{}_{\beta} = \gamma^{\alpha\mu}\eta_{\mu\beta}$, $\gamma_{\alpha\beta} = \eta_{\alpha\mu}\eta_{\beta\nu}\gamma^{\mu\nu}$, etc. Equation (3.506) is exact. Analogue of equation (3.506) with the two indices lowered, is obtained from the isomorphism identity, $\mathfrak{g}^{\alpha\nu}\mathfrak{g}_{\nu\beta} = \delta^{\alpha}_{\beta}$, and represents an infinite Taylor series

$$\mathfrak{g}_{\alpha\beta} = \eta_{\alpha\beta} - \gamma_{\alpha\beta} + \gamma^{\mu}{}_{(\alpha}\gamma_{\beta)\mu} + O\left(|\gamma_{\alpha\beta}|^3\right) \,. \tag{3.507}$$

Correspondence with the expansion of the metric tensor

$$g_{\alpha\beta} = \eta_{\alpha\beta} - h_{\alpha\beta} \,, \tag{3.508}$$

can be easily established after expansion of determinant of the metric tensor [82] in powers of $\gamma^{\alpha\beta}$

$$-g = 1 + \gamma + \frac{1}{2}\left(\gamma^2 - \gamma_{\alpha\beta}\gamma^{\alpha\beta}\right) + O\left(|\gamma_{\alpha\beta}|^3\right) \,. \tag{3.509}$$

where notation $\gamma \equiv \gamma^{\alpha}{}_{\alpha} = \eta_{\alpha\beta}\gamma^{\alpha\beta}$ has been used. Making use of this result along with equation (3.507) one obtains,

$$h_{\alpha\beta} = \gamma_{\alpha\beta} - \frac{1}{2}\eta_{\alpha\beta}\gamma^{\mu}{}_{\mu} \,. \tag{3.510}$$

The Einstein field equations are given by equation (3.501). After substitution of the metric tensor decomposition (3.506), they become

$$G^{\alpha\beta}_L = 2\kappa\Lambda^{\alpha\beta} \,, \tag{3.511}$$

where $\Lambda^{\alpha\beta}$ is the effective energy-momentum tensor density of matter and gravitational field, and

$$G^{\alpha\beta}_L = \Box\gamma^{\alpha\beta} + \eta^{\alpha\beta}\partial^2_{\mu\nu}\gamma^{\mu\nu} - \partial^{\alpha}\partial_{\nu}\gamma^{\nu\beta} - \partial^{\beta}\partial_{\nu}\gamma^{\nu\alpha} \,, \tag{3.512}$$

is a linear differential operator that includes the wave (D'Alambertian) operator,

$$\Box \equiv \eta^{\mu\nu}\partial_{\mu}\partial_{\nu} = -\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \triangle \,, \tag{3.513}$$

82) Determinant of $\mathfrak{g}^{\alpha\beta}$ is equal to the determinant of $g_{\alpha\beta}$, that is $\det[\mathfrak{g}^{\alpha\beta}] = \det[g_{\alpha\beta}] = g$.

of the Minkowski spacetime, where $\triangle$ is the Laplace operator (1.39).

The effective stress-energy tensor density is

$$\Lambda^{\alpha\beta} = \tau^{\alpha\beta} + t^{\alpha\beta} , \tag{3.514}$$

$$t^{\alpha\beta} = (-g)t^{\alpha\beta}_{LL} - \frac{1}{2\kappa}\partial^2_{\mu\nu}\left(\gamma^{\mu\nu}\gamma^{\alpha\beta} - \gamma^{\mu\alpha}\gamma^{\nu\beta}\right) , \tag{3.515}$$

where $\tau^{\alpha\beta} \equiv (-g)T^{\alpha\beta}$, $T^{\alpha\beta}$ is the metrical energy-momentum tensor of matter, and $t^{\alpha\beta}_{LL}$ is the Landau-Lifshitz pseudotensor of the gravitational field. The linear operator $G^{\alpha\beta}_L$ is conserved identically

$$\partial_\beta G^{\alpha\beta}_L \equiv 0 . \tag{3.516}$$

This identity brings about the law of conservation of the effective energy-momentum tensor, $\Lambda^{\alpha\beta}$, which is reduced to

$$\partial_\beta\left(\tau^{\alpha\beta} + t^{\alpha\beta}\right) = 0 . \tag{3.517}$$

This equation yields a local equation of motion of matter moving in the gravitational field.

Einstein's equations (3.511) can be significantly simplified if the deDonder (harmonic) gauge conditions

$$\partial_\beta\gamma^{\alpha\beta} = 0 , \tag{3.518}$$

are imposed on the metric tensor. Using these conditions one can reduce equation (3.511) to a wave equation in the Minkowski space,

$$\Box\,\gamma^{\alpha\beta} = 2\kappa\left(\tau^{\alpha\beta} + t^{\alpha\beta}\right) , \tag{3.519}$$

where

$$\begin{aligned} 2\kappa t^{\alpha\beta} = {} & \partial_\nu\gamma^{\mu\alpha}\partial_\mu\gamma^{\nu\beta} - \gamma^{\mu\nu}\partial^2_{\mu\nu}\gamma^{\alpha\beta} + \frac{1}{2}\mathfrak{g}^{\alpha\beta}\mathfrak{g}_{\mu\nu}\partial_\rho\gamma^{\mu\sigma}\partial_\sigma\gamma^{\rho\nu} \\ & - \mathfrak{g}^{\alpha\mu}\mathfrak{g}_{\nu\sigma}\partial_\rho\gamma^{\beta\sigma}\partial_\mu\gamma^{\nu\rho} + \mathfrak{g}^{\beta\mu}\mathfrak{g}_{\nu\sigma}\partial_\rho\gamma^{\alpha\sigma}\partial_\mu\gamma^{\nu\rho} + \mathfrak{g}_{\mu\nu}\mathfrak{g}^{\sigma\rho}\partial_\sigma\gamma^{\alpha\mu}\partial_\rho\gamma^{\beta\nu} \\ & + \frac{1}{8}\left(2\mathfrak{g}^{\alpha\mu}\mathfrak{g}^{\beta\nu} - \mathfrak{g}^{\alpha\beta}\mathfrak{g}^{\mu\nu}\right)\left(2\mathfrak{g}_{\sigma\rho}\mathfrak{g}_{\lambda\omega} - \mathfrak{g}_{\rho\lambda}\mathfrak{g}_{\sigma\omega}\right)\partial_\mu\gamma^{\sigma\omega}\partial_\nu\gamma^{\rho\lambda} . \end{aligned} \tag{3.520}$$

Equation of motion of matter (3.517) are valid in the harmonic gauge as a consequence of the gauge condition (3.518) and the reduced field equation (3.519).

Equation (3.519) admits a residual gauge freedom associated with a class of equivalence of diffeomorphisms

$$x'^\alpha = x^\alpha + \xi^\alpha , \tag{3.521}$$

which preserve the deDonder gauge condition (3.518) if four gauge functions ξ^α obeys a homogeneous Beltrami's equation

$$g^{\mu\nu}\partial^2_{\mu\nu}\xi^\alpha = 0 . \tag{3.522}$$

In a linear approximation with respect to the metric tensor perturbation, this exact equation yields a homogeneous wave equation

$$\Box \xi^\alpha = 0 \,. \tag{3.523}$$

Equation (3.522) admits a plenty of non-trivial solutions each of which generates a gauge transformation of gravitational potentials, $\gamma^{\alpha\beta}$, given by the exponential mapping

$$\begin{aligned} \gamma'^{\alpha\beta} &= \gamma^{\alpha\beta} + \pounds_{\boldsymbol{\xi}} \mathfrak{g}^{\alpha\beta} \\ &= \gamma^{\alpha\beta} + \partial^\alpha \xi^\beta + \partial^\beta \xi^\alpha - \eta^{\alpha\beta} \partial_\mu \xi^\mu \,, \end{aligned} \tag{3.524}$$

where $\pounds_{\boldsymbol{\xi}}$ denotes the Lie derivative along the vector field ξ^α.

Equation (3.519) is an inhomogeneous wave equation that can be formally solved by making use of the Green function for the D'Alambertian operator. There are two Green functions of the D'Alembert equation - advanced and retarded ones [Arfken and Weber, 2001]. We employ the retarded Green function as it complies with the causality principle. Retarded solution of equation (3.519) becomes

$$\gamma^{\alpha\beta}(t, \boldsymbol{x}) = -\frac{\kappa}{2\pi} \int \frac{\tau^{\alpha\beta}(s, \boldsymbol{x}') + t^{\alpha\beta}(s, \boldsymbol{x}')}{|\boldsymbol{x} - \boldsymbol{x}'|} \, d^3x' \,, \tag{3.525}$$

where

$$s = t - \frac{1}{c} |\boldsymbol{x} - \boldsymbol{x}'| \,, \tag{3.526}$$

is the retarded time owing to the finite speed of propagation of gravity from a point inside the source of gravity with spatial coordinate $\boldsymbol{x}'$ to the field point having coordinates $\boldsymbol{x}$ and the integration is performed over the entire space [83].

Generally speaking, equation (3.525) represents solution of the gravity field equations only in a formal sense since the yet unknown potentials, $\gamma^{\alpha\beta}$, enter not only left but right side of equation (3.525) via $t^{\alpha\beta}$. Nevertheless, one has reached a certain progress for the integral equation (3.525) can be resolved by iterations if gravitational field is weak, that is if $|\gamma^{\alpha\beta}| \ll 1$. We shall assume that this condition is satisfied everywhere. In case, when this assumption is not applied in some space domain, for example, inside neutron stars or black holes, equation (3.525) should be used only outside of the strong-field region. Its solution will be characterized by a number of free parameters (multipole moments) that can be fixed by matching the weak-field asymptotic expansion of the metric tensor with the interior solution of the strong gravity-field domain [Damour, 1987; D'Eath, 1975a,b; Thorne, 1980].

83) Although the fundamental speed c entering equation (3.526) is traditionally called "the speed of light", it has a physical meaning of the speed of gravity as it appears here through a solution of the Einstein gravity field equations. Various confusions arising due to misinterpretation of the meaning of the invariant speed c in different physical equations have been analyzed in [Ellis and Uzan, 2005; Kopeikin and Fomalont, 2006].

In a weak-field approximation one expands $\gamma^{\alpha\beta}$ in a power series with respect to $|\gamma^{\alpha\beta}|$ where the coupling constant κ can be formally considered as the parameter of the expansion,

$$\gamma^{\alpha\beta} = \kappa\gamma_1^{\alpha\beta} + \kappa^2\gamma_2^{\alpha\beta} + \ldots \, . \tag{3.527}$$

The process of finding functions $\gamma_n^{\alpha\beta}$ $(n = 1, 2, \ldots)$ is called the *post-Minkowskian approximations* [Blanchet and Damour, 1984a; Damour, 1987] as the iteration process of finding the metric starts from the Minkowski spacetime and functions $\gamma_n^{\alpha\beta}$ are invariant under the Lorentz transformations. Notice that slow motion of matter comprising the source of the gravitational field is not required to conduct the post-Minkowskian approximations. The iterative scheme of solving the Einstein equations that relies upon the assumptions that field is weak and motion of matter is slow, is called the *post-Newtonian approximations* [Damour, 1987]. They can be obtained by expanding the post-Minkowski gravitational potentials, $\gamma_n^{\alpha\beta}$, in additional series [Blanchet, 1999] with respect to the slow-motion parameter v/c, where v is a characteristic velocity of matter, and c is the speed of gravity as one solves the gravity field equations.

Post-Minkowski series (3.527) is substituted to the kernel of integral in equation (3.525). It leads to the post-Minkowski expansion of both the energy-momentum tensor of matter, which depends on the metric tensor, and the pseudotensor

$$\tau^{\alpha\beta} = T^{\alpha\beta} + \kappa\tau_1^{\alpha\beta} + \kappa^2\tau_2^{\alpha\beta} + \ldots \, , \tag{3.528}$$

$$t^{\alpha\beta} = \kappa t_1^{\alpha\beta} + \kappa^2 t_2^{\alpha\beta} + \ldots \, , \tag{3.529}$$

where $T^{\alpha\beta}$ denotes the energy-momentum tensor of matter in flat spacetime.

Substitution of the asymptotic expansion (3.527) to the integral equation (3.525) leads to a formal hierarchy of inhomogeneous wave equations for $\gamma_n^{\alpha\beta}$s of the type

$$\gamma_1^{\alpha\beta}(t, \boldsymbol{x}) = -\frac{1}{2\pi}\int \frac{T^{\alpha\beta}(s, \boldsymbol{x}')}{|\boldsymbol{x} - \boldsymbol{x}'|}\, d^3x' \, , \tag{3.530a}$$

$$\gamma_2^{\alpha\beta}(t, \boldsymbol{x}) = -\frac{1}{2\pi}\int \frac{\tau_1^{\alpha\beta}(s, \boldsymbol{x}') + t_1^{\alpha\beta}(s, \boldsymbol{x}')}{|\boldsymbol{x} - \boldsymbol{x}'|}\, d^3x' \, , \tag{3.530b}$$

$$\ldots\ldots\ldots\ldots \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \, , \tag{3.530c}$$

$$\gamma_n^{\alpha\beta}(t, \boldsymbol{x}) = -\frac{1}{2\pi}\int \frac{\tau_n^{\alpha\beta}(s, \boldsymbol{x}') + t_n^{\alpha\beta}(s, \boldsymbol{x}')}{|\boldsymbol{x} - \boldsymbol{x}'|}\, d^3x' \, . \tag{3.530d}$$

As soon as the energy-momentum tensor of matter $T^{\alpha\beta}$ is known, one can calculate the linear, $\gamma_1^{\alpha\beta}$, perturbation of the gravitational field from equation (3.530a). Substituting $\gamma_1^{\alpha\beta}$ to definitions of $\tau_1^{\alpha\beta}$ and $t_1^{\alpha\beta}$, and solving equation (3.530b) for $n = 2$ give us the quadratic gravitational perturbation, $\gamma_2^{\alpha\beta}$, and so on [Blanchet, 1998]. In principle, this procedure can be repeated iteratively to any order n of the post-Minkowskian approximation but technical difficulties in handling the calculations grow enormously as soon as the order of the iterations exceed $n = 3$ [Blanchet, 2002a; Blanchet and Damour, 1986; Thorne, 1980].

3.10.2 Multipolar expansion of a retarded potential

Let us consider a field $\psi = \psi(t, \boldsymbol{x})$ that is produced by some *extended* source $S(t, \boldsymbol{x})$ enclosed in a finite volume $\mathcal{V}$ of space with a characteristic size r_0. One assumes the field ψ obeys the D'Alambert equation

$$\Box\,\psi(t, \boldsymbol{x}) = -4\pi S(t, \boldsymbol{x})\,. \tag{3.531}$$

Let $\psi(t, \boldsymbol{x})$ be the retarded solution of this equation, that is

$$\psi(t, \boldsymbol{x}) = \int_{\mathcal{V}} \frac{S(s, \boldsymbol{y})}{|\boldsymbol{x} - \boldsymbol{y}|}\, d^3y\,, \tag{3.532}$$

where

$$s = t - \frac{1}{c}|\boldsymbol{x} - \boldsymbol{y}|\,, \tag{3.533}$$

is the retarded time from point $\boldsymbol{y}$ to the field point $\boldsymbol{x}$.

One is interested in the detailed structure of the solution ψ outside of the source S at distances $|\boldsymbol{x}| > r_0$. Campbell *et al.* [1977] and Blanchet and Damour [1986, 1989] had proved a powerful theorem stating that outside of the source the field can be expanded in the following (exact) multipolar series

$$\psi(t, \boldsymbol{x}) = \sum_{l=0}^{\infty} \frac{(-1)^l}{l!}\partial_L \left[\frac{M_L(u)}{r}\right]\,, \tag{3.534}$$

where $r = |\boldsymbol{x}|$ is distance from the coordinate origin to the field point, $u = t - r/c$ is the retarded time counted from the coordinate origin, multi-index $L = i_1 \ldots i_l$, the l-th order partial derivative $\partial_L \equiv \partial_{i_1} \ldots \partial_{i_l}$, and the functions $M_L(u)$ are the *symmetric trace-free* (STF) multipole moments of the source

$$M_L(u) = \int_{\mathcal{V}} d^3y\, y^{<L>} \hat{S}_l(u, \boldsymbol{y})\,, \tag{3.535}$$

with the kernel

$$\hat{S}_l(u, \boldsymbol{y}) = \frac{(2l+1)!!}{2^{l+1} l!} \int_{-1}^{+1} dz \left(1 - z^2\right)^l S\left(u + \frac{|\boldsymbol{y}|}{c} z, \boldsymbol{y}\right)\,. \tag{3.536}$$

STF combination of coordinates, $y^{<L>} = y^{<i_1 \ldots i_l>}$, in equation (3.535) is defined as

$$y^{<i_1 \ldots i_l>} = \sum_{k=0}^{[l/2]} a_k^l \delta^{(i_1 i_2} \ldots \delta^{i_{2k-1} i_{2k}} y^{i_{2k+1} \ldots i_l)} |\boldsymbol{y}|^{2k}\,, \tag{3.537}$$

where the round brackets around a group of indices denote full symmetrization with respect to permutation of the indices, $[l/2]$ denotes the integer part of $l/2$, repeated indices denotes Einstein's summation, and numerical coefficient

$$a_k^l = (-1)^k \frac{l!}{(l-2k)!(2k)!!} \frac{(2l-2k-1)!!}{(2l-1)!!}\,. \tag{3.538}$$

For practical purposes it is often useful to consider the slow-motion (post-Newtonian) expansion of function $\hat{S}_l$ in equation (3.536). This expansion assumes that any temporal change inside the source has a characteristic time much larger than the time interval taken by the field to cross the source. It means that one can expand function S entering the integrand of equation (3.536) in a Taylor series with respect to the argument z. Performing this expansion and integrating with respect to z, one obtains [Blanchet and Damour, 1989; Campbell *et al.*, 1977]

$$\hat{S}_l(u,\boldsymbol{y}) = \sum_{k=0}^{\infty} \frac{(2l+1)!!}{2^k k!(2l+2k+1)!!} \frac{|\boldsymbol{y}|^{2k}}{c^{2k}} \frac{\partial^{2k}}{\partial u^{2k}} S(u,\boldsymbol{y}) \,, \tag{3.539}$$

so that the multipole expansion (3.535) takes the following form of an asymptotic series

$$\begin{aligned} M_L(u) \quad = \quad \int_{\mathcal{V}} d^3y y^{<L>} \Bigg[S(u,\boldsymbol{y}) + \frac{1}{2(2l+3)} \frac{|\boldsymbol{y}|^2}{c^2} \frac{\partial^2 S(u,\boldsymbol{y})}{\partial u^2} \\ + \frac{1}{8(2l+3)(2l+5)} \frac{|\boldsymbol{y}|^4}{c^4} \frac{\partial^4 S(u,\boldsymbol{y})}{\partial u^4} + \dots \Bigg] \,. \end{aligned} \tag{3.540}$$

The series contains only even powers of distance, $|\boldsymbol{y}|$, and time derivatives of function $S(u,\boldsymbol{y})$.

3.10.3
Multipolar expansion of gravitational field

Multipolar expansion of the retarded potential (3.532) can be used in order to derive the multipolar expansion of gravitational field perturbations $\gamma^{\alpha\beta}$ in terms of symmetric and trace-free (STF) tensors which represent a set of irreducible harmonics of the group of rotation, $SO(3)$, in three-dimensional space [Blanchet and Damour, 1986; Coope *et al.*, 1965; Thorne, 1980]. General solution is complicated because gravitational field is non-linear, and one has to solve the (infinite) hierarchy of the post-Minkowski integrals (3.530) by successive iterations. There is a certain progress in understanding of how to solve this problem, and the achievements have been summarized in review papers by Blanchet [2002a]; Blanchet *et al.* [2001a]. For the purpose of this book, it is sufficient to describe only a linearized solution for functions $\gamma_1^{\alpha\beta}$ that has been found by Blanchet and Damour [1986]; Damour and Iyer [1991a]; Thorne [1980].

Straightforward application of formula (3.534) to equation (3.530a) yields

$$\kappa\gamma_1^{\alpha\beta}(t,\boldsymbol{x}) = -\frac{4}{c^4}\sum_{l=0}^{\infty}\frac{(-1)^l}{l!}\partial_L\left[\frac{M_L^{\alpha\beta}(u)}{r}\right] \,, \tag{3.541}$$

where

$$M_L^{\alpha\beta}(u) \quad = \quad \int_{\mathcal{V}} d^3y y^{<L>} \hat{S}_l^{\alpha\beta}(u,\boldsymbol{y}) \,, \tag{3.542}$$

$$\hat{S}_l^{\alpha\beta}(u,\boldsymbol{y}) \quad = \quad \frac{(2l+1)!!}{2^{l+1}l!}\int_{-1}^{+1} dz \left(1-z^2\right)^l T^{\alpha\beta}\left(u+\frac{|\boldsymbol{y}|}{c}z,\boldsymbol{y}\right) \,. \tag{3.543}$$

Functions $M_L^{\alpha\beta}$ are integrals from the conserved energy-momentum tensor of matter, $T^{\alpha\beta}$, taken over the finite volume $\mathcal{V}$ of the astronomical system, but they are not the multipole moments of gravitational field as yet. The reason is that gravity is a tensor field of rank two, while $M_L^{\alpha\beta}$ have been found with a retarded potential adapted for expansion of a scalar field. Hence, $M_L^{\alpha\beta}$ mixes up tensor, vector and scalar harmonics that should be decomposed in a sum of algebraically-independent pieces of three-dimensional STF tensors. Time-time component of perturbation, γ_1^{00}, has no spatial indices and behaves like a scalar under spatial rotations. Hence, it does not require any further STF decomposition. Spacetime components, γ_1^{0i}, and space-space components γ_1^{ij} behave under spatial rotations as a spatial vector and tensor of the second rank respectively, and have to be decomposed in STF harmonics with the help of the irreducible Cartesian-tensor technique introduced by Coope [1970]; Coope and Snider [1970]; Coope *et al.* [1965]. This procedure involves rather tedious algebra and cumbersome intermediate calculations which can be found in [Damour and Iyer, 1991a]. Simplifications can be achieved if one remembers that gravitational potentials $\gamma_1^{\alpha\beta}$ are subject to the harmonic gauge condition (3.518).

Implementing this condition and taking into account that some residual gauge freedom remains, one arrives to the following multipolar expansion of gravitational field of an isolated astronomical system emitting gravitational waves [Blanchet and Damour, 1986; Damour and Iyer, 1991a; Thorne, 1980]

$$\kappa\gamma_1^{\alpha\beta} = \gamma_{\rm can}^{\alpha\beta} + \partial^\alpha\xi^\beta + \partial^\beta\xi^\alpha - \eta^{\alpha\beta}\partial_\mu\xi^\mu\ , \tag{3.544}$$

where $\gamma_{1\rm can}^{\alpha\beta}$ is, the so-called, *canonical* perturbation, and ξ^α are the gauge functions. The canonical perturbation is

$$\gamma_{\rm can}^{00} = -\frac{4G\mathcal{M}}{c^2 r} - \frac{4G}{c^2}\frac{\mathcal{I}_i N^i}{r^2} - \frac{4G}{c^2}\sum_{l=2}^{\infty}\frac{(-1)^l}{l!}\partial_L\left[\frac{\mathcal{I}_L(u)}{r}\right]\ , \tag{3.545a}$$

$$\gamma_{\rm can}^{0i} = -\frac{2G}{c^3}\frac{\varepsilon_{ipq}\mathcal{S}_p N_q}{r^2} + \frac{4G}{c^3}\sum_{l=2}^{\infty}\frac{(-1)^l l}{(l+1)!}\varepsilon_{ipq}\partial_{pL-1}\left[\frac{\mathcal{S}_{qL-1}(u)}{r}\right] \tag{3.545b}$$
$$-\frac{4G}{c^2}\frac{\mathcal{P}_i}{r} + \frac{4G}{c^3}\sum_{l=2}^{\infty}\frac{(-1)^l}{l!}\partial_{L-1}\left[\frac{\dot{\mathcal{I}}_{iL-1}(u)}{r}\right]\ ,$$

$$\gamma_{\rm can}^{ij} = -\frac{4G}{c^4}\sum_{l=2}^{\infty}\frac{(-1)^l}{l!}\partial_{L-2}\left[\frac{\ddot{\mathcal{I}}_{ijL-2}(u)}{r}\right] \tag{3.545c}$$
$$-\frac{8G}{c^4}\sum_{l=2}^{\infty}\frac{(-1)^l l}{(l+1)!}\partial_{pL-2}\left[\frac{\varepsilon_{pq(i}\dot{\mathcal{S}}_{j)qL-2}(u)}{r}\right]\ .$$

where overdot denotes a time derivative, $N_i = N^i = x^i/r$ is a unit vector directed from the origin of the coordinate system to the field point, $u = t - r/c$ is the retarded time due to finite value c of the speed of propagation of gravity, and

$$\mathcal{M} = \int_{\mathcal{V}} T^{00}(t, \boldsymbol{y}) d^3 y \,, \tag{3.546}$$

$$\mathcal{I}_i = \int_{\mathcal{V}} T^{00}(t, \boldsymbol{y}) y^i d^3 y \,, \tag{3.547}$$

$$\mathcal{P}_i = \int_{\mathcal{V}} T^{0i}(t, \boldsymbol{y}) d^3 y \,, \tag{3.548}$$

$$\mathcal{S}_i = \int_{\mathcal{V}} \varepsilon_{ijk} y^j T^{0k}(t, \boldsymbol{y}) d^3 y \,, \tag{3.549}$$

are correspondingly the total mass, the dipole moment, the linear momentum and the angular momentum (spin) of the system while $\mathcal{I}_L$ and $\mathcal{S}_L$ (≤ 2) are two independent sets of mass-type and spin-type multipole moments of the gravitational field.

Mathematical technique of the retarded Green function applied for finding the multipolar expansion of gravitational field (3.545) is valid for arbitrary location of the center of mass of the system with respect to the origin of the coordinate chart used for calculations. Position of the center of mass is determined by the dipole moment, $\mathcal{I}_i$, which is zero if the center of mass is placed to the coordinate origin. In the linearized approximation the dipole moment, $\mathcal{I}_i$, of the system can be either a constant or a linear function of time because solution (3.545) is Lorentz-invariant [Fock, 1964; Pirani, 1965]. The law of conservation of the energy-momentum tensor establishes a relationship between the time derivative of the dipole moment and the linear momentum of the system,

$$\mathcal{P}_i = \dot{\mathcal{I}}_i \,. \tag{3.550}$$

In linearized approximation the linear momentum is conserved

$$\dot{\mathcal{P}}_i = \ddot{\mathcal{I}}_i = 0 \,, \tag{3.551}$$

as a consequence of the law of conservation of the energy-momentum tensor, $\partial_\mu T^{\mu\nu} = 0$. Total mass $\mathcal{M}$ and spin $\mathcal{S}_i$ of the system are conserved and remain constant in the linearized approximation,

$$\dot{\mathcal{M}} = 0 \,, \qquad \dot{\mathcal{S}}_i = 0 \,, \tag{3.552}$$

again as a consequence of the law of conservation of the energy-momentum tensor. All other multipoles can be arbitrary functions of time depending on a temporal behavior of matter comprising the system. Gravitational waves emitted by the system, reduce its total energy and spin, and change the linear momentum. Hence, in higher-order approximations the total mass, spin and the linear moment of the system become functions of time with the time derivatives being not equal to zero [Thorne, 1980].

The canonical form of the metric tensor (3.545) depends on the STF multipole moments $\mathcal{I}_L(u)$ and $\mathcal{S}_L(u)$ taken at the retarded instant of time

$$u = t - \frac{r}{c} \,, \tag{3.553}$$

that is caused by the finite speed c of propagation of gravity from the astronomical system to the field point. In the most general case of arbitrary internal motion of matter of the isolated astronomical system, the multipole moments with multipolarity, $l \geq 2$, are given by [Blanchet, 1998; Damour and Iyer, 1991a]

$$\begin{aligned}\mathcal{I}_L(u) &= \int_{\mathcal{V}} d^3y y^{<L>} \hat{\mathcal{I}}_l(u, \boldsymbol{y}) - \frac{1}{c^2}\frac{d}{du}\int_{\mathcal{V}} d^3y y^{<pL>} \hat{\mathcal{I}}^p_{l+1}(u, \boldsymbol{y}) \\ &+ \frac{1}{c^4}\frac{d^2}{du^2}\int_{\mathcal{V}} d^3y y^{<pqL>} \hat{\mathcal{I}}^{pq}_{l+2}(u, \boldsymbol{y}) ,\end{aligned} \tag{3.554}$$

$$\begin{aligned}\mathcal{S}_L(u) &= \int_{\mathcal{V}} d^3y \varepsilon^{pq<i_l} y^{L-1>p} \hat{\mathcal{S}}^q_l(u, \boldsymbol{y}) \\ &- \frac{1}{c^2}\frac{d}{du}\int_{\mathcal{V}} d^3y \varepsilon^{pq<i_l} y^{L-1>pk} \hat{\mathcal{S}}^{kq}_{l+1}(u, \boldsymbol{y}) ,\end{aligned} \tag{3.555}$$

where the kernels of the integrands are

$$\hat{\mathcal{I}}_l(u, \boldsymbol{y}) = \int_{-1}^{+1} dz \delta_l(z) \sigma\left(u + \frac{|\boldsymbol{y}|}{c}z, \boldsymbol{y}\right) , \tag{3.556a}$$

$$\hat{\mathcal{S}}^i_l(u, \boldsymbol{y}) = \int_{-1}^{+1} dz \delta_l(z) \sigma^i\left(u + \frac{|\boldsymbol{y}|}{c}z, \boldsymbol{y}\right) , \tag{3.556b}$$

$$\hat{\mathcal{I}}^i_{l+1}(u, \boldsymbol{y}) = \frac{4(2l+1)}{(l+1)(2l+3)}\int_{-1}^{+1} dz \delta_{l+1}(z) \sigma^i\left(u + \frac{|\boldsymbol{y}|}{c}z, \boldsymbol{y}\right) , \tag{3.556c}$$

$$\hat{\mathcal{S}}^{ij}_{l+1}(u, \boldsymbol{y}) = \frac{2l+1}{(l+2)(2l+3)}\int_{-1}^{+1} dz \delta_{l+1}(z) \sigma^{ij}\left(u + \frac{|\boldsymbol{y}|}{c}z, \boldsymbol{y}\right) , \tag{3.556d}$$

$$\hat{\mathcal{I}}^{ij}_{l+2}(u, \boldsymbol{y}) = \frac{2(2l+1)}{(l+1)(l+2)(2l+5)}\int_{-1}^{+1} dz \delta_{l+2}(z) \sigma^{ij}\left(u + \frac{|\boldsymbol{y}|}{c}z, \boldsymbol{y}\right) , \tag{3.556e}$$

the kernel densities

$$\sigma \equiv T^{00} + T^{kk} , \qquad \sigma^i \equiv T^{0i} , \qquad \sigma^{ij} \equiv T^{ij} , \tag{3.557}$$

and the kernel's function

$$\delta_l(z) \equiv \frac{(2l+1)!!}{2(2l)!!}\left(1 - z^2\right)^l . \tag{3.558}$$

The moments $\mathcal{I}_L(u)$ and $\mathcal{S}_L(u)$ given by equations (3.554) and (3.555) are *exact* in any post-Newtonian order and can be applied even for ultra-relativistic motion of matter inside the source of gravity. The moments are defined as integrals from densities (3.557) having compact support. Extension of this formalism to higher post-Minkowskian approximations is fulfilled by doing subsequent iterations which is formally reduced to the replacement of the compact support densities, $\sigma^{\alpha\beta}$, with the effective energy-momentum tensor, $\tau^{\alpha\beta}$, which depends on the pseudotensor of gravitational field. The pseudotensor is made of the gravitational potentials, $\gamma^{\alpha\beta}$, and is distributed all over the space. It is not surprising that its integration becomes a challenging mathematical problem. Correspondingly, in the higher-order post-Newtonian approximations the multipole moments are expressed by means of more complicated functionals [Blanchet, 2002a].

If the temporal changes of the interior of the astronomical system are slow, the multipole moments of gravitational field can be presented as post-Newtonian asymptotic series with respect to the small parameter v/c, where v is a characteristic velocity of matter's motion corresponding to a characteristic time T of the internal changes in matter's density distribution. Introducing a characteristic wavelength, $\lambda\!\!\!^{-}$, of gravitational waves emitted by the system, the slow-motion parameter becomes $v/c \sim r_0/\lambda\!\!\!^{-} \ll 1$, where r_0 is the characteristic size of the system. Smallness of this parameter indicates that the geometric size of the system is much smaller than the wavelength of gravitational waves emitted by the system.

In the first post-Newtonian approximation the multipole moments of the gravitational field read [Damour and Iyer, 1991a]

$$
\begin{aligned}
\mathcal{I}_L^{1\mathrm{PN}} &= \int_V d^3y \left[y^{<L>}\sigma + \frac{|\boldsymbol{y}|^2 y^{<L>}\partial_t^2\sigma}{2c^2(2l+3)} - \frac{4(2l+1)y^{<pL>}\partial_t\sigma_i}{c^2(l+1)(2l+3)} \right] + O\left(\frac{v^4}{c^4}\right) \quad (3.559)\\
\mathcal{S}_L^{1\mathrm{PN}} &= \int_V d^3y\,\varepsilon^{pq<i_l}y^{L-1p>}\sigma_q + O\left(\frac{v^2}{c^2}\right) .
\end{aligned}
$$

3.10.4
Gravitational field in transverse-traceless gauge

Harmonic coordinates x^α that have been used in the post-Minkowski iteration scheme, are defined as solutions of the homogeneous wave equation $\square x^\alpha = 0$ up to the gauge functions ξ^α. The canonical part of the gravitational perturbation, $\gamma^{\alpha\beta}_{\mathrm{can}}$, is defined by the condition that all gauge functions $\xi^\alpha = 0$. The canonical metric tensor (3.545) depends only on two sets of the multipole moments, $\mathcal{I}_L(u)$ and $\mathcal{S}_L(u)$, which corresponds to the existence of only two degrees of freedom (polarizations) of freely-propagating gravitational field in general relativity. A generic expression for the harmonic metric tensor is obtained by making use of gauge transformation (3.544) with the gauge functions, ξ^α, which satisfy to a homogeneous wave equation (3.523). The most general solution of this equation contains four sets of STF multipoles [Blanchet and Damour, 1986; Thorne, 1980]

$$
\xi^0 = \sum_{l=0}^{\infty} \partial_L \left[\frac{\mathcal{W}_L(u)}{r} \right] , \qquad (3.560)
$$

$$
\xi^i = \sum_{l=0}^{\infty} \partial_{iL} \left[\frac{\mathcal{X}_L(u)}{r} \right] + \sum_{l=1}^{\infty} \left\{ \partial_{L-1} \left[\frac{\mathcal{Y}_{iL-1}(u)}{r} \right] + \varepsilon_{ipq}\partial_{pL-1}\left[\frac{\mathcal{Z}_{qL-1}(u)}{r} \right] \right\} \qquad (3.561)
$$

where $\mathcal{W}_L$, $\mathcal{X}_L$, $\mathcal{Y}_{iL-1}$, and $\mathcal{Z}_{iL-1}$ are Cartesian STF tensors depending on the retarded time $u = t - r/c$. Their specific form is a matter of computational convenience for derivation and interpretation of observable effects produced by the gravitational field since, whatever choice of the gauge functions ξ^α one makes, it does not influence measurable astronomical quantities like phase or frequency of electromagnetic wave (see section 7).

One particular choice of the gauge functions is especially important. It allows us to eliminate from the metric tensor perturbation all non-propagating degrees of freedom by making use of the so-called *transverse-traceless* (TT) gauge. It is achieved by picking up the following gauge functions

$$\xi^0 = \frac{G}{c^2}\sum_{l=1}^{\infty}\frac{(-1)^l}{l!}\partial_L\left[\frac{\mathcal{I}_L^{(-1)}(u)}{r}\right], \tag{3.562}$$

$$\xi^i = \frac{G}{c^3}\sum_{l=1}^{\infty}\frac{(-1)^l}{l!}\partial_{iL}\left[\frac{\mathcal{I}_L^{(-2)}(u)}{r}\right] - \frac{4G}{c^3}\sum_{l=1}^{\infty}\frac{(-1)^l}{l!}\partial_{L-1}\left[\frac{\mathcal{I}_{iL-1}(u)}{r}\right] - \frac{4G}{c^3}\sum_{l=2}^{\infty}\frac{(-1)^l l}{(l+1)!}\varepsilon_{ipq}\partial_{pL-1}\left[\frac{\mathcal{S}_{qL-1}^{(-1)}(u)}{r}\right], \tag{3.563}$$

where one has used a shorthand notation for time integrals from STF multipoles [Blanchet and Damour, 1988]

$$\mathcal{I}_L^{(-1)}(u) \equiv \int_{-\infty}^{u}\mathcal{I}_L(\tau)d\tau\,,\quad \mathcal{S}_L^{(-1)}(u) \equiv \int_{-\infty}^{u}\mathcal{S}_L(\tau)d\tau\,,\quad \mathcal{I}_L^{(-2)}(u) \equiv \int_{-\infty}^{u}\mathcal{I}_L^{(-1)}(\tau)d\tau\,. \tag{3.564}$$

The gauge functions after being substituted to equation (3.544), transform the canonical metric tensor perturbation to a remarkably simple form [Kopeikin *et al.*, 2006]

$$\gamma_{TT}^{00} = \frac{4G\mathcal{M}}{c^2 r}, \tag{3.565a}$$

$$\gamma_{TT}^{0i} = \frac{2G}{c^3}\frac{\varepsilon_{ipq}\mathcal{S}_p N_q}{r^2}, \tag{3.565b}$$

$$\gamma_{TT}^{ij} = P^{ijpq}\gamma_{\rm can}^{pq}, \tag{3.565c}$$

where $N_q = x^q/r$ is a unit vector directed from the origin of the coordinates to the field point, the differential operator of TT-projection, P_{ijkl}, is given by [Misner *et al.*, 1973]

$$P_{ijpq} = P_{ip}P_{jq} - \frac{1}{2}P_{ij}P_{pq}, \tag{3.566}$$

with

$$P_{ij} = \delta_{ij} - \Delta^{-1}\partial_i\partial_j, \tag{3.567}$$

and Δ and Δ^{-1} denote the Euclidean Laplacian and the inverse Laplacian operators respectively. The projection operator has the following properties

$$\delta^{ij}P_{ij} = P_{jj} = 2\,,\qquad P_{ij}P_{jk} = P_{ik},\qquad \delta^{ij}P_{ijpq} = \delta^{pq}P_{ijpq} = 0\,. \tag{3.568}$$

It makes evident that the spatial components of the TT metric tensor perturbation are indeed traceless, $\delta_{ij}\gamma^{ij}_{TT} = 0$. For this reason, the metric tensor components h^{ij}_{TT} coincide with the Gothic metric components: $h^{ij}_{TT} = \gamma^{ij}_{TT}$.

When comparing the canonical metric tensor with that given by equations (3.565) it is instructive to use the following exact relationships being valid for $r \neq 0$,

$$
{}^{(-2)}\ddot{\mathcal{I}}_L(u) = \mathcal{I}_L(u)\,, \tag{3.569}
$$

$$
\Delta\left[\frac{\mathcal{I}_L(u)}{r}\right] = \frac{\ddot{\mathcal{I}}_L(u)}{r}\,. \tag{3.570}
$$

This is a consequence of the fact that function ${}^{(-2)}\ddot{\mathcal{I}}_L(u)$ is solution of the homogeneous D'Alambert equation, that is, $\Box\left[{}^{(-2)}\mathcal{I}_L(u)/r\right] = 0$ for $r \neq 0$. Previous equations immediately yield

$$
\frac{\mathcal{I}_L(u)}{r} = \Delta^{-1}\left[\frac{\ddot{\mathcal{I}}_L(u)}{r}\right], \tag{3.571}
$$

$$
\frac{{}^{(-2)}\mathcal{I}_L(u)}{r} = \Delta^{-1}\left[\frac{\mathcal{I}_L(u)}{r}\right]. \tag{3.572}
$$

The TT metric tensor perturbation (3.565) is similar to the Coulomb gauge in electrodynamics [Jackson, 1998].

Standard textbooks, like [Landau and Lifshitz, 1975; Misner *et al.*, 1973], introduce TT gauge only for a plane gravitational wave. In this case the time-time and time-space components of the TT metric are nil: $\gamma^{00}_{TT} = \gamma^{0i}_{TT} = 0$. Flanagan and Hughes [2005] remarked that this condition along with the tracelessness of the spatial components of the metric comprise five constraints on the metric, while the residual gauge freedom in the harmonic gauge is parameterized by only four functions that satisfy the wave equation. This may look contradictive and one has avoided this contradiction in the description of the TT metric tensor (3.565) that properly introduces the non-zero values of the time-time and time-space components of the metric tensor along with all other multipolar harmonics of the gravitational field generated by the isolated astronomical system. The TT metric (3.565) is valid everywhere in space outside of the system and is not limited by the plane-wave or radiative approximation that is derived in next section.

3.10.5
Gravitational radiation and detection of gravitational waves

Direct experimental detection of gravitational waves is a fascinating, yet unsolved problem of modern fundamental physics [Fairhurst *et al.*, 2010]. Enormous efforts are applied to make a progress in its solution both by theorists and experimentalists. Main theoretical efforts are presently focused on calculation of templates of the gravitational waves emitted by coalescing binary systems comprised of neutron stars and/or black holes [Blanchet, 2002a; Sathyaprakash and Schutz, 2009] as well as creation of improved filtering technique for gravitational wave detectors [Babak, 2008; Owen and Sathyaprakash, 1999] which will enable one to extract the gravitational wave signal from all kind of interferences present in the data collected by the gravitational wave observatories [Babak *et al.*, 2009; Porter, 2009]. Direct experimental efforts have led to the construction of several ground-based optical interferometers some being with arms reaching few miles [LIGO, 2010]. Certain work is under way to build super-sensitive cryogenic-bar gravitational-wave detectors of Weber's type [Aguiar *et al.*, 2008; Astone *et al.*, 2007; Barriga *et al.*, 2010; Weber, 1961]. Laser interferometric space projects such as LISA [LISA, 2010] and ASTROD [Ni, 2009] aimed to significantly increase the sensitivity of the gravitational-wave detectors are currently under intensive discussion at NASA, ESA and China.

Direct detection of gravitational waves is supposed to be done with the gravitational wave detectors which primary element is an optical interferometer. The interferometer consists of many electronic feed-back components, a laser with highly-stabilized optical frequency and mirrors suspended at the ends of the interferometric arms on thin silica strings. The laser beam is split and bounces many times in the arms between the mirrors, then recombines in the optical instrument in such a way that interference occurs between them. The mirror's suspension system isolates them from noise of terrestrial origin (industrial, geophysical, thermal, laboratory-induced, etc.) in a certain frequency band $\sim 10 \div 1000$ Hz for ground-based gravitational-wave detectors. The high-quality noise isolation allows for free movement of the mirrors in this frequency band. The experimental problem of detection of gravitational waves is reduced to the observation of motion of the mirrors which behave like test particles in the field of the incident gravitational wave. Photons of a laser beam are also freely-moving test particles. Therefore, gravitational wave interacts with both photons and mirrors and perturbs their motion. These perturbations in the motion of mirrors and photons must be explicitly calculated and clearly separated to avoid possible misinterpretations of observable effects of the gravitational wave.

It turns out that the canonical form of the metric tensor (3.545) in harmonic coordinates is extremely well-adapted for performing analytic integration of equations of motion of photons. Freely-moving mirrors also experience influence of gravitational waves emitted by the isolated astronomical system and move with respect to the harmonic coordinates in a complicated way. For this reason, the effects imposed by the gravitational waves on the light propagation get mixed up with the motion of mirrors in these coordinates. Arnowitt *et al.* [1962] (ADM) showed that there exist coordinates which have a special property such that freely-falling proof masses that are originally at rest, will not move with respect to these coordinates despite the perturbations imposed on them by the gravitational waves [84]. This means that the ADM coordinates themselves are not inertial and, although have some nice mathematical properties, should be used with care in the interpretation of gravitational-wave experiments. Making use of the ADM coordinates simplifies mathematical analysis of

84) The initial condition of non-moving test particles is crucial. Nonvanishing initial velocity leads to the respond of the test particles on the incoming gravitational wave signal in TT gauge [Bolotovskiĭ and Serov, 1994; Braginskii and Grishchuk, 1985]. Consideration was also given to the conditions under which the TT gauge can be applied and to possible existence of gravitational-wave pulses for which the relative velocity of two test particles being initially at rest, is nonzero after passage of the pulse Grishchuk and Polnarev [1989].

the gravitational-wave signal incoming to the ground-based detectors or observed by other astronomical technique like very long baseline interferometry [Kopeikin *et al.*, 2006; Pyne *et al.*, 1996], pulsar timing array [Foster and Backer, 1990; Manchester, 2008], etc. This is because proof masses (mirror) do not move with respect to the ADM coordinates.

The ADM formalism is based on the slicing of spacetime by spatial hypersurfaces with arbitrary choice of time coordinate defined by the lapse, $N \equiv 1/\sqrt{-g^{00}}$, and shift, $N_i \equiv g_{0i}$, functions [Arnowitt *et al.*, 1962]. This 3+1 spacetime split is invariant with respect to a group of kinemetric transformations thoroughly discussed by Zel′manov [1944]; Zel′manov and Agakov [1989]. One can impose a certain ADM gauge condition to restrict the freedom of Zel'manov's kinemetric transformations in order to simplify the Einstein field equations in the ADM formalism. The gauge conditions must be compatible with the general ADM decomposition of space-space components g_{ij} of the metric tensor and its corresponding conjugated momentum, π^{ij}, into their longitudinal, transverse, and transverse-traceless parts [85] [Schäfer, 1982, 1985]. In the linear approximation the ADM gauge conditions read [Kopeikin *et al.*, 1999; Schäfer, 1985]

$$2\partial_i h_{0i} - \partial_0 h_{kk} = 0\,, \qquad 3\partial_j h_{ij} - \partial_i h_{kk} = 0\,, \tag{3.573}$$

where $h_{\alpha\beta}$ is the perturbation of the Minkowski metric introduced in equation (3.508). The first ADM gauge condition in equation (3.573) is simply, $\pi_{ii} = 0$ which is reduced to a limitation on components of the Christoffel symbols, $\Gamma^0{}_{ii} = 0$. The second ADM gauge condition in equation (3.573) is a consequence of the ADM decomposition of the metric tensor perturbation

$$h^{TT}_{ij} = h_{ij} - \frac{1}{3}\delta_{ij}h_{kk}\,, \tag{3.574}$$

where h^{TT}_{ij} denotes the transverse-traceless part of h_{ij} obeying to $\partial_j h^{TT}_{ij} = 0$ and $h^{TT}_{ii} = 0$, and the trace $h_{kk} = 3h_{00}$. For comparison, the harmonic gauge conditions (3.518) in the linear approximation read,

$$2\partial_i h_{0i} - \partial_0 h_{kk} = \partial_0 h_{00}\,, \qquad 2\partial_j h_{ij} - \partial_i h_{kk} = 2\partial_0 h_{0i} - \partial_i h_{00}\,. \tag{3.575}$$

The ADM and harmonic gauge conditions (3.573) can not be satisfied simultaneously inside the regions occupied by matter, for example, in the interior of a star. However, outside of matter (in vacuum) they can comply with each other as well as with the Chandrasekhar-Nutku gauge conditions [Chandrasekhar and Nutku, 1969], frequently used in relativistic celestial mechanics [86] and gravitational-wave astronomy [Schaefer, 1983; Schäfer, 1982]

$$2\partial_i h_{0i} - \partial_0 h_{kk} = 0\,, \qquad 2\partial_j h_{ij} - \partial_i h_{kk} = -\partial_i h_{00}\,. \tag{3.576}$$

Indeed, it is straightforward to check out that the metric tensor (3.565) satisfies all these gauge conditions, at least, in the linearized post-Minkowskian approximation.

85) The momentum $\pi^{ij} = -\sqrt{-g}\left(K^{ij} - g^{ij}K\right)$ with $K_{ij} = -N\Gamma^0{}_{ij}$, and $K = K_{ii} = K_{11} + K_{22} + K_{33}$.
86) For example, PPN metric tensor [Will, 1993] obeys the Chandrasekhar-Nutku gauge conditions.

One can call the coordinates in which the metric tensor is given by equations (3.565) as the TT-harmonic coordinates [Kopeikin *et al.*, 1999] and they have mathematical advantages of both harmonic and TT coordinates. Thus, the TT-harmonic coordinates allow us to get a full analytic solution of the propagation of a laser beam in the optical interferometer and to significantly suppress the effects produced by the motion of mirrors (with respect to the coordinate grid) caused by gravitational waves. In other words, *all* physical effects of the gravitational waves are relegated in these coordinates to the solution of equations of light propagation. This conclusion is valid in the linearized approximation of general relativity and is not extended to the second post-Minkowskian approximation where gravitational-wave effects on light and motion of observers can not be disentangled because of non-linearity of gravitational field, and have to be analyzed all together. Our TT-harmonic coordinates represent an essential generalization of the standard TT coordinates [Flanagan and Hughes, 2005; Misner *et al.*, 1973; Weber, 1961] because they can be used at arbitrary distances from the isolated system emitting gravitational waves while the standard TT coordinates can be introduced only in those domain of space where gravitational wave can be considered as plane one.

The space around the astronomical system emitting gravitational waves, can be conventionally split in three concentric domains called the near zone, the buffer (intermediate) zone, and the wave zone respectively [Fock, 1964; Misner *et al.*, 1973]. The size of each zone depends on the characteristic wavelength, $\lambdabar$, of gravitational waves emitted by the system. Roughly speaking, the radius of the near zone is, $r < \lambdabar$ and the wave zone has radius $r > \lambdabar$. The buffer zone is a fairly narrow shell of overlapping between the near and wave zones. This separation can be further specialized depending on internal structure and the degree of non-stationarity of a particular astronomical system [Blanchet, 2002a; Blanchet and Damour, 1986; Thorne, 1980]. Gravitational field of the isolated system, given by equations (3.565), can be further decomposed in a Taylor series with respect to powers of the inverse distance $1/r$. Gravitational waves that can propagate freely through space are associated only with the very first term in this expansion that is proportional to $1/r$, because only these terms are essential in the wave zone and contribute to the flux of gravitational radiation coming away from the system [Misner *et al.*, 1973]. Mathematical equation for STF partial derivative from the retarded integrals is as follows [Thorne, 1980]

$$\partial_{<L>}\left[\frac{F(u)}{r}\right] = (-1)^l N_{<L>} \sum_{k=0}^{l} \frac{(l+k)!}{2^k k!(l-k)!} \frac{\overset{(l-k)}{F}(u)}{r^{k+1}} \tag{3.577}$$

where $N_i = x^i/r$ is a unit vector directed from the origin of the coordinates to the field point, $N_{<L>} = N_{<i_1}N_{i_2}...N_{i_l>}$, and the symbol $(l-k)$ above function $F(u)$ denotes time derivative with respect to the retarded time $u = t - r/c$ of the $(l-k)$-th order. Applying this formula in equation (3.565c) and taking only the leading term, one obtains the wave-zone TT projection of the metric field [Blanchet and Damour, 1989]

$$h^{ij}_{\rm rad} = \frac{4G}{c^2 r} P^{ijpq} \sum_{l=2}^{\infty} \frac{1}{c^l l!}\left[N_{L-2} \overset{(l)}{\mathcal{I}}_{pqL-2}(u) - \frac{2l}{(l+1)c} N_{kL-2}\varepsilon_{kn(p} \overset{(l)}{\mathcal{S}}_{q)nL-2}(u)\right] ,$$

(3.578)

where $N_{L-2} = N_{i_1} N_{i_2} ... N_{i_{l-2}}$, $N_{kL-2} = N_k N_{L-2}$, and the symbol (l) above the multipole moments denotes an l-th order time derivative with respect to the retarded time $u = t - r/c$, and the operator of TT projection (3.566) is simplified in the wave zone to

$$P_{ijpq} = \left(\delta_{ip} - N_i N_p\right)\left(\delta_{jq} - N_j N_q\right) - \frac{1}{2}\left(\delta_{ij} - N_i N_j\right)\left(\delta_{pq} - N_p N_q\right) . \tag{3.579}$$

Since the radiation field $h^{ij}_{\rm rad}$ is subject to four algebraic conditions

$$h^{ii}_{\rm rad} = 0 , \qquad N_i h^{ij}_{\rm rad} = 0 , \tag{3.580}$$

only two, out of six components of $h^{ij}_{\rm rad}$ are algebraically independent. They are called h_+ and $h_\times$ components respectively and characterize two states of *polarization* of gravitational radiation [Landau and Lifshitz, 1975; Misner *et al.*, 1973; Weinberg, 1972]. If z-axis of the asymptotically Cartesian coordinate chart $x^i = (x, y, z)$ is taken in the direction, N^i, of propagation of gravitational wave, the two polarizations are given by the following components of the radiative metric: $h_+ = h^{xx}_{\rm rad} = -h^{yy}_{\rm rad}$ and $h_\times = h^{xy}_{\rm rad} = h^{yx}_{\rm rad}$. Rotation around z-axis shows that the metric field is transformed as a tensor wave with helicity equal two. In quantum-mechanical language the weak gravitational wave propagating on a flat Minkowski background corresponds to a massless boson of spin 2 called *graviton* [Grishchuk, 1977].

In higher post-Minkowskian approximations, the radiative multipole moments entering equation (3.578) acquire additional relativistic corrections due to the process of the back-scattering of gravitational waves on the static curvature induced by the total mass (and spin) of the system - the effect known under the name of *tail of gravitational waves*. Radiative multipole moments with the tail-induced contributions are [Blanchet, 2002a; Blanchet *et al.*, 2001a]

$$U_L = \mathcal{I}_L + \frac{2G\mathcal{M}}{c^3}\int_0^{+\infty} d\zeta \ddot{\mathcal{I}}_L(u-\zeta)\left[\ln\zeta + \frac{2l^2+5l+4}{l(l+1)(l+2)} + \sum_{k=1}^{l-2}\frac{1}{k}\right] , \tag{3.581}$$

$$V_L = \mathcal{S}_L + \frac{2G\mathcal{M}}{c^3}\int_0^{+\infty} d\zeta \ddot{\mathcal{S}}_L(u-\zeta)\left[\ln\zeta + \frac{l-1}{l(l+1)} + \sum_{k=1}^{l-1}\frac{1}{k}\right] . \tag{3.582}$$

The radiative moments, U_L and V_L, should replace the multipole moments $\mathcal{I}_L$ and $\mathcal{S}_L$ in equation (3.578) if the contribution from the tails of gravitational waves is required to be taken into account.

Gravitational-wave detector measures variations, $\delta\varrho$, of the radar distance, ϱ, between two mirrors suspended at the ends of the interferometer's arms. The radar distance, $\varrho = c\tau$, is defined as the product of the speed of light, c, in vacuum by the proper time, τ, taken by photon to travel between the mirrors (see section 2.6.4). This definition is given in terms of invariant quantities and, for this reason, is gauge-invariant as explained in section 3.8.6.4. If wavelength $\lambda\!\!\!^{-}$ of incoming gravitational wave is much larger than ϱ, the variation of the radar distance caused by the wave is given by simple equation

$$\frac{\delta\varrho}{\varrho} = \frac{1}{2} k_i k_j h^{ij}_{\rm rad} \,, \tag{3.583}$$

where k_i is the unit vector directed along the line of propagation of the laser beam in the interferometer's arm. It should not be confused with the unit vector N_i pointing to the direction of propagation of the gravitational wave. In the coordinate chart with z-axis directed along vector N^i of propagation of the gravitational wave, the unit vector k^i has the following components: $k^i = (\cos\theta\cos\phi, \cos\theta\sin\phi, \sin\theta)$ where the spherical angles change in the range, $-\pi/2 \le \theta \le \pi/2$, $0 \le \phi < 2\pi$. In these coordinates, equation (3.583) can be expressed in terms of two polarizations of the gravitational radiation. Specifically,

$$\frac{\delta\varrho}{\varrho} = \frac{1}{2}\cos^2\theta\,(h_+\cos 2\phi + h_\times \sin 2\phi) \ . \tag{3.584}$$

Equation (3.584) describes the measured respond of the gravitational-wave detector to the gravitational-wave signal. It shows that the respond is maximal when the gravitational wave propagates perpendicular to the interferometer's arms and it vanishes in case of parallel propagation of the laser beam and the gravitational wave. Though, one has used particular coordinates in order to derive this equation, the detector's respond does not depend on their choice as its left side is defined in terms of gauge-invariant radar distance ϱ and its gauge-invariant variation $\delta\varrho$. Equation (3.583) describes the leading term in the detector's respond. The residual terms in the right side of equation (3.584) as well as the entire mathematical structure of the respond have been studied in papers by Baskaran and Grishchuk [2004] and Kopeikin and Korobkov [2005]; Kopeikin *et al.* [2006].

4
Relativistic Reference Frames

4.1 Historical Background

General theory of relativity is the most powerful theoretical tool for doing experimental gravitational physics both in the solar system and outside of its boundaries. It has passed a multitude of tests with unparallel degree of accuracy [Damour, 2000; Schäfer, 2000; Turyshev, 2009a; Will, 2006]. However, alternative theoretical models are still required for deeper understanding of the geometric nature of spacetime gravitational physics and for studying presumable violations of general relativistic relationships, which are predicted by quantum-gravity theorists. Those hypothetical violations may be observed in the near-future gravitational experiments designed for testing the principle of equivalence [Worden *et al.*, 2000], mapping astrometric positions of stars in the Milky Way with a micro-arcsecond precision [Lindegren, 2009; Lindegren *et al.*, 1995; Lindegren and Perryman, 1996], searching for extra-solar planets [SIM, 2010], testing near-zone gravitomagnetic effects associated with finite speed of propagation of gravity [Fomalont and Kopeikin, 2003; Kopeikin, 2001, 2004], detection of freely-propagating gravitational waves [LIGO, 2010; LISA, 2010], and others.

In the year 2000 the International Astronomical Union (IAU) had adopted new resolutions [Brumberg *et al.*, 1998; Soffel *et al.*, 2003] which lay down a self-consistent general relativistic foundation for applications in modern geodesy, fundamental astrometry, celestial mechanics and spacetime navigation in the solar system. These resolutions combine two independent approaches to the theory of relativistic reference frames in the solar system developed in a series of publications by Brumberg and Kopeikin (BK formalism) and Damour, Soffel and Xu (DSX formalism). The description of the BK and DSX formalisms is given on a joint theoretical platform in a review article [Soffel *et al.*, 2003], which also provides a detailed list of relevant bibliographic references.

It seems worthwhile to extend the IAU 2000 resolutions to incorporate the parameterized post-Newtonian (PPN) formalism [Nordtvedt, 1970; Will, 1993] to the IAU theory of general relativistic reference frames in the solar system, thus, widening domain of applicability of the IAU resolutions to a more general class of gravity theories. We fulfill this task in the present book, thus, making the IAU resolutions fully compatible with the JPL equations of motion used for calculation of ephemerides of major planets, Sun and Moon. These equations generalize the famous Einstein-Infeld-Hoffmann (EIH) equations [Einstein *et al.*, 1938] by including to them two PPN parameters β and γ [Seidelmann and Urban, 2010]. The IAU 2000 resolutions were compatible with the generalized EIH equations only in the case of $\beta = \gamma = 1$.

PPN parameters β and γ signifies the presence of a hypothetical scalar field, which contributes to gravitational force through the coupling to the metric tensor. This causes deviation of the metric tensor from general relativity. The scalar field has not yet been detected but its properties are so appealing that it already plays significant role in modern physics. For example, scalar field helps us to explain the origin of masses of elementary particles [Higgs, 1964], to solve various cosmological problems [Linde, 2000; Sahni and Starobinsky, 2000], to disclose the nature of dark energy in the universe [Caldwell *et al.*, 1998], to develop a gauge-invariant theory of cosmological perturbations [Kopeikin *et al.*, 2001; Ramírez and Kopeikin, 2002] uniting in a natural way the ideas contained in the original gauge-invariant formulation by Bardeen [Bardeen, 1980] with a coordinate-based approach by Lifshitz [1946]; Lifshitz and Khalatnikov [1963]. A recent book by Mukhanov [2005] reviews the results related to the development of Bardeen's theory of cosmological perturbations.

The present book employs a general class of the scalar-tensor theories of gravity initiated in the pioneering works by Jordan [Jordan, 1949, 1959], Fierz [Fierz, 1956] and, especially, Brans and Dicke [Brans and Dicke, 1961; Dicke, 1962a]. A well-written introduction to this theory and other relevant references can be found in the books [Will, 1993] and [Weinberg, 1972]. This class of theories is based on the metric tensor $g_{\alpha\beta}$, representing the gravitational field, and a scalar field ϕ that couples with the metric tensor through function $\theta(\phi)$, which remains unspecified. One stipulates that ϕ and $\theta(\phi)$ are analytic functions, which can be expanded about their cosmological background values $\bar{\phi}$ and $\bar{\theta}$. Existence of the scalar field ϕ brings about the dependence of the universal gravitational constant G on the background value of the field $\bar{\phi}$, which evolves as the universe expands. This book does not deal with cosmology and considers $\bar{\phi}$ as a constant which means that the dynamic of orbital motions take place on time scales much shorter than the Hubble cosmological time 10-15 billion of years. Time dependence of $\bar{\phi}$ may be taken into consideration later by making use of a formal replacement: $\bar{\phi} \rightarrow \bar{\phi}_0 + \dot{\bar{\phi}}(t - t_0)$, where t_0 is an initial epoch, and t is the time. Such a replacement makes the universal gravitational constant G a linear function of time which time dependence may be studied experimentally in various type of observations like timing of binary pulsars or lunar laser ranging.

The purpose of this chapter is to develop a theory of relativistic reference frames in an N-body problem (solar system) with two parameters β and γ of the PPN formalism. There is a principal difficulty in developing such a theory associated with the problem of construction of a local reference frame in the vicinity of each self-gravitating body (Sun, Earth, planet) comprising the N-body system. Standard textbook on the PPN formalism [Will, 1993] does not treat this problem beyond the Newtonian approximation. The original PPN formalism was constructed in a class of asymptotically-flat, global coordinates covering the entire spacetime and moving with respect to each other with a constant speed. The primary PPN coordinates have the origin fixed at the barycenter of the solar system. PPN formalism admits the existence of several long-range fields - scalar, vector, tensor - which are responsible for gravity besides the metric tensor of general theory of relativity. After imposing the boundary conditions on all these fields at infinity the standard PPN metric tensor combines the contribution of these fields to a single expression for the metric tensor, so that they get mixed up with the Newtonian and other general relativistic potentials. It makes technically impossible to single out the alternative fields in the matching procedure that is used to find out relativistic spacetime transformation between the local frame of a self-gravitating body (Earth, Moon, etc.) and the global PPN coordinates. This prevents us to use the law of transformation of the fields, which is a crucial element of the general-relativistic theory of reference frames. Rapidly growing precision of optical and radio astronomical observations as well as demands of gravitational wave astronomy require to work out a PPN theory of such relativistic transformations between the local and global frames generalizing the BK-DSX theory underlying the IAU 2000 resolutions.

It is quite straightforward to construct the local Fermi coordinates along a worldline of a massless particle [Ni and Zimmermann, 1978b]. Such approach can be directly applied in the PPN formalism to construct the Fermi reference frame around a worldline of, for example, an artificial satellite. However, the necessity to account for gravitational self-field effects of the particle (extended body) changes physics of the problem and introduces new mathematical aspects to the existing procedure of construction of the Fermi frame. This was recognized in papers [Klioner and Soffel, 2000; Shahid-Saless and Ashby, 1988], where possible approaches aimed to derive the relativistic transformation between the local (geocentric, planetocentric) and the PPN global coordinates were discussed in the framework of the "canonical" PPN formalism. The approach proposed in [Shahid-Saless and Ashby, 1988] is based on the formalism that was originally worked out by Ashby and Bertotti [1984, 1986] in order to construct a local inertial frame in the vicinity of a self-gravitating body that is a member of an N-body system. Fukushima [1988] has independently developed similar ideas by making use of a slightly different mathematical technique. In the Ashby-Bertotti formalism the PPN metric tensor is taken in its standard form [Will, 1993] and massive bodies are treated as point-like massive particles without rotation (massive monopoles). Construction of a local inertial frame in the vicinity of such massive particle requires to impose some specific restrictions on the worldline of the particle. Namely, the particle is assumed to be moving along a geodesic defined on the "effective" spacetime manifold which is obtained by elimination of the body under consideration from the expression for the standard PPN metric tensor. This way of introduction of the "effective" background manifold is possible but not defined uniquely, thus, bringing about an ambiguity in the construction of the "effective" manifold [Kopejkin, 1988a]. Moreover, the assumption that the bodies are point-like and non-rotating is not quite appropriate for the tasks of modern geodesy and relativistic celestial mechanics of the solar system. For example, planets in the solar system have appreciable rotational speeds and noticeable higher-order multipole moments. Gravitational interaction of the multipole moments of a celestial body with the external tidal field of other planets and Sun does not allow the body to move along the geodesic [Kopejkin, 1988a]. Deviation of the body's center-of-mass worldline from the geodesic can be significant and important in numerical calculations of planetary ephemerides (see, e.g., [Newhall *et al.*, 1983] and discussion on page 307 in [Kopejkin, 1991a]) and must be taken into account, when one constructs a theory of the relativistic reference frames in the N-body system.

Different approach to the problem of construction of a local (geocentric) reference frame in the PPN formalism was proposed in the paper by Klioner and Soffel [Klioner and Soffel, 2000]. These authors have used a phenomenological approach, which does not demand that the PPN metric tensor in local coordinates is a solution of the field equations of a specific theory of gravity. The intention was to make the parameterized formalism of the relativistic reference frames as general as possible. To this end, the paper [Klioner and Soffel, 2000] stipulated that the structure of the metric tensor written down, for example, in the local geocentric reference frame must possess the following properties:

A. gravitational field of external bodies (Sun, Moon, planets) is represented in the vicinity of the Earth in the form of relativistic tidal potentials which should reduce in the Newtonian limit to the Newtonian tidal potential,

B. switching off the tidal potentials must reduce the metric tensor of the local coordinates to its standard PPN form given in the book by Will [Will, 1993, Table 4.1]

Direct calculations revealed that under assumptions made in [Klioner and Soffel, 2000] the properties (A) and (B) work in general relativity but can not be satisfied simultaneously in the proposed extension of the PPN formalism. This is a direct consequence of a subtle inconsistency presents in the matching procedure applied in [Klioner and Soffel, 2000] in order to transform the metric tensor from the local geocentric coordinates to the global barycentric ones.

More specifically, at each step of the matching procedure four kinds of different terms in the metric tensor have been singled out and equated independently in the corresponding matching equations for the metric tensor (for more details see [Klioner and Soffel, 2000, page 024019-10]):

M1 - the terms depending on the interior structure of the body under consideration (Earth);
M2 - the terms which are functions of time only;
M3 - the terms which are linear functions of the local spatial coordinates;
M4 - the terms which are quadratic and higher-order polynomials of the local coordinates.

It is implicitly assumed by Klioner and Soffel [2000] that the application of these conditions will not give rise to contradictions with other equations and phenomenological principles of the parameterized gravitational theory in the curved spacetime. Keeping this in mind, the matching conditions are implemented in order to solve the four types of the matching equations.

Unfortunately, a scrutiny analysis [Kopeikin and Vlasov, 2004] of the procedure for construction of reference frames in the PPN formalism, proposed in [Klioner and Soffel, 2000], shows that it does not comply with the relativistic equations for gravitational field. First of all, the number of the matching equations is rigidly linked to the number of the field variables used in the gravitational theory under consideration and can not be chosen arbitrary. In general theory of relativity the only gravitational field variable is the metric tensor. Therefore, in general relativity it is necessary and sufficient to write down the matching equations for the metric tensor only. However, any alternative theory of gravity has additional long-range fields (scalar, vector, tensor), which contribute to the gravitational field as well as the metric tensor. Hence, in any of these theories one has to write down the matching equations not only for the metric tensor but also for the additional fields in order to get correspondence between the geocentric and barycentric coordinates parameterized by the PPN parameters. This problem has not been addressed in [Klioner and Soffel, 2000], which assumed that it will be sufficient (like in general relativity) to write down and to solve the matching equations merely for the metric tensor in order to obtain complete information about the structure of the parameterized post-Newtonian transformation from the local to global frames. This might incidentally work for some particular case of alternative theory of gravity but the result of matching would remain rather formal whereas the physical content and the degree of applicability of the post-Newtonian transformation between the coordinates will have remained unclear.

The present book relies upon different technique that was worked out in [Kopeikin and Vlasov, 2004] and is applicable to a quite general class of the scalar-tensor theories of gravity operating with two fields - the metric tensor and a scalar field. Our approach consistently uses the matching equation for the metric tensor along with that for the scalar field, which are direct consequences of the gravity field equations. Our results modify essentially those obtained by Klioner and Soffel [Klioner and Soffel, 2000]. This modification is a clear indication that the phenomenological construction of the reference frames in the framework of the PPN formalism introduces too many degrees of the gauge freedom which can not be uniquely fixed due to the absence of clearly formulated field equations. Phenomenological restriction of this freedom can be done in many different ways *ad liberum*, thus leading to additional (researcher-dependent) ambiguity in the interpretation of relativistic effects in the local (geocentric) reference frame. This is definitely not the goal of the physical theory of reference frames, which intends to get physically-meaningful interpretation of the results of gravitational experiments.

The main drawback of the Klioner-Soffel approach [Klioner and Soffel, 2000] to the parameterized theory of relativistic reference frames is that the metric tensor in the local coordinates is not determined from the field equations but is supposed to be found from the four matching conditions (M1)-(M4) indicated above. It forces the researchers to introduces a function Ψ [Klioner and Soffel, 2000, equation (3.33)], which is not a solution of any gravity field equations. Moreover, the first of the matching conditions (M1) requires that all internal potentials generated by Earth's matter can be fully segregated from the other terms in the metric tensor. It can be done in general relativity and in the scalar-tensor theory of gravity as a consequence of the field equations [Kopeikin and Vlasov, 2004; Kopejkin, 1988a]. We shall also demonstrate how it works in the present book. However, complete separation of the internal potentials describing interior gravitational field of the Earth from the other terms in matching equations may not work out in arbitrary alternative theory of gravity admitting existence of additional long-range vector and/or tensor fields. Thus, the overall class of gravity theories to which the first of the matching conditions (M1) can be applied remains unclear and yet, has to be determined.

Our point of view is that in order to eliminate any mathematical inconsistency and undesirable ambiguities in the construction of the PPN metric tensor in the local reference frame of the body under consideration and to apply a rigorous procedure for derivation of the relativistic coordinate transformations between the local and global coordinates, a specific theory of gravity must be unavoidably used. It makes the field equations known and the number of the field variables entering the PPN metric tensor in the local coordinates becomes exactly equal to the number of matching equations. Hence, all of them can be determined unambiguously without any additional postulates. Thus, a parameterized theory of relativistic reference frames in a self-gravitating N-body system is built by making use of the following procedure:

1) Chose a class of gravitational theories with a well-defined field variables and a system of the field equations.
2) Impose a specific gauge condition on the metric tensor and the other fields to single out a class of admissible coordinates, and to reduce the field equations to a solvable form.
3) Solve the reduced field equations in the global coordinates, $x^\alpha = (x^0, x^i)$, by imposing a fall-off boundary condition on the field variables at infinity.
4) Solve the reduced field equations in the local coordinates, $w^\alpha = (w^0, w^i)$, defined in the vicinity of a worldline of the center-of-mass of each body. This specifies N local coordinate systems.
5) Make use of the residual gauge freedom to eliminate nonphysical degrees of freedom from the field variables and to find out the most general structure of the spacetime coordinate transformation between the global and local coordinates.
6) Transform the metric tensor and the other fields from the local coordinates to the global ones by making use of the general form of the coordinate transformations found at the previous step.
7) Derive from this transformation a set of matching (first-order differential and/or algebraic) equations for all functions entering the field variables and the coordinate transformations.
8) Solve the matching equations and determine all unknown functions entering the matching equations explicitly.

This procedure works perfect in the case of general relativity [Soffel *et al.*, 2003] and is valid also in the class of the scalar-tensor theories of gravity [Kopeikin and Vlasov, 2004]. We do not elaborate on this procedure in the case of vector-tensor and tensor-tensor theories of gravity. This problem is progressively complicated and supposed to be solved somewhere else.

The scalar-tensor theory of gravity employed in this book operates with one tensor, $g_{\alpha\beta}$, and one scalar, ϕ, fields. The tensor field $g_{\alpha\beta}$ is the metric tensor of the Riemannian spacetime manifold. The scalar field ϕ couples to the metric tensor, and is generated by matter of the gravitating bodies comprising N-body system. One assumes that the N-body system (solar system, binary star) consists of extended bodies, which gravitational field is weak everywhere. Besides, the characteristic velocity of motion of matter of the bodies is slow. These assumptions allow us to use the post-Newtonian approximation (PNA) scheme developed earlier by various researchers in order to find solutions of the scalar-tensor field equations with non-singular distribution of matter in space. PNA solves the gravity field equations by making use of Taylor expansions with respect to the weak-field and slow-motion parameters. The reader is referred to the cornerstone works [Anderson and Decanio, 1975; Anderson *et al.*, 1982; Blanchet and Damour, 1984b; Brumberg, 1972, 1991; Chandrasekhar, 1965; Chandrasekhar and Esposito, 1970; Chandrasekhar and Nutku, 1969; Fock, 1964; Futamase and Itoh, 2007; Futamase and Schutz, 1983; Infeld and Plebanski, 1960] which reflect different aspects of the post-Newtonian approximations. The present book extends and improves general relativistic calculations performed in papers [Brumberg and Kopeikin, 1989; Brumberg and Kopejkin, 1989a,b; Kopeikin, 1989a,b; Kopejkin, 1988a, 1991a,b]. It takes into account the post-Newtonian definition of multipole moments of an isolated self-gravitating body (or a system of bodies) introduced by Thorne [1980] and further developed by Blanchet and Damour [1989]; Damour and Iyer [1991a].

The approach, used in this book, does not specify the internal structure of the bodies so that one's consideration is not restricted with the case of a perfect fluid as it is customary done in the PPN formalism [Will, 1993]. This extension of the formalism is important for subsequent consideration of the Newtonian and post-Newtonian respond of the gravitating bodies to elastic deformations caused by tides and other phenomena. Interesting details of the post-Newtonian development of the theory of elasticity in application to geophysics are discussed in papers [Xu *et al.*, 2005, 2003].

4.2 Isolated Astronomical Systems

4.2.1 Field equations in the scalar-tensor theory of gravity

The purpose of this book is to develop a parameterized theory of relativistic reference frames for N-body gravitating system in the framework of PPN formalism, which contains 10 phenomenological parameters [Will, 1993]. Michelson-Morley and Hughes-Drever experiments strongly restricted possible violations of the local isotropy of space, whereas Eötvös-Dicke-Braginsky experiments verified a weak equivalence principle with very high precision [Will, 1993, 2006]. These remarkable experimental achievements along with the modern theoretical attempts to unify gravity with other fundamental fields strongly restrict the class of viable alternative theories of gravity, and, very likely, reduce the number of the parameters of the standard PPN formalism [Will, 1993] from 10 to two - β and γ. Experimental testing of the Lorentz-invariance of the gravity field equations (that is, Einstein's principle of relativity for gravitational field) requires that one introduces more parameters [Kopeikin, 2004; Kostelecky, 2008; Will, 1993] but this book stipulates that the Lorentz-invariance is not violated. Parameters β and γ appear naturally in the class of alternative theories of gravity with one or several scalar fields [Damour and Esposito-Farese, 1992; Will, 1993], which can be taken as a starting point for making further generalization of the IAU 2000 resolutions on relativistic reference frames [Soffel *et al.*, 2003]. For this reason, this book deals only with the class of scalar-tensor theories of gravity assuming that additional vector and/or tensor fields do not exist. For simplicity, one focuses on the case with one real-valued scalar field ϕ loosely coupled with the tensor gravity by means of a coupling function $\theta(\phi)$ [Will, 1993].

One stipulates that both the gravitational and the scalar field are generated by matter of the only one astronomical system comprising of N extended bodies which matter occupies a finite domain in space. Such an astronomical system is called isolated [Dixon, 1979; Fock, 1964; Papapetrou, 1951a] and the solar system consisting of Sun, Earth, Moon, and other planets is its particular example. Astronomical systems like a galaxy, a globular cluster, a binary star, etc. typify other specimens of the isolated systems. A number of bodies in the N-body system which must be taken into account depends on the accuracy of astronomical observations and is determined mathematically by the magnitude of residual terms which one must retain in calculations to construct relativistic theory of reference frames being compatible with the accuracy of the observations. Since other gravitating bodies residing outside of the N-body system, are ignored, the spacetime can be considered on the global scale as asymptotically-flat so the metric tensor $g_{\alpha\beta}$ at infinity is the Minkowski metric $\eta_{\alpha\beta} = \mathrm{diag}(-1, +1, +1, +1)$.

Field equations in the scalar-tensor theory are derived from the action [Will, 1993]

$$S = \frac{c^4}{16\pi}\int\left(\phi R - \theta(\phi)\frac{\partial^\alpha\phi\partial_\alpha\phi}{\phi} - \frac{16\pi}{c^4}L(g_{\mu\nu},\psi)\right)\sqrt{-g}\,d^4x\,, \tag{4.1}$$

where the first, second and third terms in the right side of equation (4.1) are the Lagrangian densities of gravitational field, scalar field and matter respectively, $g = \det[g_{\alpha\beta}] < 0$ is the determinant of the metric tensor $g_{\alpha\beta}$, R is the Ricci scalar, ψ indicates dependence of the Lagrangian L on the matter fields, and $\theta(\phi)$ is the above-mentioned, coupling function, which is kept arbitrary. The action (4.1) makes the class of the scalar-tensor theories to be sufficiently large. For the sake of simplicity, let us postulate that the self-coupling potential of the scalar field is identically zero so that the scalar field does not interact with itself, and there is no such term in the action. This is because there is no reason to expect that this potential can lead to measurable relativistic effects within the boundaries of the solar system. However, this potential can be important in the case of a strong gravitational field and its inclusion to the post-Newtonian theory of coalescing binary neutron stars can lead to interesting physical consequences [Damour and Esposito-Farese, 1992].

Equations of the gravitational field are obtained by taking variational derivative of the action (4.1) with respect to $g_{\alpha\beta}$ and it's spatial derivatives [Misner *et al.*, 1973; Weinberg, 1972]. The variational principle yields

$$R_{\mu\nu} = \frac{8\pi}{\phi c^2}\left(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T\right) + \theta(\phi)\frac{\partial_\mu\phi\partial_\nu\phi}{\phi^2} + \frac{1}{\phi}\left(\nabla_{\mu\nu}\phi + \frac{1}{2}g_{\mu\nu}\Box_{\rm g}\phi\right), \tag{4.2}$$

where

$$\Box_{\rm g} \equiv g^{\mu\nu}\frac{\partial^2}{\partial x^\mu\partial x^\nu} - g^{\mu\nu}\Gamma^\alpha_{\mu\nu}\frac{\partial}{\partial x^\alpha} \tag{4.3}$$

is the scalar Laplace-Beltrami operator [Dubrovin *et al.*, 1984] and $T_{\mu\nu}$ is the energy-momentum tensor of matter comprising the N-body system, $T = T^\alpha{}_\alpha$ is the trace of the tensor. It is defined by the variational equation [Landau and Lifshitz, 1975; Misner *et al.*, 1973]

$$\frac{1}{2}\sqrt{-g}\,T_{\mu\nu} \equiv \frac{\partial(\sqrt{-g}L)}{\partial g^{\mu\nu}} - \frac{\partial}{\partial x^\alpha}\frac{\partial(\sqrt{-g}L)}{\partial g^{\mu\nu}{}_{,\alpha}}\,, \tag{4.4}$$

where $g^{\mu\nu}{}_{,\alpha} \equiv \partial_\alpha g^{\mu\nu}$, and will be specified more explicitly in next section.

Equation governing the scalar field is obtained by variation of the action (4.1) with respect to ϕ and it's spatial derivatives. After making use of the contracted form of equation (4.2) it yields

$$\Box_{\rm g}\phi = \frac{1}{3+2\theta(\phi)}\left(\frac{8\pi}{c^2}T - \partial_\alpha\phi\partial^\alpha\phi\,\frac{d\theta}{d\phi}\right), \tag{4.5}$$

which shows that the source of the scalar field is the trace of the energy-momentum tensor of matter as well as the kinetic energy of the scalar field itself if the derivative $d\theta/d\phi \neq 0$.

In what follows, one will also utilize another version of the Einstein equations (4.2) which is obtained after conformal transformation of the metric tensor

$$\tilde{g}_{\mu\nu} = \frac{\phi}{\phi_0} g_{\mu\nu} \qquad , \qquad \tilde{g}^{\mu\nu} = \frac{\phi_0}{\phi} g^{\mu\nu} . \tag{4.6}$$

Here ϕ_0 denotes the background value of the scalar field which is explained in more detail in (4.13). It is worth noting that the determinant $\tilde{g}$ of the conformal metric tensor relates to the determinant g of the metric $g_{\mu\nu}$ as $\tilde{g} = (\phi/\phi_0)^4 g$. The conformal transformation of the metric tensor leads to the conformal transformation of the Christoffel symbols and the Ricci tensor as well [Wald, 1984, Appendix D]. Denoting the conformal Ricci tensor by $\tilde{R}_{\mu\nu}$ one can reduce the field equations (4.2) to a simpler form

$$\tilde{R}_{\mu\nu} = \frac{8\pi}{\phi c^2}\Big(T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T\Big) + \frac{2\theta(\phi) + 3}{2\phi^2} \partial_\mu \phi \partial_\nu \phi . \tag{4.7}$$

The metric tensor $g_{\mu\nu}$ is called the physical (Jordan-Fierz-frame) metric [Will, 1993] because it is used in real measurements of time intervals and space distances. The conformal metric $\tilde{g}_{\mu\nu}$ is called the Einstein-frame metric. Its main advantage is that this metric is in many technical aspects more convenient for doing calculations than the Jordan-Fierz frame metric. Indeed, if the last (quadratic with respect to the scalar field) term in equation (4.7) was omitted, it would make it look similar to the Einstein equations of general relativity with the gravitational constant $G = 1/phi$. This dependence may explain the Dirac's hypothesis [Dirac, 1937] about the gradual change of the gravitational constant as the Universe expands [Jordan, 1959]. Nevertheless, one prefers to construct the parameterized post-Newtonian theory of reference frames for N-body problem in terms of the Jordan-Fierz-frame metric $g_{\mu\nu}$ in order to avoid the unnecessary conformal transformation to convert the results of the calculations to a physically-measurable form.

4.2.2 The energy-momentum tensor

In order to find the gravitational field and determine the motion of the bodies comprising the N-body system one needs:

1) to specify a model of matter that makes up the N-body system,
2) to specify the gauge condition imposed on the metric tensor $g_{\alpha\beta}$,
3) to simplify (reduce) the field equations by making use of the chosen gauge condition,
4) to solve the reduced field equations for the metric tensor $g_{\alpha\beta}$ and the other fields,
5) to derive equations of motion of the bodies that are consistent with the solutions of the field equations.

This program will be completed in the present book for the case of an isolated system of N bodies moving slowly and having weak gravitational field everywhere. In principle, the formalism, which will be developed, allows us to treat N-body systems consisting of black holes, neutron stars, or other compact relativistic bodies if the strong-field space domains are excluded after appropriate matching of them to the weak-field zones [Damour, 1987; D'Eath, 1975a,b]. The most important example of the weak-field and slow-motion N-body system is the solar system and one can keep this example in mind for future practical applications of the PPN formalism developed in the present book.

One assumes that the N-body system is isolated which means that the tidal influence of other matter in the Milky Way on this system is neglected. Besides, one stipulates that there is no flux of gravitational radiation incoming to the system from outside. This boundary condition picks up the retarded solution of the wave equations describing the evolution of the metric tensor and the scalar field. Thus, the spacetime is considered as asymptotically-flat very far away outside of the system so that the barycenter of the N-body system is either at rest or moves with respect to the asymptotically-flat space along a straight line with a constant velocity.

One assumes that the matter comprising the bodies of the N-body system is described by the energy-momentum tensor with some physically-reasonable equation of state, which is not specified because it does not affect the result of analytical calculations which follow. Following Fock [Fock, 1964] and Papapetrou [Papapetrou, 1951a], who developed similar ideas, the energy-momentum tensor is defined as

$$c^2 T^{\mu\nu} = \rho\left(c^2 + \Pi\right) u^\mu u^\nu + \pi^{\mu\nu} , \tag{4.8}$$

where ρ and Π are the density and the specific internal energy of matter in the co-moving frame, $u^\alpha = dx^\alpha / cd\tau$ is the dimensionless four-velocity of the matter with τ being the proper time along the worldline of the matter's particle, and $\pi^{\alpha\beta}$ is the anisotropic tensor of stresses that is symmetric and orthogonal to four-velocity

$$u^\alpha \pi_{\alpha\beta} = 0 . \tag{4.9}$$

Original PPN formalism [Will, 1993] treats the matter of the N-body system as a perfect fluid for which

$$\pi^{\alpha\beta} = \left(g^{\alpha\beta} + u^\alpha u^\beta\right) p , \tag{4.10}$$

where p is an isotropic pressure. Perfect-fluid approximation is not sufficient in the Newtonian theory of motion of the solar system bodies where tidal phenomena and dissipative forces play essential role [Zharkov and Trubitsyn, 1978]. It is also inappropriate for consideration of last stages of coalescing binary systems for which a full relativistic theory of tidal deformations must be worked out. For this reason one abandons the perfect-fluid approximation and incorporate the anisotropic stresses to the PPN formalism. General relativistic consideration of the anisotropic stresses has been also achieved in papers [Damour *et al.*, 1991, 1992, 1993; Xu *et al.*, 2005, 2003].

Conservation of the energy-momentum tensor $\nabla_\nu T^{\mu\nu} = 0$ leads both to the equation of continuity [Misner *et al.*, 1973]

$$\nabla_\alpha \left(\rho u^\alpha\right) = \frac{1}{\sqrt{-g}} \partial_\alpha \left(\rho \sqrt{-g} u^\alpha\right) = 0 , \tag{4.11}$$

and to the second law of thermodynamics that is expressed as a differential relationship between the change in the specific internal energy and the stress tensor and matter's deformation

$$\rho u^\alpha \partial_\alpha \Pi + \pi^{\alpha\beta} u_{\alpha;\beta} = 0 \,. \tag{4.12}$$

The above equations fully define the structure of the tensor of energy-momentum and will be employed later for solving the field equations, and for derivation of the orbital and rotational equations of motion of the bodies.

4.2.3
Basic principles of the post-Newtonian approximations

Field equations (4.2) and (4.5) all together represent a system of eleventh non-linear differential equations in partial derivatives and one has to find their solutions for the case of an N-body system. General relativity is obtained as a limiting case when the field ϕ is constant and, thus, unobserved. General relativity has only ten field equations for the metric tensor. Equations (4.2) and (4.5) are complicated due to their non-linearity, and can be solved only by making use of approximation methods. Two basic methods for solving these equations are known as the post-Minkowski [Blanchet and Damour, 1984b, 1986; Damour, 1987] and the post-Newtonian [Damour, 1987] approximation schemes. The post-Newtonian approximation (PNA) scheme deals with slowly-moving bodies having weak gravitational field. It makes PNA more appropriate for constructing theory of the relativistic reference frames and relativistic celestial mechanics in the solar system than the post-Minkowskian approximation (PMA) scheme, which operates with fast moving bodies and solves the gravity field equations in terms of retarded gravitational potentials. For this reason, the PNA scheme is sufficient in this chapter though some elements of the post-Minkowskian approximation (PMA) scheme will be used for definition of the multipole moments of the gravitational field. PMA will be also used for discussion of relativistic astrometry in chapter 7.

Post-Newtonian approximations are based on assumption that expansion of the metric tensor in the near zone of a source of gravity can be done in inverse powers of the fundamental speed c that is equal to the speed of light in vacuum. This expansion may be not analytic in higher post-Newtonian approximations in a certain class of coordinates [Blanchet and Damour, 1986; Kates and Kegeles, 1982]. Exact formulation of a set of basic axioms required for doing the post-Newtonian expansion was given by Rendall [Rendall, 1992]. Practically, it requires to have several small parameters characterizing the source of gravity. They are: $\epsilon_i \sim v_i/c$, $\epsilon_e \sim v_e/c$, and $\eta_i \sim U_i/c^2$, $\eta_e \sim U_e/c^2$, where v_i is a characteristic velocity of motion of matter inside a body, v_e is a characteristic velocity of the relative motion of the bodies with respect to each other, U_i is the internal gravitational potential of each body, and U_e is the external gravitational potential between the bodies. If one denotes a characteristic radius of a body as L and a characteristic distance between the bodies as R, the internal and external gravitational potentials will be $U_i \simeq GM/L$ and $U_e \simeq GM/R$, where M is a characteristic mass of the body. Due to the virial theorem of the Newtonian gravity [Binney and Tremaine, 2008] the small parameters are not independent. Specifically, one has $\epsilon_i^2 \sim \eta_i$ and $\epsilon_e^2 \sim \eta_e$. Hence, parameters ϵ_i and ϵ_e are sufficient in doing post-Newtonian approximations. Because within the solar system these parameters do not significantly differ from each other, they will be not distinguished when doing the post-Newtonian iterations. In particular, notation $\epsilon \equiv 1/c$ is used to mark the powers of the fundamental speed c in the post-Newtonian terms. This parameter is also considered as a primary parameter of the PNA scheme to each allother parameters are approximately equal, for example, $\epsilon_i = \epsilon v_i$, $\eta_i = \epsilon^2 U_i$, etc.

One assumes that the scalar field can be expanded in power series around its background value ϕ_0, that is

$$\phi = \phi_0(1+\zeta)\,, \tag{4.13}$$

where ζ is dimensionless perturbation of the scalar field around its background value. The background value ϕ_0 of the scalar field can depend on time due to the cosmological evolution of the universe but, according to Damour and Nordtvedt [1993b], such time-dependence is expected to be rather insignificant due to the, presumably, rapid decay of the scalar field in the course of cosmological evolution following immediately after the Big Bang. According to theoretical expectations [Damour and Nordtvedt, 1993b] and experimental data [Damour, 2000; Schäfer, 2000; Will, 1993] the variable part ζ of the scalar field must have a very small magnitude so that one can expand all quantities depending on the scalar field in Taylor series using the maximal value of the absolute magnitude ζ as a dimensionless small parameter. In particular, decomposition of the coupling function $\theta(\phi)$ can be written as

$$\theta(\phi) = \omega + \omega'\,\zeta + O(\zeta^2)\,, \tag{4.14}$$

where $\omega \equiv \theta(\phi_0)$, $\omega' \equiv (d\theta/d\zeta)_{\phi=\phi_0}$ are constants, and one imposes the fall-off boundary condition on the scalar field perturbation ζ that assumes ζ approaches zero as the distance from the N-body system grows to infinity.

Accounting for the decomposition of the scalar field and equation (4.5) the gravity field equations (4.2) take the following form

$$\begin{aligned} R_{\mu\nu} &= \frac{8\pi\mathcal{G}}{(1+\zeta)c^2}\left[T_{\mu\nu} - \frac{\omega+1}{2\omega+3}\,g_{\mu\nu}T\left(1+\frac{\omega'\,\zeta}{(\omega+1)(2\omega+3)}\right)\right] \\ &\quad -\frac{1}{2}g_{\mu\nu}\frac{\omega'\partial_\alpha\zeta\partial^\alpha\zeta}{2\omega+3} + \frac{\omega\partial_\mu\zeta\partial_\nu\zeta}{(1+\zeta)^2} + \frac{\nabla_{\mu\nu}\zeta}{1+\zeta}\,, \end{aligned} \tag{4.15}$$

where $\mathcal{G} = 1/\phi_0$ is the bare background value of the universal gravitational constant and one has taken into account only the linear and quadratic terms with respect to the perturbation of the scalar field, which is sufficient for developing the post-Newtonian parameterized theory of the reference frames in the solar system.

We look for solutions of the field equations (4.15) in the form of a Taylor expansion of the metric tensor and the scalar field with respect to the post-Newtonian parameter ϵ such that

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}\,, \tag{4.16}$$

and

$$h_{00} = \epsilon^2 \overset{(2)}{h}_{00} + \epsilon^4 \overset{(4)}{h}_{00} + O(\epsilon^5)\,, \tag{4.17}$$

$$h_{0i} = \epsilon \overset{(1)}{h}_{0i} + \epsilon^3 \overset{(3)}{h}_{0i} + O(\epsilon^5)\,, \tag{4.18}$$

$$h_{ij} = \epsilon^2 \overset{(2)}{h}_{ij} + \epsilon^4 \overset{(4)}{h}_{ij} + O(\epsilon^5)\,, \tag{4.19}$$

$$\zeta = \epsilon^2 \overset{(2)}{\zeta} + \epsilon^4 \overset{(4)}{\zeta} + O(\epsilon^5)\,, \tag{4.20}$$

where $\overset{(n)}{h}_{\alpha\beta}$ and $\overset{(n)}{\zeta}$ denote terms of the order ϵ^n $(n = 1, 2, 3...)$. It has been established that the post-Newtonian expansion of the metric tensor in general theory of relativity is, in fact, non-analytic [Damour, 1987; Kates and Kegeles, 1982]. However, the non-analytic terms emerge much later, in the approximations of higher post-Newtonian order and do not affect the results of this section since the first post-Newtonian approximation operates only with analytic terms $\overset{(2)}{h}_{00}$, $\overset{(4)}{h}_{00}$, $\overset{(1)}{h}_{0i}$, $\overset{(3)}{h}_{0i}$, $\overset{(2)}{h}_{ij}$ and $\overset{(2)}{\zeta}$.

The reader may have noticed that the post-Newtonian expansion of the spacetime metric perturbation h_{0i} starts from the terms of the order of ϵ. This term is fully associated with the definition of the local coordinate basis at the origin of the coordinates. The basis may be not orthogonal and/or locally rotating with angular speed Ω^i, which is the reason behind this "anomalous" ϵ-behavior of h_{0i} [Misner *et al.*, 1973; Ni and Zimmermann, 1978b]. We shall keep the linear with respect to ϵ term in h_{0i} for some time and discard it after its role in the construction of the relativistic reference frames will be completely established. In what follows, the simplified notations for the metric tensor and scalar field perturbations will be used:

$$N \equiv \overset{(2)}{h}_{00}\,, \quad L \equiv \overset{(4)}{h}_{00}\,, \quad N_i \equiv \overset{(1)}{h}_{0i}\,, \quad L_i \equiv \overset{(3)}{h}_{0i}\,, \quad H_{ij} \equiv \overset{(2)}{h}_{ij}\,, \quad H \equiv \overset{(2)}{h}_{kk}\,, \tag{4.21}$$

and

$$\varphi \equiv (\omega + 2)\, \overset{(2)}{\zeta}\;. \tag{4.22}$$

The scalar-tensor theory of gravity stipulates that the spacetime manifold admits the affine connection without torsion [Misner *et al.*, 1973; Wald, 1984; Weinberg, 1972]. Christoffel symbols for the metric tensor (4.16) are defined by standard expression

$$\Gamma^\alpha_{\beta\gamma} = \frac{1}{2} g^{\alpha\delta}\left(\partial_\gamma g_{\beta\delta} + \partial_\beta g_{\gamma\delta} - \partial_\delta g_{\beta\gamma}\right) = \frac{1}{2}\left(\eta^{\alpha\delta} - h^{\alpha\delta}\right)\left(\partial_\gamma h_{\beta\delta} + \partial_\beta h_{\gamma\delta} - \partial_\delta h_{\beta\gamma}\right) + O\left(h^3\right)\,, \tag{4.23}$$

where the metric tensor components are taken from equations (4.17)-(4.19). Calculation results in

$$\Gamma^0_{00} = -\frac{\epsilon^3}{2}\left(\overset{(2)}{h}_{00,0} + \overset{(1)}{h}_{0i}\overset{(1)}{h}_{0i,0}\right) + O(\epsilon^4)\,, \tag{4.24}$$

$$\Gamma^0_{0i} = -\frac{\epsilon^2}{2}\left[\overset{(2)}{h}_{00,i} - \overset{(1)}{h}_{0j}\left(\overset{(1)}{h}_{0i,j} - \overset{(1)}{h}_{0j,i}\right)\right] + O(\epsilon^3)\,, \tag{4.25}$$

$$\begin{aligned}\Gamma^i_{00} &= \epsilon^2\left(\overset{(1)}{h}_{0i,0} - \frac{1}{2}\overset{(2)}{h}_{00,i}\right) \\ &+\epsilon^4\left(\overset{(3)}{h}_{0i,0} - \frac{1}{2}\overset{(4)}{h}_{00,i} - \overset{(2)}{h}_{ij}\overset{(1)}{h}_{0j,0} + \frac{1}{2}\overset{(1)}{h}_{0i}\overset{(2)}{h}_{00,0} + \frac{1}{2}\overset{(2)}{h}_{ik}\overset{(2)}{h}_{00,k}\right) + O(\epsilon^5)\,,\end{aligned} \tag{4.26}$$

$$\begin{aligned}\Gamma^i_{0k} &= \epsilon\left(\overset{(1)}{h}_{0i,k} - \overset{(1)}{h}_{0k,i}\right) \\ &+\frac{\epsilon^3}{2}\left[\overset{(3)}{h}_{0i,k} - \overset{(3)}{h}_{0k,i} + \overset{(2)}{h}_{ik,0} + \overset{(2)}{h}_{ij}\left(\overset{(1)}{h}_{0k,j} - \overset{(1)}{h}_{0j,k}\right) + \overset{(1)}{h}_{0i}\overset{(2)}{h}_{00,k}\right] + O(\epsilon^4)\,,\end{aligned} \tag{4.27}$$

$$\begin{aligned}\Gamma^0_{ik} &= -\epsilon\left(\overset{(1)}{h}_{0i,k} + \overset{(1)}{h}_{0k,i}\right) \\ &-\frac{\epsilon^3}{2}\left[\overset{(3)}{h}_{0i,k} + \overset{(3)}{h}_{0k,i} - \overset{(2)}{h}_{ik,0} + \overset{(2)}{h}_{00}\left(\overset{(1)}{h}_{0i,k} + \overset{(1)}{h}_{0k,i}\right)\right. \\ &\left.- \overset{(1)}{h}_{0j}\left(\overset{(2)}{h}_{ij,k} + \overset{(2)}{h}_{kj,i} - \overset{(2)}{h}_{ik,j}\right)\right] + O(\epsilon^4)\,,\end{aligned} \tag{4.28}$$

$$\Gamma^i_{jk} = \frac{\epsilon^2}{2}\left[\overset{(2)}{h}_{ij,k} + \overset{(2)}{h}_{ik,j} - \overset{(2)}{h}_{jk,i} + \overset{(1)}{h}_{0i}\left(\overset{(1)}{h}_{0j,k} + \overset{(1)}{h}_{0k,j}\right)\right] + O(\epsilon^3)\,. \tag{4.29}$$

In what follows, the calculations neglect all terms which are quadratic with respect to $\overset{(1)}{h}_{0i}$. Dropping out the corresponding terms in equations (4.24)-(4.29) yields

$$\Gamma^0_{00} = -\frac{\epsilon^3}{2}\overset{(2)}{h}_{00,0} + O(\epsilon^4)\,, \tag{4.30}$$

$$\Gamma^0_{0i} = -\frac{\epsilon^2}{2}\overset{(2)}{h}_{00,i} + O(\epsilon^3)\,, \tag{4.31}$$

$$\Gamma^i_{00} = \epsilon^2\left(\overset{(1)}{h}_{0i,0} - \frac{1}{2}\overset{(2)}{h}_{00,i}\right) + \epsilon^4\left(\overset{(3)}{h}_{0i,0} - \frac{1}{2}\overset{(4)}{h}_{00,i} + \frac{1}{2}\overset{(2)}{h}_{ik}\overset{(2)}{h}_{00,k}\right) + O(\epsilon^5)\,, \tag{4.32}$$

$$\Gamma^i_{0k} = \epsilon\left(\overset{(1)}{h}_{0i,k} - \overset{(1)}{h}_{0k,i}\right) + \frac{\epsilon^3}{2}\left(\overset{(3)}{h}_{0i,k} - \overset{(3)}{h}_{0k,i} + \overset{(2)}{h}_{ik,0}\right) + O(\epsilon^4)\,, \tag{4.33}$$

$$\Gamma^0_{ik} = -\frac{\epsilon^3}{2}\left(\overset{(3)}{h}_{0i,k} + \overset{(3)}{h}_{0k,i} - \overset{(2)}{h}_{ik,0}\right) + O(\epsilon^4)\,, \tag{4.34}$$

$$\Gamma^i_{jk} = \frac{\epsilon^2}{2}\left(\overset{(2)}{h}_{ij,k} + \overset{(2)}{h}_{ik,j} - \overset{(2)}{h}_{jk,i}\right) + O(\epsilon^3)\,. \tag{4.35}$$

These expressions will be used in the present chapter.

The post-Newtonian expansion of the metric tensor and scalar field introduces a corresponding expansion of the energy-momentum tensor

$$T_{00} = \overset{(0)}{T}_{00} + \epsilon^2 \overset{(2)}{T}_{00} + O(\epsilon^4)\,, \tag{4.36}$$

$$T_{0i} = \epsilon \overset{(1)}{T}_{0i} + \epsilon^3 \overset{(3)}{T}_{0i} + O(\epsilon^5)\,, \tag{4.37}$$

$$T_{ij} = \epsilon^2 \overset{(2)}{T}_{ij} + \epsilon^4 \overset{(4)}{T}_{ij} + O(\epsilon^6)\,, \tag{4.38}$$

where again $\overset{(n)}{T}_{\alpha\beta}$ $(n = 1, 2, 3...)$ denote terms of order ϵ^n. In the first post-Newtonian approximation only terms $\overset{(0)}{T}_{00}$, $\overset{(2)}{T}_{00}$, $\overset{(1)}{T}_{0i}$ and $\overset{(2)}{T}_{ij}$ are needed. They are given by the following equations

$$\overset{(0)}{T}_{00} = \rho^*\,, \tag{4.39}$$

$$\overset{(1)}{T}_{0i} = -\rho^*\left(v^i + N^i\right)\,, \tag{4.40}$$

$$\overset{(2)}{T}_{ij} = \rho^*\left(v^i + N^i\right)\left(v^j + N^j\right) + \pi^{ij}\,, \tag{4.41}$$

$$\overset{(2)}{T}_{00} = \rho^*\left(\frac{v^2}{2} - v^k N_k - \frac{1}{2}N^k N_k + \Pi - N - \frac{H}{2}\right)\,. \tag{4.42}$$

Here one has used the invariant density [Fock, 1964]

$$\rho^* \equiv \sqrt{-g}u^0\rho = \rho + \epsilon^2\rho\left(\frac{1}{2}H + \frac{1}{2}v^2 + \frac{1}{2}N_k N^k + v^k N_k\right)\,, \tag{4.43}$$

that replaces density ρ and is more convenient in calculations because it satisfies the exact Newtonian-like equation of continuity (4.11) which can be recast to [Fock, 1964; Will, 1993]

$$c\rho^*_{,0} + \left(\rho^* v^i\right)_{,i} = 0\,, \tag{4.44}$$

where $v^i = cu^i/u^0$ is the 3-dimensional velocity of matter.

4.2.4
Gauge conditions and residual gauge freedom

In the physics of gauge theories, choosing a gauge denotes a mathematical procedure for removing redundant degrees of freedom in field variables. By definition, a gauge theory maps each physically distinct configuration of the field variables to an equivalence class of the configurations. Any two configurations in the same equivalence class are related by a gauge transformation in a configuration space. Most of the quantitative physical predictions of a gauge theory can only be obtained under a coherent prescription for suppressing these unphysical degrees of freedom. The scalar-tensor theory of gravity belongs to the class of the metric theories of gravity [Will, 1993] admitting the gauge transformations of the metric tensor, which are equivalent to the coordinate transformations. Fixing the gauge starts from the choice of a particular class of equivalence of the coordinates. There are many ways to do it. One of the most convenient choices was proposed by Nutku [Nutku, 1969a,b]. It generalizes the harmonic gauge of general theory of relativity [Misner *et al.*, 1973] that was used in formulation of the IAU 2000 resolutions of relativistic reference frames.

The Nutku gauge conditions imposed on the components of the metric tensor are chosen as follows

$$\partial_\nu \left(\frac{\phi}{\phi_0} \sqrt{-g}\, g^{\mu\nu} \right) = 0\,. \tag{4.45}$$

By making use of the conformal metric tensor one can recast equation (4.45) to the same form as the de Donder (or harmonic) gauge conditions in general relativity [Fock, 1964; Papapetrou, 1951a]

$$(\sqrt{-\tilde{g}}\, \partial_\nu \tilde{g}^{\mu\nu}) = 0\,. \tag{4.46}$$

In what follows, one will use a more convenient form of equation (4.45) written as

$$g^{\mu\nu}\Gamma^\alpha_{\mu\nu} = \partial^\alpha \left(\ln \frac{\phi}{\phi_0} \right)\,, \tag{4.47}$$

so that the Laplace-Beltrami operator (4.3) assumes the form

$$\Box_{\rm g} \equiv g^{\mu\nu} \left(\frac{\partial^2}{\partial x^\mu \partial x^\nu} - \frac{1}{\phi}\frac{\partial \phi}{\partial x^\mu}\frac{\partial}{\partial x^\nu} \right)\,. \tag{4.48}$$

Dependence of this operator on the scalar field is a property of the adopted gauge condition.

Any function $F(x^\alpha)$ satisfying the homogeneous Laplace-Beltrami equation, $\Box_{\rm g} F(x^\alpha) = 0$, is called harmonic. Notice, however, that $\Box_{\rm g} x^\alpha = -\partial^\alpha \ln\phi \neq 0$, so that the coordinates x^α defined by the gauge conditions (4.47) are not harmonic functions but are getting pretty close to them as the scalar field gradually switches off. Therefore, the coordinates singled out by the Nutku condition (4.45) will be called quasi-harmonic. They have many properties similar to the harmonic coordinates in general relativity. The choice of the quasi-harmonic coordinates for constructing theory of the relativistic reference frames in the scalar-tensor theory of gravity is justified by the following three factors: (1) the quasi-harmonic coordinates become harmonic when the scalar field $\phi \to \phi_0$, (2) the harmonic coordinates are used in the resolutions of the IAU 2000 [Soffel *et al.*, 2003] on relativistic reference frames, (3) the condition (4.45) significantly simplifies the field equations and makes it easier to find out their solution. One could use, of course, the harmonic coordinates too, as it has been done, for example, by Klioner and Soffel [Klioner and Soffel, 2000]. They are defined by the condition $g^{\mu\nu}\Gamma^\alpha_{\mu\nu} = 0$ but as one has found [Kopeikin and Vlasov, 2004] the field equations and the spacetime transformations in these coordinates are more complicated in contrast to the quasi-harmonic coordinates defined by the Nutku condition (4.45).

Post-Newtonian expansion of the gauge condition (4.45) yields

$$N_{k,k} = 0\,, \tag{4.49}$$

$$\frac{c}{2}\left(\frac{2\varphi}{\omega+2}+N+H-N_kN^k\right)_{,0} = \frac{N^j}{2}\left(\frac{2\varphi}{\omega+2}+N+H-N_kN^k\right)_{,j} + \left(H_{jk}N^j\right)_{,k} + L_{k,k}\,, \tag{4.50}$$

$$\frac{1}{2}\left(\frac{2\varphi}{\omega+2}-N+H-N_kN^k\right)_{,i} = -L_{i,0}+H_{ik,k}-cN_{i,0}+N^k\left(N_{i,k}-2N_{k,i}\right) \tag{4.51}$$

where the comma with subsequent index denotes a partial derivative with respect to a corresponding coordinate. It is worth noting that in the first PNA the gauge-conditions (4.49)-(4.51) do not restrict the metric tensor component $\overset{(4)}{h}_{00}\equiv L$, which is determined uniquely by the field equations and the gauge fixing at previous steps.

Gauge equations (4.49)-(4.51) define the class of equivalence of the quasi-harmonic coordinates, that is they do not fix the coordinates uniquely. Indeed, if one changes coordinates

$$x^\alpha \longrightarrow w^\alpha = w^\alpha\left(x^\alpha\right)\,, \tag{4.52}$$

keeping the gauge condition (4.47) demands only that the new coordinates w^α must satisfy the homogeneous wave equation

$$g^{\mu\nu}(x^\beta)\frac{\partial^2 w^\alpha}{\partial x^\mu \partial x^\nu} = 0\,, \tag{4.53}$$

which still has an infinite set of non-trivial solutions describing the residual gauge freedom. This residual gauge freedom is used in construction of the relativistic theory of reference frames. It is worth noticing that in the scalar-tensor theory the residual gauge freedom is described by the same equation (4.53) that is also valid in the case of the harmonic coordinates in general relativity. We shall discuss the implementation and meaning of the residual gauge freedom in applicability to the theory of astronomical reference frames in more detail in Chapter 5.

4.2.5 The reduced field equations

The covariant field equations (4.2) and (4.5) are valid in any coordinates. For this reason, their general solution contains four arbitrary functions describing the class of equivalence of the coordinate transformations. This class of equivalence can be significantly narrowed down after imposing the gauge conditions on the field variables - the metric tensor and the scalar field. The resulting equations are called the reduced field equations. The post-Newtonian form of the reduced field equations is obtained after making use of the post-Newtonian expansions, given by equations (4.17)-(4.42) and taking into account the gauge conditions (4.49)-(4.51).

In accordance with equation (4.14) the post-Newtonian approximation of the scalar-tensor theory of gravity with a variable coupling function $\theta(\phi)$ has two constant parameters, ω and ω', characterizing the degree of deviation from Einstein's general relativity. They are related to the standard PPN parameters γ and β [Will, 1993] as follows

$$\gamma = \gamma(\omega) \quad = \quad 1 - \frac{1}{\omega + 2} \,, \tag{4.54}$$

$$\beta = \beta(\omega) \quad = \quad 1 + \frac{\omega'}{(2\omega + 3)(2\omega + 4)^2} \,. \tag{4.55}$$

General relativity is obtained as a limiting case of the scalar-tensor theory when parameters $\gamma = \beta = 1$. In order to obtain this limit parameter ω must go to infinity with ω' growing slower than ω^3. If it were not the case one could get $\lim_{\omega\to\infty} \gamma = 1$ but $\lim_{\omega\to\infty} \beta \neq 1$ which is not a general relativistic limit.

The background value of the scalar field, ϕ_0, and the parameter of coupling ω determine the observed numerical value of the universal gravitational constant

$$G = \frac{2\omega + 4}{2\omega + 3}\, \mathcal{G} \,, \tag{4.56}$$

where $\mathcal{G} \equiv 1/\phi_0$. Had the background value of the scalar field been driven by cosmological evolution, the measured value of the universal gravitational constant would depend on time, and one could expect to detect it experimentally. The best upper limit on time variability of G is imposed by lunar laser ranging (LLR) as $|\dot{G}/G| \leq 0.5 \times 10^{-11}$ yr^{-1} [Schäfer, 2000].

One draws attention of the reader that the book [Will, 1993] uses a different normalization factor in definition of parameter β. Specifically, parameter $\Lambda \equiv \beta - 1$ is introduced as $\Lambda = \omega'(2\omega+3)^{-2}(2\omega+4)^{-1}$ [Will, 1993, equation (5.36)]. The difference from definition (4.55) arises due to the different definitions of the derivative of the coupling function θ with respect to the scalar field, that is $(\omega')_{\rm Will} = \phi_0^{-1}(\omega')_{\text{this book}}$. Taking into account equation (4.56) and the system of units, where $G = 1$, reveals that our definition of parameter β coincides with that given in [Will, 1993]. One also notices that the scalar field perturbation (4.22) is expressed in terms of parameter γ as

$$\overset{(2)}{\zeta} = (1 - \gamma)\varphi \,. \tag{4.57}$$

This excludes factor $\gamma - 1$ in denominator of the right side of the field equation for variable φ (see below).

After making use of the definition of the tensor of energy-momentum, equations (4.39)-(4.42), and that of the PPN parameters, equations (4.54)-(4.56), one obtains the final form of the reduced field equations:

$$\Box\,\varphi = -4\pi G\rho^* \,, \tag{4.58}$$

$$\Box\left\{N + \epsilon^2\left[L + \frac{N^2}{2} + 2(\beta-1)\varphi^2\right]\right\} = \tag{4.59}$$
$$-8\pi G\rho^* + \frac{1}{2}\left(N_{i,k} - N_{k,i}\right)\left(N_{i,k} - N_{k,i}\right) + \epsilon^2\Big\{H_{<ij>}N_{,ij}$$
$$-8\pi G\rho^*\left[\left(\gamma + \frac{1}{2}\right)v^2 + \Pi + \gamma\frac{\pi^{kk}}{\rho^*} - \frac{H}{6} - (2\beta - \gamma - 1)\varphi\right]\Big\}\,,$$

$$\Box\,N_i = 0\,, \tag{4.60}$$

$$\Box\,L_i = 8\pi G\rho^*\left[(1+\gamma)v^i + N^i\right] - 2cN^k N_{i,0k}\,, \tag{4.61}$$

$$\Box\,H_{ij} = -8\pi G\gamma\rho^*\delta_{ij} + N_{k,i}\left(N_{k,j} - N_{j,k}\right) - N_{i,k}\left(N_{j,k} + N_{k,j}\right)\,, \tag{4.62}$$

where $\Box \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu$ is the D'Alembert (wave) operator of the Minkowski spacetime, and $H_{<ij>} \equiv H_{ij} - \delta_{ij}H/3$ is the quadrupole symmetric trace-free (STF) part of the spatial components of the metric tensor perturbation [Thorne, 1980]. We keep in the field equations the terms being quadratic in N_i, but discard the cubic terms and ones being proportional to the products of N_i and the metric perturbations.

Equations (4.58)-(4.62) are valid in any quasi-harmonic coordinates, which are defined up to the residual gauge transformation satisfying to the wave equation (4.53). We shall study the degrees of the residual gauge freedom in constructing the global coordinates for the entire N-body system and the local coordinates for each of the bodies. Global coordinates in the solar system are identified with the barycentric reference frame and the local coordinates are associated with the Earth and planets. The most interesting case of practical applications is the geocentric coordinate frame attached to the Earth.

4.3
Global Astronomical Coordinates

4.3.1
Dynamic and kinematic properties of the global coordinates

When one considers an isolated astronomical system, consisting of N extended bodies – the solar system is a typical example. In the simplest case the N-body system has no hierarchical structure and is comprised of single bodies separated by large distance from each other, like it is shown in Figure 4.1. However, in the most general case the N-body system has more complicated hierarchic structure, which consists of a sequence of sub-systems enclosed to each other like in a Russian nesting doll. Each subsystem is comprised of M_p bodies, where p is a serial number of the sub-system. In its own turn each of the sub-systems can contain several sub-sub-systems, and so on. The solar system with the Sun and planetary sub-systems has this hierarchical structure. In order to describe dynamical behavior of the entire N-body system and reduction of astrometric observations of distant reference sources of light, one needs to introduce a global four-dimensional coordinate system. We denote such global coordinates $x^\alpha = (x^0, x^i)$, where $x^0 = ct$ is time coordinate and $x^i \equiv \boldsymbol{x}$ are spatial coordinates. Adequate description of internal dynamical behavior of the sub-systems of the bodies and/or one of the celestial bodies requires introducing a set of local coordinates attached to each of the sub-systems or to a separate body (see Fig 4.2). The hierarchic structure of the coordinate charts in the sub-system repeats that of the N-body system and is fully compatible with the mathematical concept of the differentiable spacetime manifold [Dubrovin *et al.*, 1984; Misner *et al.*, 1973; Schouten, 1954]. We shall discuss local coordinates later in section 4.5.

Let us define the metric tensor perturbation with respect to the Minkowski metric in accordance with equation (4.16)

$$h_{\alpha\beta}(t, \boldsymbol{x}) \equiv g_{\alpha\beta}(t, \boldsymbol{x}) - \eta_{\alpha\beta} \ . \tag{4.63}$$

We demand that quantities $rh_{\alpha\beta}$ and $r^2\partial_\gamma h_{\alpha\beta}$ are bounded, and that at the null-ray infinity the perturbation decays as follows

$$\lim_{\substack{r\to\infty \\ t+r/c=\mathrm{const.}}} h_{\alpha\beta}(t, \boldsymbol{x}) = 0 \ , \tag{4.64}$$

where $r = |\boldsymbol{x}|$. Additional boundary condition at the null-ray infinity must be imposed on the derivatives of the metric tensor to prevent existence of a non-physical radiative solution associated with the advanced wave potentials. It is defined as [Fock, 1964]

$$\lim_{\substack{r\to\infty \\ t+r/c=\mathrm{const.}}} \left[\left(rh_{\alpha\beta}\right)_{,r} + \left(rh_{\alpha\beta}\right)_{,0}\right] = 0 \ . \tag{4.65}$$

equation (4.65) is known as a "no-incoming-radiation" boundary condition [Damour, 1983; Fock, 1964]. In the case of an isolated astronomical system this condition singles out a causal solution of the gravitational field equation depending on the retarded

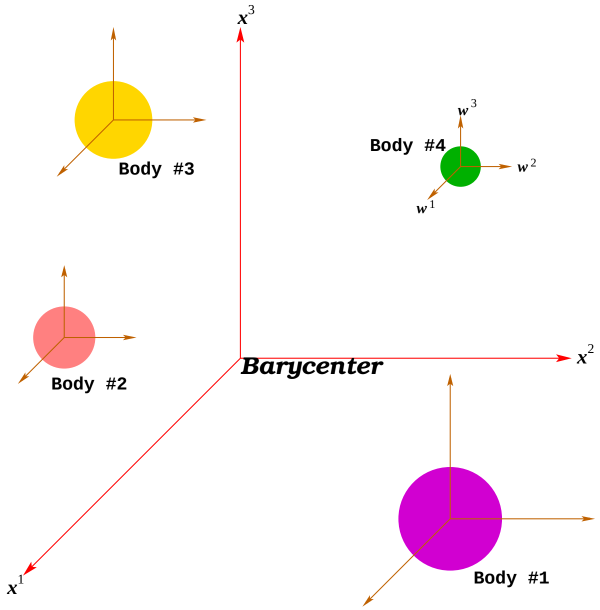

Figure 4.1 The picture illustrates an astronomical N-body system and coordinate charts associated with it. Global coordinates, $x^\alpha = (ct, x^i)$, cover the entire space, have origin at the barycenter of the system, and match the inertial coordinates of the Minkowski spacetime at infinity. Each body is affiliated with its own local coordinate chart, $w^\alpha = (cu, w^i)$, having origin at the center of mass of the body. Local coordinates are not asymptotically Minkowskian at large distance from the body and do not cover the entire space.

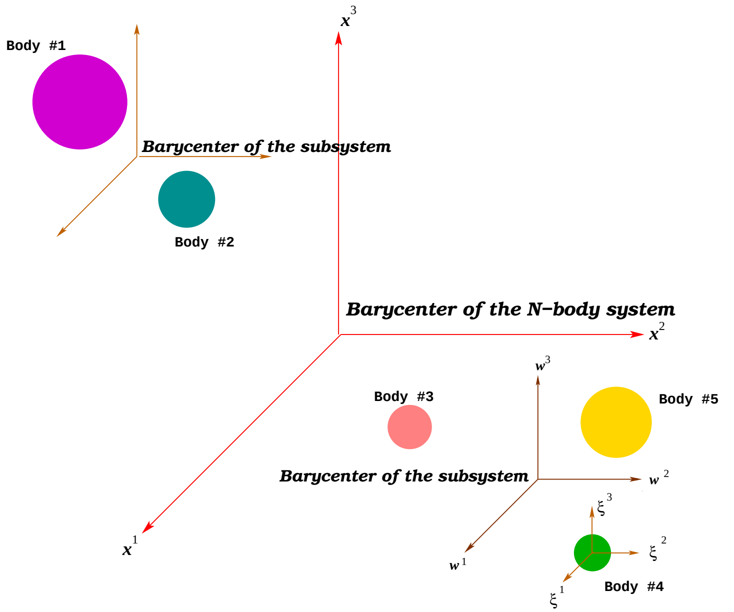

Figure 4.2 The picture illustrates a hierarchy of coordinate charts existing in N-body astronomical system (solar system) consisting of several sub-systems: M_1, M_2,...,M_p. One global coordinate chart, $x^\alpha = (ct, x^i)$, covers the entire space and has its origin at the barycenter of the N-body system. Each sub-system consists of several gravitationally-bounded bodies (a planet and its moons) and has its own local coordinate chart, $w^\alpha = (cu, w^i)$, having its origin at the center of mass of the sub-system. At the same time, each body from the sub-system possesses its own local coordinate chart, $\xi^\alpha = (cs, \xi^i)$. The hierarchy can have as many levels as the number of the consecutive sub-systems enclosed to each other.

time $t - r/c$ only. Similar boundary conditions are imposed on the perturbation of the scalar field defined in equation (4.14),

$$\lim_{\substack{r\to\infty \\ t+r/c=\text{const.}}} \zeta(t, \boldsymbol{x}) = 0\,, \tag{4.66}$$

$$\lim_{\substack{r\to\infty \\ t+r/c=\text{const.}}} \left[(r\zeta)_{,r} + (r\zeta)_{,0}\right] = 0\,. \tag{4.67}$$

In principle, the boundary conditions (4.65) and (4.67) are not explicitly required in the first post-Newtonian approximation for solving the field equations (4.58)-(4.62) because the gravitational potentials in this approximation are time-symmetric. However, they are convenient for doing calculations and are physically motivated by the principle of causality applied to the gravitational field. Therefore, in what follows the radiative boundary conditions (4.65) and (4.67) will be used for giving precise definition of the multipole moments of the gravitational field of the isolated astronomical system.

The global coordinates x^α cover the entire spacetime and they set up a primary basis for construction of the theory of relativistic reference frames in the N-body system [Kopejkin, 1988a]. We shall stipulate that the origin of the global coordinates coincides with the barycenter of the N-body system at any instant of time. This condition can be satisfied after choosing a suitable definition of the post-Newtonian dipole moment $\mathbb{D}^i$ of the N-body system and equating its numerical value to zero along with its first time derivative (see section 4.4.5). This can be always done in general relativity in low orders of the post-Newtonian approximation scheme if one neglects the octuple and higher-order multipole gravitational radiation [Blanchet, 1998]. In the scalar-tensor theory of gravity one has also to take into account gravitational wave emission in the form of scalar field modes [Damour and Esposito-Farese, 1992] but it does not affect the first post-Newtonian approximation. There are alternative theories of gravity, which violate the third Newton's law of the gravitational interaction between celestial bodies. The dipole moment $\mathbb{D}^i$ of N-body system is not conserved in such theories even in the first post-Newtonian approximation [Will, 1993]. We do not consider such theories which are physically implausible.

We shall also assume that the spatial axes of the global coordinates do not rotate in space either kinematically or dynamically [Brumberg and Kopejkin, 1989b]. Notice that the angular velocities of dynamic and kinematic rotations of a reference frame are not distinguished in classic celestial mechanics. However, they have different meaning already in the first post-Newtonian approximation due to the presence of the relativistic geodetic precession caused by the orbital motion of the body around the center of mass of the N-body system. Spatial axes of a coordinate chart are called kinematically non-rotating if their orientation is fixed with respect to the Minkowski coordinates defined at the infinite past and infinite distance from the solar system called the past null infinity [Misner *et al.*, 1973]. Such kinematically non-rotating coordinate system can be built on the stellar sky by making use of quasars as reference objects with accuracy better than 100 μas (see [Johnston *et al.*, 2000] and references therein). Quasars are uniformly distributed all over the sky and have negligibly small parallax and proper motion that is defined as its transverse motion in the plane of the sky being orthogonal to the line of sight of observer located at the barycenter of the solar system. Thus, kinematically non-rotating character of coordinates can be determined only through the experimental analysis of the global properties of the spacetime manifold including its topology. This consideration reveals that the theory of reference frames in N-body system based on the assumption that the spacetime is asymptotically-flat may be corrupted by the influence of some cosmological effects. Hence, a more appropriate approach to the reference frames taking into account that the background spacetime is the Friedman-Robertson-Walker universe is to be developed. One has started a constructive work in this direction in papers [Kopeikin *et al.*, 2001; Ramírez and Kopeikin, 2002]. More work is required to match observational predictions of the post-Newtonian approximations with the cosmological model.

Dynamically non-rotating coordinate chart is defined by the condition that equations of motion of test particles moving with respect to these coordinates do not have any term that might be interpreted as either the Coriolis or the centripetal force [Brumberg and Kopejkin, 1989b]. This definition relies only upon the local properties of spacetime and does not require the astronomical observations of distant celestial objects like stars or quasars. Dynamical definition of the spatially non-rotating coordinates is used in construction of modern ephemerides of the solar system bodies which are based on radar and laser ranging measurements to planets and Moon as well as on their radio and optical observations (see [Johnston *et al.*, 2000; Kovalevsky *et al.*, 1989; Seidelmann and Urban, 2010] and references therein). Because of the assumption that the N-body system under consideration is isolated, one can theoretically postulate that the global coordinates do not rotate in any sense.

4.3.2
The metric tensor and scalar field in the global coordinates

The metric tensor $g_{\alpha\beta}(t, \boldsymbol{x})$ is obtained by solving the field equations (4.58)-(4.62) after imposing the boundary conditions (4.64)-(4.67). In what follows, a trivial solution of the homogeneous equation (4.60) for spacetime component g_{0i} of the metric tensor is chosen, $N_i = 0$. This is because N_i describes rotation (angular velocity) of the spatial axes of the coordinate chart but the previous section postulates that the global coordinates are not rotating. Solution of the other field equations are obtained by applying the Green function method for elliptic equations. They are given in the following form

$$\varphi(t, \boldsymbol{x}) = U(t, \boldsymbol{x}) \,, \tag{4.68}$$

$$N(t, \boldsymbol{x}) = 2\,U(t, \boldsymbol{x}) \,, \tag{4.69}$$

$$L(t, \boldsymbol{x}) = 2\Phi(t, \boldsymbol{x}) - 2\beta U^2(t, \boldsymbol{x}) - c^2\chi_{,00}(t, \boldsymbol{x}) \,, \tag{4.70}$$

$$L_i(t, \boldsymbol{x}) = -2(1+\gamma)\,U_i(t, \boldsymbol{x}) \,, \tag{4.71}$$

$$H_{ij}(t, \boldsymbol{x}) = 2\gamma\,\delta_{ij}\,U(t, \boldsymbol{x}) \,, \tag{4.72}$$

where

$$\Phi(t,\boldsymbol{x}) \equiv (\gamma + \frac{1}{2})\Phi_1(t,\boldsymbol{x}) + (1-2\beta)\Phi_2(t,\boldsymbol{x}) + \Phi_3(t,\boldsymbol{x}) + \gamma\Phi_4(t,\boldsymbol{x}) \,, \tag{4.73}$$

and the gravitational potentials U, U^i, χ, and Φ_k $(k = 1, ..., 4)$ can be represented as linear combinations of the gravitational potentials of each body, that is

$$U = \sum_A U^{(A)} \,, \qquad U_i = \sum_A U_i^{(A)} \,, \qquad \Phi_k = \sum_A \Phi_k^{(A)} \,, \qquad \chi = \sum_A \chi^{(A)} \,. \tag{4.74}$$

Herein, the gravitational potentials of body A are defined as integrals taken over the volume of this body

$$U^{(A)}(t,\boldsymbol{x}) = G\mathbf{I}_{-1}^{(A)}\{\rho^*\} \,, \tag{4.75}$$

$$U_i^{(A)}(t,\boldsymbol{x}) = G\mathbf{I}_{-1}^{(A)}\left\{\rho^* v^i\right\} \,, \tag{4.76}$$

$$\chi^{(A)}(t,\boldsymbol{x}) = -G\mathbf{I}_{1}^{(A)}\{\rho^*\} \,, \tag{4.77}$$

$$\Phi_1^{(A)}(t,\boldsymbol{x}) = G\mathbf{I}_{-1}^{(A)}\left\{\rho^* v^2\right\} \,, \tag{4.78}$$

$$\Phi_2^{(A)}(t,\boldsymbol{x}) = G\mathbf{I}_{-1}^{(A)}\{\rho^* U\} \tag{4.79}$$

$$\Phi_3^{(A)}(t,\boldsymbol{x}) = G\mathbf{I}_{-1}^{(A)}\{\rho^* \Pi\} \,, \tag{4.80}$$

$$\Phi_4^{(A)}(t,\boldsymbol{x}) = G\mathbf{I}_{-1}^{(A)}\left\{\pi^{kk}\right\} \,, \tag{4.81}$$

where notation $\mathbf{I}_n^{(A)}\{f\}$ $(n = 1, 2, 3...)$ is used to define the volume integral

$$\mathbf{I}_n^{(A)}\{f\}\,(t,\boldsymbol{x}) = \int\limits_{V_A} f(t,\boldsymbol{x}')|\boldsymbol{x}-\boldsymbol{x}'|^n \, d^3x' \,, \tag{4.82}$$

with n being an integer, and V_A - the volume of integration. Potential χ is determined as a particular solution of the inhomogeneous equation

$$\nabla^2\chi = -2U \,, \tag{4.83}$$

with the right side defined in a whole space. Nevertheless, its solution given by equation (4.77) has a compact support concentrated inside the volumes of the bodies [Fock, 1964]. It is worthwhile to emphasize that all the integrals defining the metric tensor in the global coordinates are taken over the hypersurface of constant coordinate time t. Spacetime transformations can change the hypersurface of the coordinate time, hence transforming the volume of integration in the integrals. This important issue should be taken into account properly. It will be further discussed in section 5.2.3.

4.4
Gravitational Multipoles in the Global Coordinates

4.4.1
General description of multipole moments

In what follows, a set of certain parameters, describing properties of gravitational and scalar fields and depending on the integral characteristics of the N-body system, will be indispensable. These integral parameters are called multipole moments. In the Newtonian approximation they are uniquely defined as coefficients in the Taylor expansion of the Newtonian gravitational potential U in powers of $1/r$, where $r = |\boldsymbol{x}|$ is the radial coordinate distance from the origin of a coordinate system to a field point. The Newtonian multipole moments are functions of time in the most general astronomical situations because the body can oscillate, wobble and change its internal structure due to radiation or other processes. However, very often one assumes that mass is conserved and the center of mass of the N-body system is located at the origin of the coordinate chart under consideration. Provided that these assumptions are satisfied, the monopole and dipole multipole moments remain constant.

General relativistic multipolar expansion of gravitational field is in many aspects similar to the Newtonian multipolar decomposition. However, due to the non-linearity and tensorial character of gravitational interaction in general theory of relativity and the scalar-tensor theory of gravity, the definition of relativistic multipole moments gets much more complicated in contrast to the Newtonian theory. Furthermore, the gauge freedom existing in the general theory of relativity clearly indicates that any multipolar decomposition of gravitational field will be coordinate-dependent making the definition of the multipoles affected by the choice of the gauge. Hence, a certain care is required for unambiguous physical interpretation of various relativistic effects associated with the multipoles For example, section 6.3 demonstrates how the appropriate choice of coordinates allows us to eliminate a number of coordinate-dependent terms in equations of motion of spherically-symmetric bodies depending on the "quadrupoles" defined in the global coordinate system.

The papers [Blanchet, 1987; Blanchet and Damour, 1986; Thorne, 1980] proved that in general relativity the multipolar expansion of the gravitational field of an isolated gravitating system is characterized by only two independent sets of multipoles called as mass-type and current-type moments. In particular, Thorne [Thorne, 1980] had systematized and significantly perfected works of previous researchers [Brown, 1896; Epstein and Wagoner, 1975; Pirani, 1965; Sachs, 1961; Wagoner, 1979], and defined the two sets of the post-Newtonian multipole moments as follows [Thorne, 1980, equations (5.32a) and (5.32b)]

$$\mathcal{I}^L_{\rm Thorne} = \int \left(\tau_{00} x^L + A^{l0} r^2 x^{<L-2}\tau^{a_{l-1}a_l>} + B^{l0} x^{j<L-1}\tau^{a_l>j} + C^{l0} x^L \tau^{jj}\right) d^3x , \tag{4.84}$$

$$\mathcal{S}^L_{\rm Thorne} = \varepsilon^{pq<a_l} \int \left(x^{L-1>p}\tau^{0q} + E^{l0} r^2 x^{L-2}\partial_t \tau^{a_{l-1}>j} x^j + F^{l0} x^{L-1>kp}\partial_t \tau^{kq}\right) d^3x , \tag{4.85}$$

where the numerical coefficients

$$A^{l0} = \frac{l(l-1)(l+9)}{2(l+1)(2l+3)}\,, \qquad B^{l0} = -\frac{6l(l-1)}{(l+1)(2l+3)}\,, \tag{4.86}$$

$$C^{l0} = \frac{2l(l-1)}{(l+1)(2l+3)}\,, \qquad E^{l0} = \frac{(l-1)(l+4)}{2(l+2)(2l+3)}\,, \tag{4.87}$$

$$F^{l0} = -\frac{l-1}{(l+2)(2l+3)}\,, \tag{4.88}$$

and the multipolar integer-valued index l runs from 0 to infinity. In these expressions

$$\tau^{\alpha\beta} = \left(1 + 4\epsilon^2 U\right) T^{\alpha\beta} + \frac{c^4}{16\pi G} N^{\alpha\beta}\,, \tag{4.89}$$

is the effective energy-momentum tensor of Landau-Lifshitz [Landau and Lifshitz, 1975] evaluated at the post-Newtonian order in the harmonic gauge with

$$N^{00} = -\frac{14}{c^4} U_{,p} U_{,p}\,, \tag{4.90}$$

$$N^{0i} = \frac{4}{c^5}\left[4U_{,p}\left(U^p{}_{,i} - U^i{}_{,p}\right) - 3U_{,i} U^p{}_{,p}\right]\,, \tag{4.91}$$

$$N^{ij} = \frac{2}{c^4}\left(2U_{,i} U_{,j} - \delta_{ij} U_{,p} U_{,p}\right)\,, \tag{4.92}$$

and U, U^i are gravitational potentials of the isolated astronomical system defined in equations (4.74).

Thorne [Thorne, 1980] had systematically neglected all surface terms in the solution of the boundary-value problem of gravitational field equations in order to derive the multipoles (4.84), (4.85). However, the part $N^{\alpha\beta}$ of the effective energy-momentum tensor $\tau^{\alpha\beta}$ falls off as $1/r^4$ as a coordinate distance r from the isolated system grows. For this reason, the multipole moments defined in equations (4.84), (4.85) are formally divergent. This divergency is, in fact, artificial and can be completely eliminated if one makes use of a more rigorous mathematical technique developed by Blanchet and Damour [1989] for the mass-type multipole moments and by Damour and Iyer [1991b] for the spin-type multipoles. This technique is based on the theory of distributions (generalized functions) [Gelfand, 1964] and consists in the replacement of the energy-momentum pseudotensor $\tau^{\alpha\beta}$ with the effective localized source $\tau_c^{\alpha\beta}$, which has a compact support inside the region occupied by matter of the isolated system [Blanchet and Damour, 1989; Damour and Iyer, 1991b]. Blanchet and Damour [1989] proved that formal integration by parts of the integrands of Thorne's multipole moments (4.84), (4.85) with subsequent discarding all surface terms recovers the multipole moments derived by Blanchet and Damour by making use of the compact-support effective source $\tau_c^{\alpha\beta}$. It effectively demonstrates that Thorne's post-Newtonian multipole moments are physically (and computationally) meaningful provided that one operates only with compact-support terms in the integrands of equations (4.84), (4.85) after their rearrangement and integration by parts of the non-linear source of gravitational field $N^{\alpha\beta}$ given by equations (4.90)-(4.92). This transformation was done by Blanchet and Damour [1989] who extracted the non-divergent core of Thorne's multipole moments. We shall use their result in this book.

In the scalar-tensor theory of gravity the multipolar series gets more involved because of the presence of the scalar field. This brings about an additional set of scalar multipole moments which are intimately related with the multipolar decomposition of the scalar field outside of the gravitating system. We emphasize that definition of the multipole moments in the scalar-tensor theory of gravity depends not only on the choice of the gauge conditions but also on the freedom of conformal transformation of the metric tensor as was pointed out by Damour and Esposito-Farese [1992] who also derived (in the global coordinates) the set of multipole moments for an isolated astronomical system in the framework of two-parametric scalar-tensor theory of gravity.

We shall study the problem of the multipolar decomposition of gravitational and scalar fields both of the whole N-body system and of each body comprising the system in the framework of the scalar-tensor theory of gravity. We shall follow the line of study outlined and elucidated in works [Blanchet and Damour, 1989; Damour and Esposito-Farese, 1992; Thorne, 1980; Will, 1993]. The multipole moments under discussion will include the sets of *active*, *conformal* and *scalar* multipole moments. These three sets are constrained by one identity (see equation (4.120)). Hence, only two of the sets are algebraically (and physically) independent. The multipole moments will be defined in various reference frames associated both with an isolated astronomical system and with a separated body (or a sub-system of the bodies) comprising the isolated system.

4.4.2
Active multipole moments

Let us introduce the metric tensor potentials

$$V = \frac{1}{2}\left\{N + \epsilon^2\left[L + \frac{N^2}{2} + 2(\beta - 1)\varphi^2\right]\right\}, \tag{4.93}$$

$$V^i = -\frac{L_i}{2(1+\gamma)}, \tag{4.94}$$

which enter $g_{00}(t,\boldsymbol{x})$ and $g_{0i}(t,\boldsymbol{x})$ components of the metric tensor respectively. Furthermore, throughout this chapter $N_i = 0$, and the spatial metric component H_{ij} is chosen to be isotropic, that is its symmetric and trace-free (STF) part $H_{<ij>} = 0$ [Kopeikin and Vlasov, 2004]. Then, the field equations for these potentials follow from equations (4.59), (4.61) and read

$$\Box V = -4\pi G\sigma, \tag{4.95}$$

$$\Box V^i = -4\pi G\sigma^i, \tag{4.96}$$

where one has introduced the *active* mass density

$$\sigma = \rho^*\left\{1 + \epsilon^2\left[(\gamma + \frac{1}{2})v^2 + \Pi + \gamma\frac{\pi^{kk}}{\rho^*} - \frac{H}{6} - (2\beta - \gamma - 1)\varphi\right]\right\}, \tag{4.97}$$

and the *active* current mass density

$$\sigma^i = \rho^* v^i \,. \tag{4.98}$$

It is worthwhile to observe that in the global coordinates one has $H = H_{ii} = 6\gamma U(t, \boldsymbol{x})$ and $\varphi(t, \boldsymbol{x}) = U(t, \boldsymbol{x})$. Hence, the expression (4.97) for the *active* mass density in these coordinates is simplified and reduced to

$$\sigma = \rho^* \left\{ 1 + \epsilon^2 \left[(\gamma + \frac{1}{2}) v^2 + \Pi + \gamma \frac{\pi^{kk}}{\rho^*} - (2\beta - 1) U \right] \right\} . \tag{4.99}$$

Solutions of equations (4.95) and (4.96) are retarded wave potentials [Landau and Lifshitz, 1975] satisfying the boundary conditions (4.64)-(4.65) and determined up to the solution of a homogeneous wave equation. Taking into account that potentials V and V_i are, in fact, components of the metric tensor, solutions of equations (4.95) and (4.96) can be written down as

$$V(t, \boldsymbol{x}) = G \int_{\mathcal{D}} \frac{\sigma(t - \epsilon|\boldsymbol{x} - \boldsymbol{x}'|, \boldsymbol{x}')}{|\boldsymbol{x} - \boldsymbol{x}'|} d^3x' + c^2 \partial_0 \xi^0 \,, \tag{4.100}$$

$$V^i(t, \boldsymbol{x}) = G \int_{\mathcal{D}} \frac{\sigma^i(t - \epsilon|\boldsymbol{x} - \boldsymbol{x}'|, \boldsymbol{x}')}{|\boldsymbol{x} - \boldsymbol{x}'|} d^3x' + \frac{c^3}{2(1+\gamma)} \left[\partial_0 \xi^i - \partial_i \xi^0 \right] , \tag{4.101}$$

where $\mathcal{D}$ designates a domain of integration going over the entire space, and the gauge functions ξ^0 and ξ^i are solutions of the homogeneous wave equation. One notices that because the densities σ and σ^i vanish outside the bodies the integration in equations (4.100) and (4.101) is performed only over the volume occupied by matter of the bodies.

We take a special choice of the gauge functions as proposed in [Blanchet and Damour, 1989]

$$\xi^0 = -2(1+\gamma)\epsilon^3 G \sum_{l=0}^{\infty} \frac{(-1)^l}{(l+1)!} \frac{2l+1}{2l+3} \left[\frac{1}{r} \int_{\mathcal{D}} \sigma^k(t - \epsilon r, \boldsymbol{x}') \, x'^{<kL>} \, d^3x' \right]_{,L} \tag{4.102}$$

$$\xi^i = 0 \,. \tag{4.103}$$

Such gauge transformation preserves the gauge condition (4.45) and does not change the post-Newtonian form of the scalar multipole moments, which will be discussed in the next section. Then, one can show that potentials V and V^i can be expanded outside of the N-body system in a multipolar series as follows [Blanchet and Damour, 1989]

$$V(t, \boldsymbol{x}) = G \sum_{l=0}^{\infty} \frac{(-1)^l}{l!} \left[\frac{I_{<L>}(t - \epsilon r)}{r} \right]_{,L} , \tag{4.104}$$

$$V^i(t, \boldsymbol{x}) = G \sum_{l=0}^{\infty} \frac{(-1)^l}{(l+1)!} \left\{ \left[\frac{\dot{I}_{<iL>}(t - \epsilon r)}{r} \right]_{,L} - \frac{l+1}{l+2} \varepsilon_{ipq} \left[\frac{S_{<pL>}(t - \epsilon r)}{r} \right]_{,qL} \right\} , \tag{4.105}$$

where the dot above function denotes differentiation with respect to time t. equations (4.104) and (4.105) define the *active* mass multipoles, I_L, and the spin multipoles, S_L, which can be expressed in the first PNA in terms of integrals over the N-body system's matter as follows

$$I_{<L>}(t) = \int_{\mathcal{D}} \sigma(t,\boldsymbol{x}')x'^{<L>}\, d^3x' + \frac{\epsilon^2}{2(2l+3)}\left[\frac{d^2}{dt^2}\int_{\mathcal{D}} \sigma(t,\boldsymbol{x}')x'^{<L>}x'^2\, d^3x' - 4(1+\gamma)\frac{2l+1}{l+1}\frac{d}{dt}\int_{\mathcal{D}} \sigma^i(t,\boldsymbol{x}')x'^{<iL>}\, d^3x'\right], \quad (4.106)$$

$$S_{<L>}(t) = \int_{\mathcal{D}} \varepsilon^{pq<a_l}\hat{x}'^{L-1>p}\sigma^q(t,\boldsymbol{x}')\, d^3x' \,. \quad (4.107)$$

As one can see the mass and spin multipole moments of the scalar-tensor theory define the tensor part of gravitational field of the metric tensor outside N-body system as perfect as in general relativity [Blanchet and Damour, 1989; Thorne, 1980]. When $\beta = \gamma = 1$ these multipole moments coincide with their general relativistic expressions [Blanchet and Damour, 1989]. However, in order to complete the multipole decomposition of the gravitational field in the scalar-tensor theory one needs to obtain a multipolar expansion of the scalar field as well.

4.4.3 Scalar multipole moments

In order to find out the post-Newtonian definition of the multipole moments of the scalar field one again uses the technique developed in works [Blanchet and Damour, 1989; Thorne, 1980]. One takes equation (4.5) and write it down with the post-Newtonian accuracy by making use of a new (scalar) potential

$$\bar{V} = c^2\zeta + \frac{\epsilon^2}{2}\Big[\eta - (\gamma-1)(\gamma-2)\Big]\varphi^2 \,. \quad (4.108)$$

Then, equation (4.5) assumes the form

$$\Box\, \bar{V} = -4\pi G\bar{\sigma}\,, \quad (4.109)$$

where notation $\eta \equiv 4\beta - \gamma - 3$ stands for the Nordtvedt parameter [Will, 1993] and the *scalar* mass density $\bar{\sigma}$ is defined as

$$\bar{\sigma} = (1-\gamma)\rho^*\left\{1 - \epsilon^2\left[\frac{1}{2}v^2 - \Pi + \frac{\pi^{kk}}{\rho^*} + \frac{H}{6}\right]\right\} - \epsilon^2\Big[\eta + \gamma(\gamma-1)\Big]\rho^*\varphi \,. \quad (4.110)$$

One can easily check out that in the global coordinates, where $H = 6\gamma U(t,\boldsymbol{x})$ and $\varphi(t,\boldsymbol{x}) = U(t,\boldsymbol{x})$, the *scalar* mass density is simplified and is given by

$$\bar{\sigma} = (1-\gamma)\rho^*\left[1 - \epsilon^2\left(\frac{1}{2}v^2 - \Pi + \frac{\pi^{kk}}{\rho^*}\right)\right] - \epsilon^2\eta\rho^* U \,. \quad (4.111)$$

Solution of equation (4.109) is the retarded scalar potential

$$\bar{V}(t,\boldsymbol{x}) = G\int_{\mathcal{D}} \frac{\bar{\sigma}(t-\epsilon|\boldsymbol{x}-\boldsymbol{x}'|,\,\boldsymbol{x}')}{|\boldsymbol{x}-\boldsymbol{x}'|}\,d^3x'\,. \tag{4.112}$$

Multipolar decomposition of the potential (4.112) has the same form as in equation (4.104) with the *scalar* mass multipole moments defined as integrals over a volume occupied by matter of the N-body system

$$\bar{I}_{<L>}(t) = \int_{\mathcal{D}} \bar{\sigma}(t,\boldsymbol{x}')x'^{<L>}\,d^3x' + \frac{\epsilon^2}{2(2l+3)}\frac{d^2}{dt^2}\int_{\mathcal{D}} \bar{\sigma}(t,\boldsymbol{x}')x'^{<L>}x'^2\,d^3x'\,. \tag{4.113}$$

We conclude that in the scalar-tensor theory of gravity the multipolar decomposition of gravitational field requires introduction of three sets of multipole moments - the *active* mass moments I_L, the *scalar* mass moments $\bar{I}_L$, and the spin moments S_L. Neither the *active* nor the *scalar* mass multipole moments alone lead to the laws of conservation of energy, linear momentum, etc. of an isolated system; only their linear combination does. This linear combination of the multipole moments can be derived after making conformal transformation of the metric tensor, solving the Einstein equations for the conformal metric, and finding its multipolar decomposition in the similar way as it was done in section 4.4.2.

4.4.4 Conformal multipole moments

Let us now define the *conformal* metric potential

$$\tilde{V} = \frac{1}{1+\gamma}\left[\tilde{N} + \epsilon^2\left(\tilde{L} + \frac{\tilde{N}^2}{2}\right)\right]\,. \tag{4.114}$$

The conformal field equations (4.7) in the quasi-harmonic gauge of Nutku (4.46) yield

$$\Box\,\tilde{V} \;=\; -4\pi G\tilde{\sigma}\,, \tag{4.115}$$

where one has introduced a *conformal* mass density

$$\tilde{\sigma} = \rho^*\left\{1 + \epsilon^2\left[\frac{3}{2}v^2 + \Pi + \frac{\pi^{kk}}{\rho^*} - \frac{H}{6} - (1-\gamma)\varphi\right]\right\}\,, \tag{4.116}$$

which has been calculated directly from equation (4.7) by making use of the definition of the conformal metric (4.6) and the post-Newtonian expansions of corresponding quantities described in section 4.2.3. Remembering that in the global coordinates $H = 6\gamma U(t,\boldsymbol{x})$ and $\varphi(t,\boldsymbol{x}) = U(t,\boldsymbol{x})$ one can simplify expression for the *conformal* mass density, which assumes the form

$$\tilde{\sigma} = \rho^*\left[1 + \epsilon^2\left(\frac{3}{2}v^2 + \Pi + \frac{\pi^{kk}}{\rho^*} - U\right)\right]\,. \tag{4.117}$$

This equation coincides precisely with the post-Newtonian mass density as defined in general relativity [Damour and Esposito-Farese, 1992; Fock, 1964; Will, 1993]. The *conformal* current density $\tilde{\sigma}^i$ is defined in the approximation under consideration by the same equation as equation (4.98), that is $\tilde{\sigma}^i = \sigma^i$. The field equation for the conformal vector potential $\tilde{V}^i$ has the form (4.96), therefore in this approximation $\tilde{V}^i = V^i$.

Solution of equation (4.115) gives the retarded *conformal* potential

$$\tilde{V}(t,\boldsymbol{x}) = G\int_{\mathcal{D}} \frac{\tilde{\sigma}(t-\epsilon|\boldsymbol{x}-\boldsymbol{x}'|,\,\boldsymbol{x}')}{|\boldsymbol{x}-\boldsymbol{x}'|}\,d^3x'\,. \tag{4.118}$$

Multipolar expansion of *conformal* potentials $\tilde{V}$ and $\tilde{V}^i$ is done in the same way as it was done previously in section 4.4.2. It turns out that the conformal spin moments coincide with the active spin moments (4.107), and the expansion of the potential $\tilde{V}(t,\boldsymbol{x})$ acquires the same form as that given in equation (4.104) but with all active multipole moments replaced with the conformal multipoles, $\tilde{I}_{<L>}$, defined as follows

$$\begin{aligned}\tilde{I}_{<L>}(t) \quad = \quad & \int_{\mathcal{D}} \tilde{\sigma}(t,\boldsymbol{x}')x'^{<L>}d^3x' + \frac{\epsilon^2}{2(2l+3)}\left[\frac{d^2}{dt^2}\int_{\mathcal{D}} \tilde{\sigma}(t,\boldsymbol{x}')x'^{<L>}x'^2 d^3x'\right. \\ & \left. -8\,\frac{(2l+1)}{l+1}\frac{d}{dt}\int_{\mathcal{D}} \sigma^i(t,\boldsymbol{x}')x'^{<iL>}d^3x'\right]\,.\end{aligned} \tag{4.119}$$

These *conformal* mass multipole moments coincide exactly with those introduced in general relativity by Blanchet and Damour [1989] who also proved [Blanchet and Damour, 1989, appendix A] that their definition is compatible after formal discarding all surface integrals with the mass multipole moments introduced originally in the first post-Newtonian approximation in general relativity by Thorne [Thorne, 1980].

There is a simple algebraic relationship between the three mass multipole moments, I_L, $\bar{I}_L$ and $\tilde{I}_L$ in the global frame. Specifically, one has [Kopeikin and Vlasov, 2004]

$$I_{<L>} = \frac{1+\gamma}{2}\tilde{I}_{<L>} + \frac{1}{2}\bar{I}_{<L>}\,. \tag{4.120}$$

We shall show later in section 4.5.3 that relationship (4.120) between the multipole moments obtained in the global coordinates $(t,\boldsymbol{x})$ for the case of an isolated astronomical N-body system preserves its form in the local coordinates for each gravitating body (a sub-system of the bodies) as well.

4.4.5
Post-Newtonian conservation laws

It is crucial for the following analysis to discuss the laws of conservation for an isolated astronomical N-body system in the framework of the scalar-tensor theory of gravity. These laws will allow us to formulate the post-Newtonian definitions of mass, the center of mass, the linear and the angular momenta for the isolated system which are used in derivation of equations of motion of the bodies comprising the system. In order to derive the laws of conservation one employs a general relativistic approach developed in [Landau and Lifshitz, 1975] and extended to the Brans-Dicke theory by Nutku [Nutku, 1969a].

To this end, it is convenient to recast the field equations (4.2) to the form

$$\Theta^{\mu\nu} \equiv (-g)\frac{\phi}{\phi_0}\left[c^2T^{\mu\nu}+t^{\mu\nu}\right] = \frac{c^4}{16\pi\phi_0}\partial_{\alpha\beta}\left[(-g)\phi^2(g^{\mu\nu}g^{\alpha\beta}-g^{\mu\alpha}g^{\nu\beta})\right], \tag{4.121}$$

where $t^{\mu\nu}$ is an analog of the Landau-Lifshitz pseudo-tensor of the gravitational field [Landau and Lifshitz, 1975] in the scalar-tensor theory of gravity. This pseudotensor is defined by the equation [Nutku, 1969a]

$$t^{\mu\nu} = \frac{c^4}{16\pi}\frac{\phi^3}{\phi_0^2}\tilde{\tau}^{\mu\nu}_{LL} + \frac{c^4}{16\pi}\frac{2\theta(\phi)+3}{\phi}\left(\partial^\mu\phi\partial^\nu\phi - \frac{1}{2}g^{\mu\nu}\partial_\lambda\phi\partial^\lambda\phi\right), \tag{4.122}$$

where $\tilde{\tau}^{\mu\nu}_{LL}$ is the (standard) Landau-Lifshitz pseudotensor [Landau and Lifshitz, 1975] expressed in terms of the conformal metric $\tilde{g}_{\alpha\beta}$ and its derivatives.

The conservation laws are now obtained from equation (4.121)

$$\partial_\nu\Theta^{\mu\nu} \equiv \partial_\nu\left[(-g)\frac{\phi}{\phi_0}(c^2T^{\mu\nu}+t^{\mu\nu})\right] = 0\,. \tag{4.123}$$

They are a direct consequence of anti-symmetry of the right side of equation (4.121) with respect to the upper indices ν and α. In what follows, one concentrates on the laws of conservation in the first post-Newtonian approximation only. Hence, one neglects the energy, linear and angular momenta taken away from the system by gravitational waves (see [Damour and Esposito-Farese, 1992] where the role of the gravitational waves is discussed). Under these circumstances the mass $\mathbb{M}$, the linear momentum $\mathbb{P}^i$, and spin $\mathbb{S}^i$ of the isolated gravitating N-body system are conserved and defined as

$$\mathbb{M} = \epsilon^2\int_{\mathcal{D}}\Theta^{00}\,d^3x\,, \tag{4.124}$$

$$\mathbb{P}^i = \epsilon\int_{\mathcal{D}}\Theta^{0i}\,d^3x\,, \tag{4.125}$$

$$\mathbb{S}^i = \epsilon\int_{\mathcal{D}}\varepsilon^i{}_{jk}w^j\Theta^{0k}\,d^3x\,, \tag{4.126}$$

where the integration is performed over the whole space. Let us remark that the integrals (4.124)-(4.126) are finite (not divergent) since in the first PNA the integrands Θ^{00} and Θ^{0i} are of $O(r^{-4})$ for large r and decay faster than the volume of integration grows. Moreover, in this approximation the domain of integration can be reduced from the entire space to the volume of the bodies comprising the system. This is because the integrand in equations (4.124)-(4.126) can be integrated by parts with all the surface terms going to zero at infinity. After all, what remains are the integrals shown below in equations (4.129)-(4.131) where the functions under the integrals have only the compact support inside the bodies.

Taking into account the asymptotic behavior of Θ^{00} and the differential law of conservation (4.123), one can prove that the linear momentum $\mathbb{P}^i$ can be represented as the time derivative of the function

$$\mathbb{D}^i = \epsilon^2 \int_{\mathcal{D}} \Theta^{00} x^i \, d^3x \,, \tag{4.127}$$

which is interpreted as the integral of the center of mass of the N-body system. Hence, integrating one more time one gets

$$\mathbb{D}^i(t) = \mathbb{P}^i \, t + \mathbb{K}^i \,, \tag{4.128}$$

where $\mathbb{K}^i$ is a constant vector defining a constant displacement of the barycenter of the N-body system from the origin of the global coordinate frame. One can chose $\mathbb{K}^i = 0$ and $\mathbb{P}^i = 0$. In such case $\mathbb{D}^i = 0$, and the center of mass of the N-body system will always coincide with the origin of the global reference frame. Such global reference frame is called the Barycentric Celestial Reference Frame (BCRF) and it is used in the IAU 2000 resolutions [Soffel *et al.*, 2003] discussed in chapter 9 of this book. BCRF is a primary tool in description of ephemerides of the solar system bodies, navigation of spacecrafts in deep space and reduction of astronomical observations of various types.

Direct calculations of the pseudotensor (4.122) with subsequent comparison with the conformal multipole moments (4.119) reveal that for the isolated system the post-Newtonian conserved quantities are

$$\mathbb{M} \equiv \tilde{I} = \int_{\mathcal{D}} \rho^* \left[1 + \epsilon^2 \left(\Pi + \frac{v^2}{2} - \frac{U}{2} \right) \right] d^3x + O(\epsilon^4) \,, \tag{4.129}$$

$$\mathbb{D}^i \equiv \tilde{I}^i = \int_{\mathcal{D}} \rho^* x^i \left[1 + \epsilon^2 \left(\Pi + \frac{v^2}{2} - \frac{U}{2} \right) \right] d^3x + O(\epsilon^4) \,, \tag{4.130}$$

$$\mathbb{P}^i = \dot{\mathbb{D}}^i = \int_{\mathcal{D}} \left\{ \rho^* v^i \left[1 + \epsilon^2 \left(\Pi + \frac{v^2}{2} - \frac{U}{2} \right) \right] + \epsilon^2 \pi^{ik} v^k - \frac{\epsilon^2}{2} \rho^* W^i \right\} d^3x + O(\epsilon^4) \tag{4.131}$$

where by definition

$$W^i(t, \boldsymbol{x}) = G \int_{\mathcal{D}} \frac{\rho^*(t, \boldsymbol{x}') \boldsymbol{v}' \cdot (\boldsymbol{x} - \boldsymbol{x}')(x^i - x'^i)}{|\boldsymbol{x} - \boldsymbol{x}'|^3} \, d^3x' \,, \tag{4.132}$$

and the integration is performed over the hypersurface of constant global coordinate time t. It is evident from equations (4.129) and (4.130) that it is the *conformal* moments, $\tilde{I}$ and $\tilde{I}^i$, which define the conserved mass $\mathbb{M}$ and linear momentum $\mathbb{D}^i$ of the N-body system. The *active* monopole and dipole moments defined by equation (4.106) for $l = 0, 1$ are not consistent with the laws of conservation and, hence, can not serve to define the conserved quantities in the scalar-tensor theory of gravity. We fix position of the center of mass (barycenter) of the N-body system in the global coordinates by equating the *conformal* dipole moment of the system to zero, that is $\tilde{I}^i = 0$, for any instant of time.

Now one can proceed further and construct local coordinates in the vicinity of a gravitating body or a sub-system of bodies which are members of the entire N-body system. For concreteness and for the sake of simplicity the focus will be on the construction of the local coordinate chart around one body (Earth, planet, etc.). Extension of this formalism to the planetary sub-system like the Earth-Moon is given in paper by [Xie and Kopeikin, 2010].

4.5
Local Astronomical Coordinates

4.5.1
Dynamic and kinematic properties of the local coordinates

A local coordinate chart (local coordinates) is constructed in the vicinity of the worldline of each body comprising the N-body system. Thus, in principle, N local coordinate charts w^α must be introduced in addition to one global coordinate system x^α (see Figure 4.1). In the case of N-body system, which is divided into sub-systems of bodies, the number of the local coordinates increases in accordance with the underlying hierarchic structure of the N-body system (see Figure 4.2). An example is the Earth-Moon sub-system that is gravitationally bounded and requires introduction of the local coordinates associated with the center of mass of the sub-system [Kopeikin *et al.*, 2008; Xie and Kopeikin, 2010] besides the local coordinates attached to the Earth and Moon. The principles of construction of the local coordinates are the same for any weakly gravitating body. A sub-system of bodies can be also viewed as a single body with a time-dependent internal structure governed by the internal gravitational field of the sub-system. For this reason, it is sufficient to work out the description of one local coordinate chart, $w^\alpha = (cu, \boldsymbol{w})$, around a single self-gravitating body as the other local coordinate charts have a similar mathematical structure [Kopejkin, 1991b]. For practical applications in the solar system the most important local coordinates are associated with the Earth and they are called geocentric coordinates. They play an important role in formulation of the IAU 2000 resolutions on the reference frames in the solar system [Soffel *et al.*, 2003]. The important physical difference between the global and local coordinates is that the local coordinates are not asymptotically-Minkowski far away from the body because the gravitational field of the body under consideration must smoothly match with the tidal gravitational field of external bodies in a buffer region of space between the bodies. It means that the metric tensor in the local coordinates diverges as the coordinate distance from the body grows [Ni and Zimmermann, 1978b], [Misner *et al.*, 1973, see pages 172, 327]. Further discussion will be focused primarily on the construction of the local geocentric reference frame.

Mathematical development of the theory requires to assume that each body consists of viscoelastic matter which admits continuous distribution of mass density, anisotropic stresses, and internal velocity field. If one had “turned off" gravitational field of all external bodies (Moon, Sun, planets) the gravitational field of the body under consideration (Earth) would be described by a set of the (internal) multipole moments defined by equations given in previous section. However, one can not neglect gravitational field of the external bodies if one wants to take into account the Newtonian [Melchior, 1983] and relativistic effects associated with tides [Kopejkin, 1988a; Mashhoon, 1975; Xu *et al.*, 2003]. The tidal deformation of the body is comprehensively large, for example, at the latest stage of coalescence of neutron stars in binary systems emitting gravitational radiation and, no doubt, must be taken into account in calculation of templates of gravitational waves emitted by such systems.

Terrestrial tides are not so dramatically large. However, modern astronomical observations are so accurate that the tidal gravitational field influences many aspects of the rotational and orbital motions of the Earth [Melchior, 1983], which require a very accurate theoretical modeling. Gravitational potential of the tidal force is represented in the local coordinates as a Taylor series with respect to spatial coordinates with time-dependent coefficients, which are called the external (tidal) multipole moments [Kopejkin, 1988a; Suen, 1986; Thorne and Hartle, 1985]. This series usually starts in the Newtonian approximation from the second order (quadratic) polynomial term because the monopole and dipole external multipole moments are not physically associated with the tidal force, which origin in relativity is explained in terms of the Riemann curvature of spacetime that is given by the second derivatives from the metric tensor (see sections 1.2.6 and 3.7). In general relativity this monopole-dipole effacing property of the external gravitational field is extended from the Newtonian gravity to the post-Newtonian approximations as a consequence of the Einstein principle of equivalence (EEP) [Will, 1993, see §2.3]. In particular, EEP suggests that it is always possible to chose the local coordinates in such a way that all first derivatives of the metric tensor (i.e., the Christoffel symbols) will vanish along a geodesic worldline of a freely-falling test particle [Ni and Zimmermann, 1978b]. This is equivalent to making a suitable coordinate transformation on the spacetime manifold from the global to local frame [Dubrovin *et al.*, 1984; Schouten, 1954]. In general relativity EEP is also valid for a self-gravitating body moving in external gravitational field of other bodies. The original proof was given for black holes in papers [D'Eath, 1975a,b] and for extended bodies it was elaborated in [Breuer and Rudolph, 1982; Brumberg and Kopejkin, 1989a; Damour *et al.*, 1991, 1992; Kopeikin, 1985; Kopejkin, 1991b].

As contrasted with general relativity, the scalar-tensor theory of gravity has a scalar component of the gravitational field, which can not be eliminated by a coordinate transformation to the local frame of the body being in a free fall. This is because the scalar field does not change its numerical value under pointwise coordinate transformations and can not be eliminated if it has a non-zero value at least in one coordinate chart on spacetime manifold. It means that the scalar field can not obey the principle of equivalence and the gravitational field in the scalar-tensor theory can not be reduced in the local coordinates to the pure tidal field. In particular, this was the reason why Einstein had rejected a scalar theory of gravity based exclusively on a scalar potential (for more detail see discussion in [Misner *et al.*, 1973; Will, 1993]). One of the consequences of the violation of the principle of equivalence in scalar-tensor theory of gravity is the appearance of, the so-called, Nordtvedt effect [Nordtvedt, 1968a,b] that is also discussed in [Will, 1993, see §8.1]. The Nordtvedt effect can be observed only for massive, self-gravitating bodies. This is because the scalar field is not observed directly but is organically incorporated to the metric tensor, which can be reduced to the Minkowski metric at the origin of the local coordinates as long as the body's gravitational field is negligibly small.

Standard parameterized post-Newtonian formalism pretends to describe any possible deviation from general relativity phenomenologically by introducing a metric tensor parameterized with 10 parameters [Nordtvedt, 1970; Nordtvedt and Will, 1972; Will, 1993]. However, this leads to a major problem in experimental gravity as it remains unclear how to incorporate a local coordinate system to the standard PPN formalism in a self-consistent way. Theoretical principles demand that one needs to know the nature of the fundamental fields (scalar, vector, tensor, spinor, etc.) entering the metric tensor because these fields have different behavior under coordinate transformations. To construct a local coordinate system, solution of the field equations for the fundamental fields must be found directly in the local coordinate system. Then, this solution must be matched to the solution of the same field equations in the global coordinates. Finally, the transformation law of the fields must be used along with the transformation law of the metric tensor in order to find out the relativistic spacetime transformation between the global and local coordinate systems. The standard PPN formalism misses all these steps, thus, making correct interpretation of more subtle gravitational experiments difficult or even impossible.

We demand that the origin of the local coordinates coincides with the body's center of mass at any instant of time. This requires a precise definition of the center of mass of each body with respect to its local coordinates. However, when one takes into account the post-Newtonian corrections, the concept of the body's center of mass becomes ambiguous because it can be chosen in several different ways depending on what kind of definition of the internal dipole moment of the body in the multipolar expansion of the local metric tensor is chosen. One has proven by straightforward calculations [Kopeikin and Vlasov, 2004] that it is the *conformal* dipole moment given in equation (4.130), which gives a physically meaningful definition of the body's center of mass. The reason is that this is the only definition which allows us to derive equations of translational motion of the body where all self-accelerated terms violating the Newton's third law are suppressed and vanish identically. This property of the *conformal* dipole moment is closely related to its conservation for an isolated system of N bodies as demonstrated in section 4.4.5.

In general, the body (Earth) as a part of the N-body system is not isolated and interacts gravitationally with other bodies of the Solar system (Moon, Sun, etc.). For this reason, the second and higher order time derivatives of the conformal dipole moment of the body are not equal to zero. It means that there is a local force exerted on the body by external gravitational field, which prevents its linear momentum (the first time derivative of the body's dipole moment) to remain constant. Nevertheless, if one chooses a local frame moving with a certain acceleration all time derivatives of the body's dipole moment can be made equal to zero. Such a local frame does not move along a geodesic worldline [Kopejkin, 1988a] bringing about a non-zero value of the external dipole moment in the multipolar expansion of the homogeneous solution of the gravitational field equations (see section 6.1.4). A proper choice of the body's center of mass and the local acceleration of the body-centered frame allow us to eliminate the coordinate-dependent terms in the equations of motion of the body and clarify the origin of the Nordtvedt effect for extended bodies.

We admit that the local coordinates can be, in general, dynamically rotating. It means that translational equations of motion of a test particle written down in the local coordinates can include the Coriolis and centrifugal forces. If one excludes the dynamic rotation of the local coordinates, their spatial axes will slowly precess in the kinematic sense with respect to the spatial axes of the global coordinates [Brumberg *et al.*, 1998; Soffel *et al.*, 2003]. This effect is called the geodetic precession, and it obeys the law of parallel transport of vectors on a curved spacetime manifold [Misner *et al.*, 1973]. Nowadays, the IAU recommends that one uses a kinematically non-rotating geocentric system with the spatial axes that are anchored to distant quasars being considered as reference points of the international celestial reference system (ICRS) (see [Mignard, 2000] for more detail). The metric tensor of the kinematically non-rotating geocentric coordinates has an external dipole moment in $\hat{g}_{0i}(u, \boldsymbol{w})$ component of the geocentric metric tensor describing the dynamic rotation of the spatial axes of the geocentric coordinates. This dipolar term would be equal to zero if the geocentric coordinates were chosen to be dynamically non-rotating. The angular velocity of the dynamical rotation is equal to that of the geodetic precession of the Earth's angular momentum vector as it moves around the Sun. Numerical value of the geodetic precession is fixed by the corresponding IAU resolution in agreement with observations of the reference quasars. At this step of development of the theory, the angular velocity of the dynamical rotation is not specified in order to keep the formalism as general as possible.

4.5.2
The metric tensor and scalar field in the local coordinates

We denote the local (for example, geocentric) coordinates by $w^\alpha = (w^0, w^i) = (cu, w^i)$ where u stands for the local coordinate time. All quantities related to the body around which the local coordinate frame is constructed, will be labeled by subindex B standing for "body". One is looking for the solution of the field equations (4.58)-(4.62) inside a world tube containing the worldline of the body's center of mass and spreading up to the nearest external body, so that the only source of matter inside the region covered by the local frame is the matter of the central body. Thus, the right side of equations (4.58)-(4.62) contains the energy-momentum tensor of the body's matter only. Spatial domain of applicability of the local coordinates can be extended after finding the spacetime transformation from the local to global coordinates as it will be explained in Section 5.1.3.

Solution of the differential equations (4.58)-(4.62) is a linear combination of general solution of the homogeneous equation and a particular solution of the inhomogeneous equation. Solution for a scalar field in the local coordinates is written as

$$\hat{\varphi}(u, \boldsymbol{w}) = \hat{\varphi}^{(B)}(u, \boldsymbol{w}) + \hat{\varphi}^{(E)}(u, \boldsymbol{w}) + \hat{\varphi}^{(C)}(u, \boldsymbol{w}) , \tag{4.133}$$

whereas the metric tensor, $\hat{g}_{\mu\nu}(u, \boldsymbol{w}) = \eta_{\mu\nu} + \hat{h}_{\mu\nu}(u, \boldsymbol{w})$, is given in the form

$$\hat{h}_{\mu\nu}(u, \boldsymbol{w}) = \hat{h}^{(B)}_{\mu\nu}(u, \boldsymbol{w}) + \hat{h}^{(E)}_{\mu\nu}(u, \boldsymbol{w}) + \hat{h}^{(C)}_{\mu\nu}(u, \boldsymbol{w}) , \tag{4.134}$$

where the terms with sub-index B refer to the body in question and describe the internal solution of the inhomogeneous equations, terms with sub-index E refer to the external bodies (Moon, Sun, etc.) and describe the external solution of the homogeneous equations, and terms with sub-index C, which stands for coupling, arise because of the non-linearity of the gravity field equations for the metric tensor. One notices that in the first post-Newtonian approximation the coupling terms appear only in $\hat{g}_{00}(u, \boldsymbol{w})$ component of the metric tensor and are not required in calculation of the scalar field.

We do not impose any other specific limitations on the structure of the metric tensor in the local coordinates. All information about its structure can be obtained from the solution of the field equations (4.58)-(4.62). One draws attention of the reader that a hat over all quantities is a label referred to the local coordinates w^α. It is used because the functional dependence of one and the same quantity looks different in different coordinates. For example, for any function $F(x)$ and coordinate transformation $x = x(w)$ one has $F(x) = F[x(w)] \equiv \hat{F}(w)$. One can not use the same notation for this function in all coordinates because the numerical value of $F(w)$ differs from $F(x)$ [Dubrovin *et al.*, 1984; Schouten, 1954].

4.5.2.1 ***The Scalar Field: Internal and External Solutions***

equation (4.58) gives internal, $\hat{\varphi}^{(B)}(u, \boldsymbol{w})$, and external, $\hat{\varphi}^{(E)}(u, \boldsymbol{w})$, solutions for the scalar field in the following form

$$\hat{\varphi}^{(B)}(u, \boldsymbol{w}) = \hat{U}^{(B)}(u, \boldsymbol{w}) , \tag{4.135}$$

$$\hat{\varphi}^{(E)}(u, \boldsymbol{w}) = \sum_{l=0}^{\infty} \frac{1}{l!} P_L w^L . \tag{4.136}$$

Here $\hat{U}^{(B)}(u, \boldsymbol{w})$ is the Newtonian potential defined in equation (4.141), $P_L \equiv P_L(u)$ are external STF multipole moments [Thorne, 1980] in the multipolar decomposition of the scalar field generated by the bodies which are external with respect to the central body. These external moments are defined on the worldline of the body in question and are functions of the local time u only. The internal solution $\hat{\varphi}^{(B)}(u, \boldsymbol{w})$ describes the scalar field, which is generated by the body only.

4.5.2.2 ***The Metric Tensor: Internal Solution***

The boundary conditions imposed on the internal solution for the metric tensor are identical with those given in equations (4.64)-(4.65). For this reason the internal solution for the metric tensor looks formally similar to that obtained in the global coordinates in section 4.3.2, where all quantities must be referred now only to the body under consideration. One obtains

$$\hat{N}^{(B)}(u, \boldsymbol{w}) = 2\hat{U}^{(B)}(u, \boldsymbol{w}) , \tag{4.137}$$

$$\hat{L}^{(B)}(u, \boldsymbol{w}) = 2\hat{\Phi}^{(B)}(u, \boldsymbol{w}) - 2\beta\left[\hat{U}^{(B)}(u, \boldsymbol{w})\right]^2 - c^2\hat{\chi}^{(B)}_{,00}(u, \boldsymbol{w}) , \tag{4.138}$$

$$\hat{L}^{(B)}_i(u, \boldsymbol{w}) = -2(1+\gamma)\hat{U}^{(B)}_i(u, \boldsymbol{w}) , \tag{4.139}$$

$$\hat{H}^{(B)}_{ij}(u, \boldsymbol{w}) = 2\gamma\delta_{ij}\hat{U}^{(B)}(u, \boldsymbol{w}) , \tag{4.140}$$

where all gravitational potentials of the central body are taken over the volume of the body's matter defined as a cross-section of the body's world tube with the hypersurface of a constant local coordinate time u. Specifically, one has

$$\hat{U}^{(B)}(u,\boldsymbol{w}) = G\hat{\mathbf{I}}^{(B)}_{-1}\{\rho^*\}\ , \tag{4.141}$$

$$\hat{U}^{(B)}_i(u,\boldsymbol{w}) = G\hat{\mathbf{I}}^{(B)}_{-1}\left\{\rho^* v^i\right\}\ , \tag{4.142}$$

$$\begin{aligned}\hat{\Phi}^{(B)}(u,\boldsymbol{w}) &= (\gamma+\frac{1}{2})\hat{\Phi}^{(B)}_1(u,\boldsymbol{w}) + (1-2\beta)\hat{\Phi}^{(B)}_2(u,\boldsymbol{w})\\ &\quad +\hat{\Phi}^{(B)}_3(u,\boldsymbol{w}) + \gamma\hat{\Phi}^{(B)}_4(u,\boldsymbol{w}),\end{aligned} \tag{4.143}$$

where

$$\hat{\Phi}^{(B)}_1(u,\boldsymbol{w}) = G\hat{\mathbf{I}}^{(B)}_{-1}\left\{\rho^* v^2\right\}\ , \tag{4.144}$$

$$\hat{\Phi}^{(B)}_2(u,\boldsymbol{w}) = G\hat{\mathbf{I}}^{(B)}_{-1}\left\{\rho^* \hat{U}^{(B)}\right\}\ , \tag{4.145}$$

$$\hat{\Phi}^{(B)}_3(u,\boldsymbol{w}) = G\hat{\mathbf{I}}^{(B)}_{-1}\{\rho^*\Pi\}\ , \tag{4.146}$$

$$\hat{\Phi}^{(B)}_4(u,\boldsymbol{w}) = G\hat{\mathbf{I}}^{(B)}_{-1}\left\{\pi^{kk}\right\}\ , \tag{4.147}$$

$$\hat{\chi}^{(B)}(u,\boldsymbol{w}) = -G\hat{\mathbf{I}}^{(B)}_{1}\{\rho^*\}\ , \tag{4.148}$$

the symbol $v^i = dw^i/du$ is the velocity of the body's matter with respect to the origin of the local coordinates, and one has introduced a special notation

$$\hat{\mathbf{I}}^{(B)}_n\{f\}(u,\boldsymbol{w}) = \int\limits_{V_B} f(u,\boldsymbol{w}')|\boldsymbol{w}-\boldsymbol{w}'|^n\, d^3w'\ , \tag{4.149}$$

for integrals over the body's volume. We emphasize once again that the integrand of $\hat{\mathbf{I}}^{(B)}_n\{f\}(u,\boldsymbol{w})$ is a function, which is taken over the intersection of the hypersurface of the constant time u with the world tube of the body. Changing the hypersurface of integration alters the value of the integral.

The local metric given by equations (4.137)-(4.140) must obey the gauge condition (4.45) which yields

$$\frac{1}{c}\frac{\partial\hat{U}^{(B)}}{\partial u} + \frac{\partial\hat{U}^{(B)}_k}{\partial w^k} = O(\epsilon^2)\,. \tag{4.150}$$

This is the only gauge condition, which can be imposed on the local metric in the first post-Newtonian approximation. We note that equation (4.150) is satisfied due to the validity of the equation of continuity (4.44).

4.5.2.3 ***The Metric Tensor: External Solution***

Solution of the homogeneous field equations for the metric tensor given in this section extends the multipolar formalism for description of vacuum gravitational fields developed in [Suen, 1986; Thorne and Hartle, 1985]. Brief introduction to this formalism is given in Appendix A. Boundary conditions imposed on the external solution must ensure its convergency on the worldline of the origin of the local coordinates, where $\boldsymbol{w} = 0$. However, the external solution for the metric tensor diverges as the radial distance $r = |\boldsymbol{w}|$ from the origin of the local coordinates grows. This is because the gravitational field of the external bodies has a tidal character and does not asymptotically vanish in the local coordinates for large r [Suen, 1986; Thorne and Hartle, 1985].

Explicit form of the external solution for the linearized metric tensor perturbation in the local coordinates is given by

$$\hat{N}^{(E)}(u,\boldsymbol{w}) = 2\sum_{l=0}^{\infty}\frac{1}{l!}Q_L w^L + \Omega^2 w^2 - \Omega^p\Omega^q w^p w^q \,, \tag{4.151}$$

$$\hat{N}_i^{(E)}(u,\boldsymbol{w}) = \mathcal{V}_i + \varepsilon_{ipq}\Omega_p w^q \,, \tag{4.152}$$

$$\hat{L}_i^{(E)}(u,\boldsymbol{w}) = \sum_{l=1}^{\infty}\frac{1}{l!}\varepsilon_{ipq}C_{pL-1}w^{<qL-1>} + \sum_{l=0}^{\infty}\frac{1}{l!}Z_{iL}w^L + \sum_{l=0}^{\infty}\frac{1}{l!}S_L w^{<iL>} \,, \tag{4.153}$$

$$\begin{aligned}\hat{H}_{ij}^{(E)}(u,\boldsymbol{w}) = {} & 2\delta_{ij}\sum_{l=0}^{\infty}\frac{1}{l!}Y_L w^L + \sum_{l=0}^{\infty}\frac{1}{l!}B_L w^{<ijL>} + \frac{1}{3}\left(\delta_{ij}\Omega^2 - \Omega^i\Omega^j\right)w^2 \\ & + \sum_{l=1}^{\infty}\frac{1}{l!}\left(D_{iL-1}w^{<jL-1>} + \varepsilon_{ipq}E_{pL-1}w^{<jqL-1>}\right)^{\mathrm{Sym}(ij)} \\ & + \sum_{l=2}^{\infty}\frac{1}{l!}\left(F_{ijL-2}w^{L-2} + \varepsilon_{pq(i}G_{j)pL-2}w^{<qL-2>}\right),\end{aligned} \tag{4.154}$$

where Ω_i is the angular velocity of kinematic rotation of the local frame with respect to the global coordinates, $\mathcal{V}^i$ is the linear velocity of the local frame with respect to the local frame moving along geodesic worldline (see below), and symbol "Sym(ij)" around indices denote symmetry with respect to the indices that is equivalent to the round brackets around indices, for instance, $[T_{ijL}]^{\mathrm{Sym}(ij)} \equiv T_{(ij)L} = (1/2)[T_{ijL} + T_{jiL}]$. Equations (4.151)-(4.154) keep rotational terms of the linear, $O(\Omega)$, and quadratic, (Ω^2), orders, which are relevant in discussion of the Newtonian motion of test particles.

It is worth to notice that the external solution for the metric tensor in the local coordinates contains the monopole terms: Q and Y. We accept that time intervals and spacial distances are measured in the international system of units (SI) irrespectively of the choice of coordinates. Then, term Q defines the rate of the coordinate time u at the origin of the local frame, and Y defines the scale of the spatial coordinates. Both these terms are equated to zero in most theoretical works but one prefers to keep them in one's equations for generality. This is because the IAU resolutions [Soffel *et al.*, 2003] implicitly introduce the non-zero values of Q and Y through the scaling transformations of time and space coordinates [Brumberg and Kopeikin, 1990; Klioner *et al.*, 2009a]. Currently, IAU adopts functions $Q = Y \equiv L_C = 1.48082686741 \times 10^{-8} \pm (2 \times 10^{-17}$ [Fukushima, 2002; Irwin and Fukushima, 1999]. Papers [Brumberg and Kopeikin, 1990; Klioner *et al.*, 2009a; Kopeikin, 2010a] provide more detailed theoretical review of the problem of the relativistic time scales in the solar system.

In order to understand physical meaning of various components of the external solution for the metric tensor in the local coordinates it is instructive to write down the Newtonian equation of motion of a test particle falling freely in the gravitational field, defined only by the external metric. This equation is a geodesic worldline, which is defined by the law of parallel transport of the particle's four velocity. After calculation of the Christoffel symbols (4.24)-(4.29) the particle's acceleration reads

$$\frac{d^2w^i}{du^2} = Q_i - \dot{\mathcal{V}}_i - 2\varepsilon_{ijk}\Omega^j v^k - \varepsilon_{ijk}\dot{\Omega}^j w^k + \left(\Omega^2\delta_{ij} - \Omega_i\Omega_j\right) w^j + Q_{ij}w^j + \sum_{l=2}^{\infty}\frac{1}{l!}Q_{iL}w^L + O\left(\epsilon^2\right) , \tag{4.155}$$

where $v^i \equiv dw^i/du$, and one has neglected the post-Newtonian corrections. Linear combination of the first two terms in the right side of this equation, $Q_i - \dot{\mathcal{V}}_i$, describe kinematic acceleration of the particle with respect to the coordinate chart having the origin moving along another geodesic. The third term, $2\varepsilon_{ijk}\Omega^j v^k$, in the right side of equation (4.155) is the famous Coriolis acceleration [Landau and Lifshitz, 1969] caused by the motion of the particle and rotation of the spatial axes of the local frame with the angular velocity Ω^i. The forth term, $\varepsilon_{ijk}\dot{\Omega}^j w^k$, in the right side of equation (4.155) is an acceleration due to the non-uniform rotation of the local frame. The fifth term, $\left(\Omega^2\delta_{ij} - \Omega_i\Omega_j\right) w^j$, describes a centrifugal acceleration of the particle. The sixth term, $Q_{<ij>}w^j$, is a quadrupole tidal acceleration due to the presence of the external gravitational field of the other bodies. Last term in the right side of equation (4.155) is the tidal acceleration due to the higher order multipoles of the external gravitational field of the other bodies. It is interesting to note that the centrifugal and the quadrupole tidal accelerations have similar structure. The difference, however, is that the matrix of the centrifugal acceleration, $\Omega^2\delta_{ij} - \Omega_i\Omega_j$, is not trace-free as oppose to the tidal matrix, Q_{ij}. However, the trace-free part of $\Omega^2\delta_{ij} - \Omega_i\Omega_j$ can be singled out and absorbed to the definition of Q_{ij}, which indicates that the definition of Q_{ij} should be always specified along with the specification of the matrix of the centrifugal acceleration.

It is convenient to construct the external part of the metric tensor in such a way that the space and time axes become orthogonal at the origin of the local coordinates. This can be achieved, if one chooses function $\mathcal{V}^i = 0$. This condition also allows us to give a unique interpretation of the dipole term Q_i as being equal to the inertial force per unit mass exerted on the free falling particle due to the accelerated motion of the origin of the local frame under consideration with respect to a geodesic worldline. In other words, the metric tensor with $\mathcal{V}^i = 0$ and $Q_i \neq 0$ specifies a local coordinate chart such that its origin moves with acceleration Q_i with respect to a geodesic worldline defined on the background spacetime, which is determined exclusively by the external part of the metric tensor. We also notice that the dipole term Z_i in equation (4.153) is just a post-Newtonian correction to $\mathcal{V}^i$ and would also destroy orthogonality of spatial axes of the local frame at its origin. Thus, in addition to the condition, $\mathcal{V}^i = 0$, one demands $Z^i = 0$. Therefore, the acceleration of the origin of the local coordinates (the test particle coordinates $w^i = 0$) is now defined by a single function Q_i in equation (4.151).

A set of eleven external STF multipole moments P_L, Q_L, C_L, Z_L, S_L, Y_L, B_L, D_L, E_L, F_L, G_L (the sign of angular brackets $<>$ is omitted for simplicity, i.e. $P_L \equiv P_{<L>}$, etc.) is defined on the worldline of the origin of the local coordinates so that these multipoles are functions of the local coordinate time u only. Furthermore, the external multipole moments are symmetric and trace-free (STF) geometric objects with respect to any of two indices, which are transformed as tensors with respect to a linear transformation of spatial coordinates. In what follows, the angular velocity of rotation of the local frame, Ω^i, is assumed to be so small that the metric tensor component $\epsilon\hat{N}_i^{(E)}$ is comparable with the post-Newtonian correction $\epsilon^3\hat{L}_i^{(E)}$. For this reason, all terms which are quadratic with respect to $\hat{N}_i^{(E)}$ or are products of $\hat{N}_i^{(E)}$ with either $\hat{N}^{(E)}$, or $\hat{L}_i^{(E)}$ or $\hat{H}_{ij}^{(E)}$, will be neglected. Only linear with respect to $\hat{N}_i^{(E)}$ terms and their first derivatives will be retained in the calculations which follow.

Imposing the gauge conditions (4.49)-(4.51) on the metric tensor given by equations (4.151)-(4.154) reveals that only 7 from 11 external multipole moments are algebraically independent. More specifically, the gauge condition (4.51) leads to the following relationship between the moments

$$D_L = \frac{2l(2l-1)}{2l+1}\Big[Y_L + (1-\gamma)P_L - Q_L\Big], \qquad (l \geq 1). \tag{4.156}$$

Three other relationships are obtained after accounting for equation (4.156) in the gauge condition (4.50), which yields

$$S_L = \dot{Y}_L + (1-\gamma)\frac{2l^2+l+1}{(l+1)(2l+3)}\dot{P}_L - \frac{2l^2-3l-1}{(l+1)(2l+3)}\dot{Q}_L, \qquad (l \geq 0), \tag{4.157}$$

and

$$E_i = \frac{2}{5}\dot{\Omega}_i, \tag{4.158}$$

$$E_L = 0, \qquad (l \geq 2) \tag{4.159}$$

$$B_L = 0, \qquad (l \geq 0). \tag{4.160}$$

equations (4.156)-(4.160) allow us to eliminate the external multipole moments B_L, E_L, D_L, S_L from the local metric so that the spacetime and space-space components of the external metric tensor assume a simpler form

$$\hat{L}_i^{(E)}(u,\boldsymbol{w}) = \sum_{l=1}^{\infty}\frac{1}{l!}\varepsilon_{ipq}C_{pL-1}w^{<qL-1>} + \sum_{l=1}^{\infty}\frac{1}{l!}Z_{iL}w^L + \sum_{l=0}^{\infty}\frac{1}{l!}\left[\dot{Y}_L + (1-\gamma)\frac{2l^2+l+1}{(l+1)(2l+3)}\dot{P}_L - \frac{2l^2-3l-1}{(l+1)(2l+3)}\dot{Q}_L\right]w^{<iL>}\,, \tag{4.161}$$

$$\hat{H}_{ij}^{(E)}(u,\boldsymbol{w}) = \frac{1}{5}\left(\varepsilon_{ipq}\dot{\Omega}^p w^{<jq>} + \varepsilon_{jpq}\dot{\Omega}^p w^{<iq>}\right) + 2\delta_{ij}\sum_{l=0}^{\infty}\frac{1}{l!}Y_L w^L + 2\sum_{l=0}^{\infty}\frac{2l+1}{(2l+3)\,l!}\left[\left(Y_{iL} + (1-\gamma)P_{iL} - Q_{iL}\right)w^{<jL>}\right]^{\mathrm{Sym}(ij)} + \sum_{l=0}^{\infty}\frac{1}{(l+2)!}F_{ijL}w^L + \sum_{l=0}^{\infty}\frac{1}{(l+2)!}\varepsilon_{pq(i}G_{j)pL}w^{<qL>}\,. \tag{4.162}$$

Remaining seven multipole moments P_L, Q_L, Z_L, C_L, Y_L, F_L, G_L and the angular velocity of rotation, Ω_i, can not be constrained by imposing the gauge conditions. However, the residual gauge freedom described by the differential equation (4.53) allows us to find out further limitations on the remaining 7 sets of the multipole moments, which are explicitly shown in the right side of equations (4.161) and (4.162). Examination of the residual gauge degrees of freedom makes it clear which multipole moments are, in fact, physically meaningful, that is can not be excluded by infinitesimal coordinate transformations.

In order to eliminate the gauge-dependent external multipoles, which do not carry out gravitational degrees of freedom, one uses the property of the gauge-invariance of the linearized Riemann tensor [Misner *et al.*, 1973, see Box 18.2] under infinitesimally small coordinate transformations $\hat{w}^\alpha = w^\alpha + \xi^\alpha(x^\beta)$

$$R_{\alpha\beta\gamma\delta}(\hat{w}) = R_{\alpha\beta\gamma\delta}(w)\,. \tag{4.163}$$

equation (4.163) must be understood as invariance of the functional form of the Riemann tensor with respect to infinitesimally small gauge transformation. Computing all components of the Riemann tensor for the external metric tensor (4.151), (4.161) and (4.162), one finds that the external part of the Riemann tensor depends only on three sets of the external multipole moments P_L, Q_L, C_L while all other multipole moments vanish. More specifically, the components of the Riemann tensor computed by making use of the external metric tensor only, are [1)]

$$
\begin{aligned}
R^{(E)}_{0i0j} = & -\epsilon^2 \sum_{l=0}^{\infty} \frac{1}{l!} Q_{ijL} w^{<L>} \\
& +\epsilon^4 \Bigg\{ \delta^{ij} \Bigg[(1-\gamma) \left(\frac{\ddot{P}}{3} + \sum_{l=1}^{\infty} \frac{1}{l!} \ddot{P}_L w^{<L>} \right) \\
& - \sum_{l=1}^{\infty} \frac{2(l+2)}{(2l+3)l!} \ddot{Q}_L w^{<L>} \\
& - \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \frac{1}{l!k!} Q_{mL} \Big(Q_{mK} + (\gamma-1) P_{mK} \Big) w^{<L>} w^{<K>} \Bigg] \\
& + \sum_{l=0}^{\infty} \frac{1}{l!} \dot{C}_{pL} \varepsilon^{pq(i} \delta^{j)<q} w^{L>} + \sum_{l=0}^{\infty} \frac{1}{l!} \dot{Z}_{ijL} w^{<L>} \\
& +2 \sum_{l=1}^{\infty} \frac{2l+1}{(2l+3)l!} [2\ddot{Q}_L + (\gamma-1)\ddot{P}_L] \delta^{(i<j)} w^{L>} \\
& -2 \sum_{l=0}^{\infty} \frac{1}{(2l+5)l!} \ddot{Q}^{L(i} w^{j)} w^{<L>} \\
& -\frac{1}{2} \sum_{l=0}^{\infty} \frac{1}{(2l+7)l!} \ddot{Q}_{ijL} w^{<L>} w^2 \\
& +3 \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \frac{1}{l!k!} Q_{iL} Q_{jK} w^{<L>} w^{<K>} \\
& +2 \sum_{l=0}^{\infty} \sum_{k=1}^{\infty} \frac{1}{l!k!} Q_{ijL} Q_K w^{<L>} w^{<K>} \\
& +2(\beta-1) \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \frac{1}{l!k!} P_{iL} P_{jK} w^{<L>} w^{<K>} \\
& +2(\beta-1) \sum_{l=0}^{\infty} \sum_{k=1}^{\infty} \frac{1}{l!k!} P_{ijL} P_K w^{<L>} w^{<K>} \\
& +2(\gamma-1) \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \frac{1}{l!k!} Q_{L(i} P_{j)K} w^{<L>} w^{<K>} \Bigg\} ,
\end{aligned}
\tag{4.164}
$$

1) Equation (4.165) corrects equation (B.2) from paper [Kopeikin and Vlasov, 2004].

$$
\begin{aligned}
R^{(E)}_{0ijk} \quad = \quad & \epsilon^3 \Bigg\{ \sum_{l=1}^{\infty} \frac{l}{(l+2)\,l!} \left[\delta_{ij} \dot{Q}_{kL} - \delta_{ik} \dot{Q}_{jL} \right] w^L \\
& + 2 \sum_{l=0}^{\infty} \frac{1}{(l+3)\,l!} \left[\dot{Q}_{ijL} w^{kL} - \dot{Q}_{ikL} w^{jL} \right] \\
& + (\gamma - 1) \Bigg[\sum_{l=0}^{\infty} \frac{l+1}{(l+2)\,l!} \left(\delta_{ij} \dot{P}_{kL} - \delta_{ik} \dot{P}_{jL} \right) w^L \\
& + \sum_{l=0}^{\infty} \frac{1}{(l+3)\,l!} \left(\dot{P}_{ijL} w^{kL} - \dot{P}_{ikL} w^{jL} \right) \Bigg] \\
& + \sum_{l=0}^{\infty} \frac{l+1}{(l+2)!} \left(\varepsilon_{pjk} C_{ipL} - \frac{1}{2} \varepsilon_{ipj} C_{kpL} - \frac{1}{2} \varepsilon_{ipk} C_{jpL} \right) w^L \\
& - \frac{1}{2} \sum_{l=0}^{\infty} \frac{l}{(l+3)\,l!} \left(\varepsilon_{pqj} C_{kipL} - \varepsilon_{pqk} C_{jipL} \right) w^{<qL>} \Bigg\} ,
\end{aligned} \tag{4.165}
$$

$$
\begin{aligned}
R^{(E)}_{ijkp} \quad = \quad & \epsilon^2 \Bigg\{ \sum_{l=0}^{\infty} \frac{1}{l!} \Big[\delta_{ip} Q_{jkL} + \delta_{jk} Q_{ipL} - \delta_{ik} Q_{jpL} - \delta_{jp} Q_{ikL} \Big] w^L \\
& + (\gamma - 1) \sum_{l=0}^{\infty} \frac{1}{l!} \Big[\delta_{ip} P_{jkL} + \delta_{jk} P_{ipL} - \delta_{ik} P_{jpL} - \delta_{jp} P_{ikL} \Big] w^L \Bigg\} .
\end{aligned} \tag{4.166}
$$

Here, the multipole moments P_L are caused by the presence of the scalar field. One can notice that only three sets of the external moments have real physical meaning. In what follows, the moments P_L, Q_L, and C_L are chosen as the primary external multipoles. Other four multipole moments Y_L, Z_L, F_G and G_L can be chosen arbitrary which corresponds to the presence of four *residual* gauge degrees of freedom generated by the coordinate transformations confined by equation (4.53).

Hereafter one assumes that the angular velocity of dynamic rotation of the local frame

$$
\Omega_i = 0 , \tag{4.167}
$$

which gives $\hat{N}^{(E)}_i(u, \boldsymbol{w}) = 0$. This assumption greatly simplifies subsequent calculations without missing any significant physics. One has to notice, however, that rotating local coordinates have a certain practical value for satellite geodesy [Kaula, 1966; Milani *et al.*, 1987] and global positioning system (GPS) [Ashby, 1998; Ries, 2009].

Various authors used the residual gauge freedom differently in order to take advantage of one or another property of the metric tensor in their analysis. We shall follow the convention accepted in papers [Damour *et al.*, 1991; Kopejkin, 1988a; Soffel *et al.*, 2003; Thorne and Hartle, 1985], which postulate that the space-space component $\hat{g}_{ij}(u, \boldsymbol{w})$ of the local metric tensor must be proportional to the unit matrix (the Kronecker symbol) δ_{ij}. This postulate is in agreement with the standard PPN formalism, where the standard gauge is used to make the space-space component of the metric tensor isotropic [Will, 1993, see §4.2]. The isotropic gauge allows us to eliminate the non-linear term $H_{<ij>}N_{,ij}$ in equation (4.59) in the time-time component of the metric tensor because the STF part $H_{<ij>} = 0$ in this gauge.

In order to diagonalize $\hat{g}_{ij}(u,\boldsymbol{w})$ one chooses the external multipoles F_L, G_L as follows

$$F_L = 0\,, \tag{4.168}$$

$$G_L = 0\,, \tag{4.169}$$

for all $l \geq 0$. Furthermore, one makes the choice of

$$Y_L = Q_L + (\gamma - 1)P_L\,, \tag{4.170}$$

for all Y_L with $l \geq 1$. The monopole moment Y (l=0) is left arbitrary. The gauge restrictions (4.168)-(4.170) imposed on the external multipoles makes the space-space component of the metric tensor diagonal in the entire domain of validity of the local coordinates. In what follows, one prefers to preserve some gauge freedom by leaving the set of the external multipoles Z_L ($l \geq 2$) arbitrary.

Finally, the external metric tensor in the isotropic gauge assumes a "canonical" form

$$\hat{N}^{(E)}(u,\boldsymbol{w}) = 2\sum_{l=0}^{\infty}\frac{1}{l!}Q_L w^L\,, \tag{4.171}$$

$$\begin{aligned}\hat{L}_i^{(E)}(u,\boldsymbol{w}) &= \left(\dot{Y} + \frac{1}{3}\dot{Q} + \frac{1-\gamma}{3}\dot{P}\right)w^i + \sum_{l=1}^{\infty}\frac{1}{l!}\varepsilon_{ipq}C_{pL-1}w^{<qL-1>} \\ &+ 2\sum_{l=1}^{\infty}\frac{2l+1}{(2l+3)(l+1)!}\left[2\dot{Q}_L + (\gamma-1)\dot{P}_L\right]w^{<iL>} + \sum_{l=1}^{\infty}\frac{1}{l!}Z_{iL}w^L\,,\end{aligned} \tag{4.172}$$

$$\hat{H}_{ij}^{(E)}(u,\boldsymbol{w}) = 2\delta_{ij}\left\{Y + \sum_{l=1}^{\infty}\frac{1}{l!}\left[Q_L + (\gamma-1)P_L\right]w^L\right\}\,. \tag{4.173}$$

where the (time-dependent) monopole terms $Q \neq 0$ and $Y \neq 0$.

Now one can compute $\hat{h}_{00}^{(E)}(u,\boldsymbol{w})$ component of the external metric tensor up to the post-Newtonian order by making use of equation (4.59). The most general solution of this equation is determined up to that of a homogeneous wave equation, which has the same form as equation (4.171). For this reason, it is not written explicitly as it is absorbed to the post-Newtonian corrections to the multipole moments Q_L in equation (4.171). Hence, one obtains

$$\hat{L}^{(E)}(u,\boldsymbol{w}) = -2\left(\sum_{l=1}^{\infty}\frac{1}{l!}Q_L w^L\right)^2 - 2(\beta-1)\left(\sum_{l=1}^{\infty}\frac{1}{l!}P_L w^L\right)^2 + \sum_{l=0}^{\infty}\frac{1}{(2l+3)l!}\ddot{Q}_L w^L w^2 \,. \tag{4.174}$$

It is interesting to note that the summation in the first two terms in the right side of equation (4.174) originally starts from $l = 0$ as appears in equations (4.136) and (4.151). However, the product of a harmonic polynomial with the monopoles Q and P represents a homogeneous solution of the Laplace equation and, therefore, can be absorbed to the Newtonian-like polynomial $Q_L w^L$ in equation (4.171) by means of a corresponding re-definition of the multipoles Q_L. This can be always done because the precise mathematical structure of the multipoles Q_L as functions of the coordinates and velocities of the external bodies has not yet been specified. This observation also helps us to recognize that the scalar field affects the external solution of the metric tensor in the local coordinates starting from quadratic, with respect to spatial coordinates $\boldsymbol{w}$, terms only.

External scalar field *per se* can not be eliminated by a pointwise coordinate transformation but it enters the external metric tensor in such a way that it is absorbed to the multipole moments Q_L of the metric tensor. Hence, the external multipoles P_L do not contribute explicitly to the Newtonian equations of translational motion of test particles and extended bodies - only their non-linear combination is observable (see section 6.1).

4.5.2.4 *The Metric Tensor: The Coupling Terms*

The coupling terms in the metric tensor in local coordinates are given as a particular solution of the inhomogeneous equation (4.59) with the right side taken as a product of the internal and external solutions having been found at previous step of the approximation procedure. Solving equation (4.59) yields the coupling terms of the metric tensor in the local coordinates

$$\begin{aligned}\hat{L}^{(C)}(u,\boldsymbol{w}) = &-2\hat{U}^{(B)}(u,\boldsymbol{w})\times\\ &\times\left\{Y + (2\beta-\gamma-1)P + 2\sum_{l=0}^{\infty}\frac{1}{l!}\Big[Q_L + (\beta-1)P_L\Big]w^L\right\}\\ &-2G\sum_{l=1}^{\infty}\frac{1}{l!}\Big[Q_L + 2(\beta-1)P_L\Big]\hat{\mathbf{I}}^{(B)}_{-1}\left\{\rho^* w^L\right\} \,.\end{aligned} \tag{4.175}$$

This completes derivation of the metric tensor in the local coordinates.

4.5.3 Multipolar expansion of gravitational field in the local coordinates

The local coordinates are introduced in the vicinity of each of the gravitating body comprising the N-body system. We consider one of them and call it the "central" body which is indexed by the letter 'B'. This body, for example, can be the Earth, and, in such case, the local coordinates are called the geocentric coordinates [Soffel *et al.*, 2003,?]. Gravitational field of the central body taken alone, that is when all other (external) bodies are ignored, is described in the local coordinates in terms of the metric tensor and scalar field which depend on the internal field potentials $\hat{U}^{(B)}$, $\hat{U}^i_{(B)}$, $\hat{\Phi}^{(B)}_1$, etc., defined in equations (4.141)-(4.148). Multipolar decomposition of the internal metric tensor of the central body is basically equivalent to the procedure of the multipolar decomposition of the gravitational field of N-body in the global coordinates described in section 4.4. However, in contrast to the whole N-body system, the central body is not gravitationally isolated from the other bodies of this system and interacts with their tidal fields. This interaction brings about the coupling terms (4.175) to the metric tensor in the local coordinates, which can contribute to the numerical values of the body's multipole moments in the multipolar decomposition of the local metric tensor. The presence of the coupling terms introduces ambiguity in the definition of the multipoles in the local coordinates and rises a question about how it should be resolved. This problem was first noticed by Thorne and Hartle [Thorne and Hartle, 1985].

The solution of the ambiguity problem can be found only after equations of motion of the central body will have been derived with taking into account for all its multipoles. One has two possibilities: either to include or to exclude the contribution of the coupling terms to the multipole moments of the body, and one has explored both of them [Kopeikin and Vlasov, 2004]. It turns out that the final form of the equations of motion can be significantly simplified if the coupling terms are included to the definition of the multipole moments. In fact, if one excludes the contribution of the coupling terms from the definition of the internal multipoles, it produces a lot of additional terms in the equations of motion, which vanish after suitable re-definition of the multipole moments. One can check that the final form of such re-normalized equations of motion coincides with that, which would be obtained, if one included the coupling terms in the local metric tensor to the definition of the internal multipoles of the central body from the very beginning. The significantly simple form of the re-normalized equations of motion is a direct indication that the coupling terms must be included to the definition of the multipole moments of the body in the local coordinates. This resolves the Thorne-Hartle ambiguity [Thorne and Hartle, 1985] in the definition of the multipole moments.

Thus, the formal procedure of the multipolar decomposition of the gravitational field in the local coordinates is based on the same field equations (4.95), (4.109) and (4.115) for *active*, *scalar* and *conformal* potentials, whose right sides depend on the *active*, *scalar* and *conformal* mass densities defined by equations (4.97), (4.110) and (4.116) respectively. Each density depends on the trace of the space-space component of the metric tensor, H, and the scalar field, φ. In accordance with the definition of the multipole moments in the local coordinates, these functions must include the contribution of the external gravitational and scalar fields. In other words, computation of the mass densities in equations (4.97), (4.110) and (4.116) in the local coordinates must rely upon the trace of the metric tensor, H, defined by the sum of equations (4.140) and (4.173), and the scalar field, φ, defined by the sum of equations (4.135) and (4.136). Solving equations (4.95), (4.109), (4.115) with the above-defined mass densities and expanding the metric potentials in the multipolar series, yield the internal multipole moments of the central body in the local coordinates.

One has constructed three sets of the mass multipole moments - active, scalar, and conformal - in the global coordinates. The same type of the multipoles for each body presents in the local coordinates as well. The *active* STF mass multipole moments of the central body are [Kopeikin and Vlasov, 2004]

$$\begin{aligned} \mathcal{I}_L &= \int_{V_B} \sigma_{B}(u,\boldsymbol{w}) w^{<L>}\, d^3w + \frac{\epsilon^2}{2(2l+3)} \times \\ &\times \left[\frac{d^2}{du^2} \int_{V_B} \sigma_{B}(u,\boldsymbol{w}) w^{<L>} w^2\, d^3w - 4(1+\gamma)\frac{2l+1}{l+1}\frac{d}{du}\int_{V_B} \sigma^i_{B}(u,\boldsymbol{w}) w^{<iL>}\, d^3w \right] \\ &- \epsilon^2 \int_{V_B} d^3w\, \sigma_{B}(u,\boldsymbol{w}) \times \\ &\times \left\{ Y + (2\beta-\gamma-1)P + \sum_{k=1}^{\infty}\frac{1}{k!}\Big[Q_K + 2(\beta-1)P_K \Big] w^K \right\} w^{<L>}\ , \end{aligned} \tag{4.176}$$

where V_B denotes the volume of the body under consideration and the *active* mass density in the body's interior is defined as

$$\sigma_{B} = \rho^* \left\{ 1 + \epsilon^2 \left[(\gamma + \frac{1}{2})v^2 + \Pi + \gamma\frac{\pi^{kk}}{\rho^*} - (2\beta-1)\hat{U}^{(B)} \right] \right\}\ , \tag{4.177}$$

where $\hat{U}^{(B)}$ is the gravitational potential of the body given by equation (4.141).

The *scalar* STF mass multipole moments of the body are defined as

$$\begin{aligned} \bar{\mathcal{I}}_L &= \int_{V_B} \bar{\sigma}_B(u,\boldsymbol{w}) \left\{ 1 - \epsilon^2 \left[Y - \gamma P + \sum_{k=1}^{\infty}\frac{1}{k!} Q_K w^K \right] \right\} w^{<L>}\, d^3w \\ &+ \frac{\epsilon^2}{2(2l+3)}\frac{d^2}{du^2}\int_{V_B} \bar{\sigma}_B(u,\boldsymbol{w}) w^{<L>} w^2\, d^3w \\ &+ \epsilon^2 \int_{V_B} \rho^*(u,\boldsymbol{w}) \left[4(1-\beta)\sum_{k=1}^{\infty}\frac{1}{k!} P_K w^K - \eta P \right] w^{<L>} d^3w\ , \end{aligned} \tag{4.178}$$

where the *scalar* mass density of the body's matter is defined by

$$\bar{\sigma}_B = (1-\gamma)\rho^* \left[1 - \epsilon^2 \left(\frac{1}{2}v^2 - \Pi + \frac{\pi^{kk}}{\rho^*} \right) \right] - \epsilon^2 \eta \rho^* \hat{U}^{(B)}\ . \tag{4.179}$$

The *conformal* STF mass multipole moments of the body are

$$\begin{aligned} \tilde{\mathcal{I}}_L &= \int_{V_B} \tilde{\sigma}_B(u,\boldsymbol{w}) \left\{ 1 - \epsilon^2 \left[Y + (1-\gamma)P + \sum_{k=1}^{\infty}\frac{1}{k!} Q_K w^K \right] \right\} w^{<L>}\, d^3w \\ &+ \frac{\epsilon^2}{2(2l+3)} \left[\frac{d^2}{du^2}\int_{V_B} \tilde{\sigma}_B(u,\boldsymbol{w}) w^{<L>} w^2\, d^3w - \frac{8(2l+1)}{l+1}\frac{d}{du}\int_{V_B} \sigma^i_{B}(u,\boldsymbol{w}) w^{<iL>}\, d^3w \right]\ , \end{aligned} \tag{4.180}$$

with the *conformal* mass density of the body's matter defined as

$$\tilde{\sigma}_{_B} = \rho^* \left[1 + \epsilon^2 \left(\frac{3}{2}\nu^2 + \Pi + \frac{\pi^{kk}}{\rho^*} - \hat{U}^{(B)}\right)\right] . \tag{4.181}$$

Notice that the *conformal* density does not depend on the PPN parameters β and γ.

The density of mass current is defined in the local coordinates by

$$\sigma^i_{_B} = \rho^* \nu^i , \tag{4.182}$$

and the spin multipole moments of the body are determined by the formula

$$S_L = \int_{V_B} \varepsilon^{pq<a_l} \hat{w}^{L-1>p} \sigma^q_{_B}(u, \boldsymbol{w})\, d^3w . \tag{4.183}$$

Extrapolation of this definition to the post-Newtonian approximation is discussed in section 6.2.

It is important to emphasize that the algebraic relationship (4.120) preserves its form for the internal mass multipoles taken for each body separately, that is

$$\mathcal{I}_L = \frac{1+\gamma}{2}\tilde{\mathcal{I}}_L + \frac{1}{2}\bar{\mathcal{I}}_L . \tag{4.184}$$

Validity of this relationship can be checked out by a straightforward calculation [Kopeikin and Vlasov, 2004]. We also draw attention of the reader that the hypersurface of the integration in equations (4.176), (4.178), (4.180) is that of a constant local coordinate time u, which does not coincide with the hypersurface of the constant coordinate time t in the global coordinates (see Figure 4.3). This remark is of a great importance for correct implementation of the matching procedure of the local and global coordinates discussed in the next section. It will be also important in section 6.3 for an adequate mathematical derivation of equations of motion of the body, which requires comparison of the definition of the multipole moments of the body in the local and global coordinates.

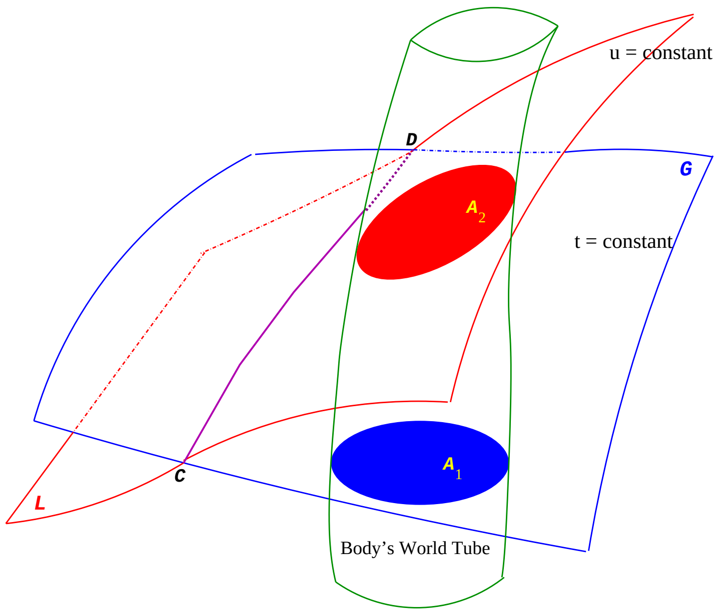

Figure 4.3 Two 3-dimensional hypersurfaces of constant time related to the global (G) and local (L) coordinates are shown. The two hypersurfaces do not coincide because the origin of the local coordinates moves with respect to the global one. The body's world tube embraces a worldline of the origin of the local coordinates. The world tube cross-section, A_1, is a part of the hypersurface of constant time t of the global coordinates, and that, A_2, is a part of the hypersurface of constant time u of the local coordinates. The line CD marks the two-dimensional intersection of the two hypersurfaces. All integrals depending on density, velocity and other internal characteristics of the body, are performed in the global coordinates over the cross-section A_1 whereas in the local coordinates the integration is performed over the cross-section A_2.

5
Post-Newtonian Coordinate Transformations

5.1 The Transformation from the Local to Global Coordinates

5.1.1 Preliminaries

General relativistic post-Newtonian coordinate transformation from local, w^α, to global, x^α, coordinates are used in standard algorithms of data processing of various astronomical observations to reduce the observable quantities to the barycentric coordinates of the solar system [Kovalevsky *et al.*, 1989; Seidelmann and Urban, 2010] where they are stored (cataloged) at a certain astrometric epoch recommended for the international usage by IAU resolutions. The most commonly used epochs are J2000 and B1950, that is, instantaneous orientation of the barycentric coordinates is fixed by the position of vernal equinox on the sky in 2000AD or 1950AD.

Post-Newtonian coordinate transformation from the local to global coordinates is an inalienable part in the procedure of derivation of general relativistic equations of motion of test particles (like artificial satellites), extended bodies in the solar system [Ashby and Bertotti, 1984, 1986; Brumberg and Kopeikin, 1989; Brumberg and Kopejkin, 1989a,b; Damour *et al.*, 1991, 1992; Klioner and Soffel, 2000; Klioner and Voinov, 1993; Kopeikin, 1989a,b; Kopejkin, 1988a, 1991b; Mashhoon, 1985; Shahid-Saless and Ashby, 1988] as well as compact relativistic stars in binary systems emitting gravitational waves [Damour, 1983, 1987; D'Eath, 1975a,b; Grishchuk and Kopeikin, 1986; Kopeikin, 1985]. Though calculations of equations of relativistic celestial mechanics are done mostly in the framework of general relativity, it is widely accepted that scalar fields exist since they provide elegant solution of various problems encountered in modern theoretical physics of fundamental particles, gravitation, and cosmology [Macías *et al.*, 2001]. Therefore, it is natural to generalize the existing form of general-relativistic post-Newtonian transformation from the local to global coordinates [Soffel *et al.*, 2003] to make them compatible with the scalar-tensor theory of gravity [Kopeikin, 2010a]. Solution of this problem in terms of the PPN parameters β and γ has been found in paper [Kopeikin and Vlasov, 2004] and is explained in the present section. Alternative approach was worked out by Klioner and Soffel [Klioner and Soffel, 2000]. It is compared it with the approach of the present book in sections 4.1 and 5.2.1.

The PPN coordinate transformations from the local geocentric (GCRS) coordinates to global barycentric (BCRS) coordinates can be found by making use of the mathematical technique of matching of asymptotic expansions [Eckhaus, 1973] as proposed by D'Eath [1975a,b]. The idea is that the metric tensor represents one and the same solution of the field equations, which is given in the form of two different post-Newtonian expansions depending on whether it is expressed in terms of the global or local coordinates. However one chooses the coordinates the solution must describe the same physical situation. Therefore, the two post-Newtonian expansions of the metric tensor must match smoothly in space domain, where both coordinate charts overlap. This domain of the matching is defined as a region in which the post-Newtonian expansion of the metric tensor exists and is not divergent. In the case of a weak gravitational field this domain extends from the origin of the local coordinates associated with the body under consideration (Earth) up to the closest celestial body (Moon). In N-body system comprised of the bodies with a strong gravitational field like neutron stars and/or black holes, the local coordinates overlap with the global ones in a smaller domain, called the *buffer* region, in which the gravitational field of the body and the tidal gravitational field of external bodies are comparatively weak [D'Eath, 1975a,b; Thorne and Hartle, 1985]. The domain of applicability of the local coordinates can be stretched out to a larger distance if the post-Newtonian expansion of the metric tensor in the local coordinates is written down in a self-closed form [Klioner and Voinov, 1993].

A special relativistic part of transformation from the local to global coordinates is a linear Poincaré transformation (see Section 2.3.6) that takes into account only kinematic aspects of the transformation depending on the velocity of motion of the local frame with respect to the global one. The special relativistic transformation is described at each instant of time by 10 parameters characterizing the structure of the tangent Minkowski spacetime at each point of the worldline of the origin of the local coordinates [Fock, 1964; Landau and Lifshitz, 1975; Misner *et al.*, 1973]. These parameters include 3 space and 1 time translations, 3 spatial rotations, and 3 Lorentz boosts depending on 3-velocity of the origin of the local frame with respect to the global coordinates. General relativity generalizes the Poincaré transformation by accounting for the presence of dynamic effects (acceleration) in the motion of the local frame as well as for the effects of the external gravitational field of the other bodies and the background curvature of the spacetime. Papers [Ashby and Bertotti, 1986; D'Eath, 1975a,b; Kopejkin, 1988a] demonstrated that in general relativity the post-Newtonian coordinate transformation from the local to global coordinates is, actually, non-linear and has more than 10 parameters, all depending on time. Scalar-tensor theory of gravity brings about additional complications to the relativistic theory of reference frames in the solar system that is caused by the dependence of the post-Newtonian transformation on the scalar field, which is described by two more parameters, γ and β. However, the scalar-tensor theory does not change the basic approach to finding the parameterized post-Newtonian (PPN) transformation, which structure remains the same as that used in general theory of relativity. A new feature that must be carefully incorporated to the theory is the matching of the post-Newtonian expansions of the scalar field besides the metric tensor [Kopeikin and Vlasov, 2004].

The PPN coordinate transformation between the global and local coordinates belongs to the class of transformations which must comply with the gauge condition (4.45). Therefore, one should begin with finding the most general structure of such coordinate transformations. As soon as this structure is known it can be further specialized by reducing the number of the gauge degrees of freedom. This is achieved by making use of the matching technique applied for establishing a one-to-one correspondence between the two asymptotic expansions of the metric tensor and scalar field written down in the global and local coordinates. This procedure allows us to deduce the PPN coordinate transformation with the PPN parameters β and γ in the form which is consistent with the covariant nature of the scalar-tensor theory of gravity. We shall show that the PPN transformation from the local to global coordinates can be represented as an expansion in a Taylor series with respect to two small parameters, $\epsilon = 1/c$ and r/R, where r is the distance from the origin of the local coordinates (the body) to the matching point, and R is a characteristic distance from the body to the nearest external body. Coefficients of the power series expansion of the PPN transformation are symmetric and trace-free (STF) tensors depending on time, which are determined in the course of the matching procedure simultaneously with the external multipole moments, P_L, Q_L, C_L of the metric tensor and scalar field. The STF coefficients of the PPN coordinate transformation are functions of time that are 'pinned down' to the origin of the local coordinates. The matching procedure shows that the STF coefficients of the most general form of the PPN coordinate transformation couples linearly with the external STF multipole moments of the local metric tensor in equations (4.151)-(4.154). For this reason the matching allows us to derive only a set of equations defining their algebraic sum. It reflects the presence of the residual gauge freedom, which can be used in order to simplify either the structure of the PPN coordinate transformation or that of the metric tensor in the local coordinates. One has used this gauge freedom in section 4.5.2.3 to suppress the number of the external multipole moments, which have no direct physical meaning. Elimination of the non-physical multipole moments from the metric tensor leads to more simple structure of the PPN transformation as well. These subtle issues are discussed in the next sections in more detail.

5.1.2
General structure of the coordinate transformation

The most general structure of the PPN coordinate transformation from the local, $w^\alpha = (cu, w^i)$, to global, $x^\alpha = (ct, x^i)$, coordinates in the weak-field and slow-motion post-Newtonian approximation is given by two equations:

$$u = t + \epsilon^2 \xi^0(t, \boldsymbol{x}) \,, \tag{5.1}$$

$$w^i = R^i_{B} + \epsilon^2 \xi^i(t, \boldsymbol{x}) \,, \tag{5.2}$$

where ξ^0 and ξ^i are the post-Newtonian corrections to the Galilean transformation: $u = t$, $w^i = R^i_B$, $R^i_B = x^i - x^i_B(t)$, and $x^i_B(t)$ is the position of the origin of the local frame at time t with respect to the origin of the global coordinates. We shall prove later on that the origin of the local coordinates can be always chosen at any instant of time at the center of mass of the central body around which the local coordinates have been constructed. In what follows, velocity and acceleration of the origin of the local coordinates are denoted as $v^i_B \equiv \dot{x}^i_B$ and $a^i_B \equiv \ddot{x}^i_B$ respectively, where here and everywhere else, the overdot must be understood as an ordinary time derivative with respect to time t.

Pointwise matching equations for the scalar field, the metric tensor, and the Christoffel symbols are given by the law of their transformation from one set of coordinates to another [Dubrovin *et al.*, 1984; Misner *et al.*, 1973; Schouten, 1954]

$$\varphi(t,\boldsymbol{x}) = \hat{\varphi}(u,\boldsymbol{w}) \,, \tag{5.3}$$

$$g_{\mu\nu}(t,\boldsymbol{x}) = \hat{g}_{\alpha\beta}(u,\boldsymbol{w})\frac{\partial w^\alpha}{\partial x^\mu}\frac{\partial w^\beta}{\partial x^\nu} \,, \tag{5.4}$$

$$\Gamma^\mu_{\alpha\beta}(t,\boldsymbol{x}) = \hat{\Gamma}^\nu_{\rho\sigma}(u,\boldsymbol{w})\frac{\partial x^\mu}{\partial w^\nu}\frac{\partial w^\rho}{\partial x^\alpha}\frac{\partial w^\sigma}{\partial x^\beta} + \frac{\partial x^\mu}{\partial w^\nu}\frac{\partial^2 w^\nu}{\partial x^\alpha \partial x^\beta} \,. \tag{5.5}$$

One recalls that g_{0i} component of the metric tensor does not contain terms of the order $O(\epsilon)$ because one has stipulated that both the global and the local frames are not dynamically rotating. This fact, after having been implemented in equation (5.4), implies that function $\xi^0(t,\boldsymbol{x})$ from equation (5.1) must be subject to the following restriction

$$\xi^0{}_{,k} = -v^i_{\scriptscriptstyle B} + O(\epsilon^2) \tag{5.6}$$

. This is a partial differential equation, which can be integrated yielding the following result as

$$\xi^0(t,\boldsymbol{x}) = -\mathcal{A}(t) - v^k_{\scriptscriptstyle B} R^k_{\scriptscriptstyle B} + \epsilon^2 \kappa(t,\boldsymbol{x}) + O(\epsilon^4) \,, \tag{5.7}$$

where $\mathcal{A}(t)$ and $\kappa(t,\boldsymbol{x})$ are analytic but otherwise yet unspecified functions except that function $\mathcal{A}(t)$ depends only on time t.

Let us now use the gauge conditions in the form of equation (4.47) in order to impose further restrictions of the PPN functions ξ^0 and ξ^i entering equations (5.1) and (5.2). The gauge conditions (4.47) can be written in arbitrary coordinates as an exact equality

$$g^{\alpha\beta}\Gamma^\mu_{\alpha\beta} = \frac{\zeta^{,\mu}}{1+\zeta} \,, \tag{5.8}$$

where ζ is the perturbation of the scalar field defined in equation (4.13). The law of transformation of the Christoffel symbols, equation (5.5), after being substituted to equation (5.8) yields a partial differential equation of the second order

$$g^{\alpha\beta}(t,\boldsymbol{x})\frac{\partial^2 w^\mu}{\partial x^\alpha \partial x^\beta} = 0 \,, \tag{5.9}$$

which fully agrees with equation (4.53, and describes a residual gauge freedom of the PPN transformation in the scalar-tensor theory of gravity.

Let us now substitute functions $u \equiv w^0/c$ and w^i from equations (5.1) and (5.2), and ξ^0 from equation (5.7) to equation (5.9). One obtains two Poisson equations

$$\nabla^2\kappa(t,\boldsymbol{x}) = 3v^k_{\scriptscriptstyle B} a^k_{\scriptscriptstyle B} - \ddot{\mathcal{A}} - \dot{a}^k_{\scriptscriptstyle B} R^k_{\scriptscriptstyle B} + O(\epsilon^2) \,, \tag{5.10}$$

$$\nabla^2\xi^i(t,\boldsymbol{x}) = -a^i_{\scriptscriptstyle B} + O(\epsilon^2) \,, \tag{5.11}$$

which define two functions κ and ξ^i. General solution of these, elliptic-type equations can be written down in the form of a Taylor expansion with respect to the scalar and vector spherical harmonics as explained in appendix A. Furthermore, the solution for functions $\kappa(t, \boldsymbol{x})$ and $\xi^i(t, \boldsymbol{x})$ consist of two parts - a general solution of the homogeneous Laplace equation and a particular solution of the inhomogeneous Poisson equation. We exclude a singular part of the general solution of the homogeneous equation, which diverges at the origin of the local coordinates $w^i = 0$. Such ill-behaved terms may appear in other alternative theories of gravity, which admit violation of the fundamental law of conservation of the linear momentum. The scalar-tensor theory of gravity does not admit the appearance of such singular functions in the local metric as they have no counterparts and can be matched neither with scalar field nor the metric tensor expressed in the global coordinates.

After taking into account this remark, integration of equations (5.10) and (5.11) results in

$$\kappa = \Big(\frac{1}{2}v^k_{\scriptscriptstyle B}a^k_{\scriptscriptstyle B} - \frac{1}{6}\ddot{\mathcal{A}}\Big)R^2_{\scriptscriptstyle B} - \frac{1}{10}\dot{a}^k_{\scriptscriptstyle B}R^k_{\scriptscriptstyle B}R^2_{\scriptscriptstyle B} + \Xi(t, \boldsymbol{x}) \,, \tag{5.12}$$

$$\xi^i = -\frac{1}{6}a^i_{\scriptscriptstyle B}R^2_{\scriptscriptstyle B} + \Xi^i(t, \boldsymbol{x}) \,, \tag{5.13}$$

where functions Ξ and Ξ^i are non-singular solutions of the homogeneous Laplace equation. These solutions can be written down in the form of harmonic polynomials

$$\Xi(t, \boldsymbol{x}) = \sum_{l=0}^{\infty}\frac{1}{l!}\mathcal{B}_{<L>}R^L_{\scriptscriptstyle B} \,, \tag{5.14}$$

$$\Xi^i(t, \boldsymbol{x}) = \sum_{l=1}^{\infty}\frac{1}{l!}\mathcal{D}_{<iL>}R^L_{\scriptscriptstyle B} + \sum_{l=0}^{\infty}\frac{\varepsilon_{ipq}}{(l+1)!}\mathcal{F}_{<pL>}R^{<qL>}_{\scriptscriptstyle B} + \sum_{l=0}^{\infty}\frac{1}{l!}\mathcal{E}_{<L>}R^{<iL>}_{\scriptscriptstyle B} \,, \tag{5.15}$$

where the polynomial coefficients $\mathcal{B}_{<L>}$, $\mathcal{D}_{<L>}$, $\mathcal{F}_{<L>}$, and $\mathcal{E}_{<L>}$ are time-dependent STF tensors. These tensors are defined on the worldline of the origin of the local coordinates and depend only on the global time t. Explicit form of these functions will be obtained later in the process of matching of the global and local metric tensors as well as the scalar field. Finally, one draws attention of the reader that STF functions entering the coordinate transformations (5.14) and (5.15) have been denoted by capital calligraphic letters, while the STF multipole moments entering the expressions for the local metric, equations (4.171)-(4.173), and the scalar field (4.136) have been denoted by the capital Roman letters. These notations are supposed to help to distinguish the different mathematical origination of these functions.

5.1.3
Transformation of the coordinate basis

Further derivation of the PPN coordinate transformation makes use of the matching equation (5.4) for the metric tensor applied in the domain of overlapping of the local and global coordinates. This equation contains the matrix of transformation $\Lambda^\beta{}_\alpha =$

$\partial w^\beta/\partial x^\alpha$ between the two coordinate bases, $\hat{\boldsymbol{e}}_\alpha \equiv \partial/\partial w^\alpha$ and $\boldsymbol{e}_\alpha \equiv \partial/\partial x^\alpha$, in the local, w^α, and global, x^α, coordinates respectively. Transformation between the bases reads

$$\boldsymbol{e}_\alpha = \Lambda^\beta{}_\alpha \hat{\boldsymbol{e}}_\beta \;, \tag{5.16}$$

and one assumes that the matrix of the transformation $\Lambda^\beta{}_\alpha$ is non-singular. It means that it can be inverted so that the inverse matrix $\text{V}^\beta{}_\alpha$ is defined by a standard rule [Misner *et al.*, 1973]

$$\Lambda^\alpha{}_\beta \text{V}^\beta{}_\gamma = \delta^\alpha_\gamma \;. \tag{5.17}$$

The inverse matrix is required to get the inverse PPN transformation between the two bases. The matrix $\Lambda^\alpha{}_\beta$ can be expanded in the post-Newtonian series, which is a consequence of the post-Newtonian expansion of the coordinate transformation described in a previous section. The post-Newtonian expansion of the matrix of the transformation is as follows

$$\Lambda^0{}_0 = 1 + \epsilon^2 \mathfrak{B}(t,\boldsymbol{x}) + \epsilon^4 \mathfrak{D}(t,\boldsymbol{x}) + O(\epsilon^5) \;, \tag{5.18}$$

$$\Lambda^0{}_i = -\epsilon v^i_{\scriptscriptstyle B} + \epsilon^3 \mathfrak{B}^i(t,\boldsymbol{x}) + O(\epsilon^5) \;, \tag{5.19}$$

$$\Lambda^i{}_0 = -\epsilon v^i_{\scriptscriptstyle B} + \epsilon^3 \mathfrak{P}^i(t,\boldsymbol{x}) + O(\epsilon^5) \;, \tag{5.20}$$

$$\Lambda^i{}_j = \delta^i_j + \epsilon^2 \mathfrak{R}^i{}_j(t,\boldsymbol{x}) + O(\epsilon^4) \;, \tag{5.21}$$

where the coefficients of the expansion are the following functions of the global coordinates:

$$\mathfrak{B}(t,\boldsymbol{x}) = v_B^2 - a_B^k R_B^k - \dot{\mathcal{A}}\,, \tag{5.22}$$

$$\begin{aligned}\mathfrak{D}(t,\boldsymbol{x}) = &\left(\frac{1}{3}\ddot{\mathcal{A}} - v_B^k a_B^k + \frac{1}{5}\dot{a}_B^k R_B^k\right)\left(v_B^j R_B^j\right)\\ &+\left(\frac{1}{2}a_B^2 + \frac{3}{5}v_B^k \dot{a}_B^k - \frac{1}{10}\ddot{a}_B^k R_B^k - \frac{1}{6}\dddot{\mathcal{A}}\right)R_B^2\\ &+\sum_{l=0}^{\infty}\frac{1}{l!}\left(\dot{\mathcal{B}}_{<L>} - v_B^k \mathcal{B}_{<kL>}\right)R_B^L\,,\end{aligned} \tag{5.23}$$

$$\mathfrak{B}^i(t,\boldsymbol{x}) = \left(v_B^k a_B^k - \frac{1}{3}\ddot{\mathcal{A}}\right)R_B^i - \frac{1}{10}\dot{a}_B^i R_B^2 - \frac{1}{5}\dot{a}_B^k R_B^k R_B^i + \sum_{l=0}^{\infty}\frac{1}{l!}\mathcal{B}_{<iL>}R_B^L\,, \tag{5.24}$$

$$\begin{aligned}\mathfrak{P}^i(t,\boldsymbol{x}) = &\frac{1}{3}a_B^i v_B^k R_B^k - \frac{1}{6}\dot{a}_B^i R_B^2 + \sum_{l=1}^{\infty}\frac{1}{l!}\left(\dot{\mathcal{D}}_{<iL>} - v_B^k \mathcal{D}_{<ikL>}\right)R_B^L - v_B^k \mathcal{D}_{<ik>}\\ &+\varepsilon_{ipq}\sum_{l=1}^{\infty}\frac{1}{l!}\left(\dot{\mathcal{F}}_{<pL-1>} - \frac{l}{l+1}v_B^k \mathcal{F}_{<pkL-1>}\right)R_B^{qL-1}\\ &-v_B^k \varepsilon_{ipk}\sum_{l=0}^{\infty}\frac{1}{(l+1)!}\mathcal{F}_{<pL>}R_B^L + \sum_{l=0}^{\infty}\frac{1}{l!}\dot{\mathcal{E}}_{<L>}R_B^{<iL>}\\ &-\sum_{l=0}^{\infty}\frac{l+1}{l!}\mathcal{E}_{<L>}v_B^{<i}R_B^{L>}\,,\end{aligned} \tag{5.25}$$

$$\begin{aligned}\mathfrak{R}^i{}_j(t,\boldsymbol{x}) = &-\frac{1}{3}a_B^i R_B^j + \sum_{l=0}^{\infty}\frac{1}{l!}\left(\mathcal{D}_{<ijL>} + \delta_{ij}\mathcal{E}_{<L>} + \frac{1}{l+1}\varepsilon_{ipj}\mathcal{F}_{<pL>}\right)R_B^L\\ &+\varepsilon_{ipq}\sum_{l=0}^{\infty}\frac{l+1}{(l+2)!}\mathcal{F}_{<jpL>}R_B^{qL} + \sum_{l=0}^{\infty}\frac{1}{l!}\mathcal{E}_{<jL>}R_B^{<iL>}\\ &-2\sum_{l=0}^{\infty}\frac{1}{(2l+3)l!}\mathcal{E}_{<iL>}R_B^{<jL>}\,.\end{aligned} \tag{5.26}$$

Components of the inverse matrix $\daleth^\alpha{}_\beta$ can be deduced from equations (5.22)-(5.26) by applying relationship (5.17).

Formulas (5.22)-(5.26) allow us to evaluate the size of the entire domain of applicability of the local coordinates. The radius of this domain is determined by the condition that the determinant of the matrix $\Lambda^\alpha{}_\beta$ of the coordinate transformation is zero. Calculating the determinant of the matrix one obtains

$$\begin{aligned}\det\left(\Lambda^\alpha{}_\beta\right) = &1 + \epsilon^2\left[-\dot{\mathcal{A}} + 3\,\mathcal{E} - \frac{4}{3}\left(a_B^k - \frac{5}{2}\mathcal{E}_k\right)R_B^k\right.\\ &\left.+\sum_{l=2}^{\infty}\frac{(l+1)(2l+3)}{(2l+1)l!}\mathcal{E}_{<L>}R_B^L\right] + O\left(\epsilon^4\right)\,.\end{aligned} \tag{5.27}$$

Radius of convergence of the polynomial in the right side of equation (5.27) crucially depends on the choice of functions $\mathcal{E}_L$. Later on, one will derive equations (5.64), (5.65)) which will be used to prove that it is possible to make function $\mathcal{E}_i = a^i_{_B}$, and all other functions $\mathcal{E}_L = 0$ for any $l \geq 2$. Then, if one also takes into account equations (5.57) and (5.63) from the next section (putting for simplicity $Q = Y = 0$ in there), the determinant (5.27) turns to zero at the distance $R_{_B} \approx c^2/(2a_{_B})$. In case of the local geocentric frame attached to the Earth and moving around the Sun with acceleration $a_{_B} \simeq 0.6cm/s^2$, the distance $R_{_B}$ is about 10^{21} cm. Hence, the local geocentric frame covers a region which includes the entire solar system. In case of a binary pulsar with a characteristic size of the orbit $\sim 10^{10}$ cm, the local coordinate system attached to the pulsar covers the region with a radius about 10^{14} cm, which also significantly exceeds the distance between the pulsar and its companion. This remark can be important for researchers doing analysis of the physical processes going on in pulsar's magnetosphere [Mestel, 2000].

This consideration suggests that the metric tensor defined originally in the local coordinates only in the domain restricted by the distance to the nearest external gravitating body can be extrapolated far beyond this boundary. Such extrapolation can be accomplished by choosing another form of the solution of the homogeneous field equations in the local coordinates describing the background gravitational field of external bodies. Some research in this direction had been pursued by Klioner and Voinov [1993].

5.2
Matching Transformation of the Metric Tensor and Scalar Field.

5.2.1
Historical background

PPN coordinate transformations (5.18)-(5.26)) introduced in the previous chapter depend on a number undetermined functions that are to be specified for practical astronomical calculations in the solar system. Moreover, the metric tensor in the local coordinates (4.171)-(4.175) is expressed in terms of the external multipole moments Q_L, P_L, etc., which have not been specified in terms of coordinates and velocities of the solar system either. On the other hand, the metric tensor (4.68)-(7.62.7) in the global coordinates is fully defined in terms of the integrals taken over the distribution of matter of the N body system. It is clear that the two post-Newtonian expressions of the metric tensor must match each other if the PPN coordinate transformation of the metric tensor is applied. The same rule should be true for the scalar field. The matching determines the external multipole moments in the local coordinates and the functions in the PPN coordinate transformation, thus, fixing the residual gauge freedom. This mathematical technique is known in the theory of differential equations as the method of matched asymptotic expansions (also called the boundary layer method [Zwillinger, 1998]), which has been originally developed for finding solutions of the ordinary and partial differential equations long time ago [Eckhaus, 1973; Hinch, 1991].

The idea of implementing this method in general relativity goes back to a earlier work by Einstein and Rosen [1935], where the authors discussed the problem of motion of gravitationally-interacting particles by treating them as topological structures ('bridges') on spacetime manifold endowed with metric that is a regular solution of the Einstein field equations. Fock [Fock, 1964] had applied the matching technique to join two expansions of the metric tensor in the near and far (radiative) zones of an isolated astronomical system emitting gravitational radiation [1] Manasse [1963] studied a radial fall of a small black hole onto a massive gravitating body and calculated a tidal distortion of the black hole's horizon by making use of the matching technique. Thorne [Thorne, 1969] and Burke [Burke, 1971] suggested to use the matching technique for imposing an outgoing-wave radiation condition on the post-Newtonian metric tensor for an isolated N body system emitting gravitational waves. This method helps to pick up a causal solution of the homogeneous Einstein equations in the post-Newtonian approximation scheme and to postpone the appearance of ill-behaved (divergent) integrals, at least, up to the fourth PNA [Anderson *et al.*, 1982; Anderson and Madonna, 1983; Kates and Madonna, 1982].

Demiański and Grishchuk [1974] used the matching technique to show that a black hole orbits its companion of a comparable mass in accordance with the Newtonian equations of motion. At about the same time, D'Eath [1975a,b] had explored the idea of matching, proposed by Hawking, to work out a detailed post-Newtonian solution

1) Blanchet [Blanchet, 1987] discusses Fock's procedure in more detail.

of the problem of motion of two Kerr black holes comprising a binary system. He made use of the matching of the internal (local coordinates) and external (global coordinates) solutions of the Einstein equations and successfully derived general-relativistic (*a la* Einstein-Infeld-Hoffmann) equations of motion of the black holes in the first post-Newtonian (1 PN) approximation. Kates [1980a,b] extended the analysis and obtained gravitational radiation-reaction force (2.5 PNA) for two black holes making up a binary system. He has also elaborated on a rigorous mathematical treatment of the matched asymptotic expansions technique for various applications in general relativity [Kates, 1981]. Damour [1983] used the asymptotic matching technique to solve the problem of motion of two spherically-symmetric and non-rotating compact bodies with the gravitational radiation reaction force taken into account. He proved that the body's mass which appears in the external solution of the two-body problem as a constant parameter, is the same as that characterizing the Schwarzschild metric of a non-rotating black hole. Thorne and Hartle [1985] applied the matching technique to study the problem of translational motion and precession of compact bodies possessing the intrinsic quadrupole moments. Their method being combined with the mathematical technique of D'Eath [1975a], was employed by Kopejkin [1988a] to derive the post-Newtonian equations of motion of extended bodies in N-body system in the weak-field and slow-motion approximation. The paper [Kopejkin, 1988a] also demonstrates how to construct a local coordinate system in the post-Newtonian approximation with the origin moving exactly along the worldline of the center of mass of the extended body having arbitrary shape and rotating.

The matching technique used by Kopejkin [1988a] led to the development of the Brumberg-Kopeikin (BK) formalism in the theory of astronomical reference frames for the solar system. Later on, it has been improved in a series of subsequent publications [Brumberg and Kopeikin, 1989, 1990; Brumberg and Kopejkin, 1989a,b; Klioner and Voinov, 1993; Kopeikin and Vlasov, 2004; Kopejkin, 1991b]. Similar matching technique was used in papers by Damour *et al.* [1991, 1992] resulting in Damour-Soffel-Xu (DSX) formalism of the post-Newtonian celestial mechanics of an isolated astronomical system. Both BK and DSX formalisms are complementary and were used in making resolutions on the relativistic reference frames and time scales in the solar system adopted by the 20-th General Assembly of the IAU [Soffel *et al.*, 2003]. In the present book the technique of the matched asymptotic expansions is applied in the framework of the scalar-tensor theory of gravity. It helps us to incorporate the IAU resolutions on reference frames in the solar system to the parameterized post-Newtonian (PPN) formalism [Will, 1993] linking experimental gravity and fundamental astronomy.

Both BK and DSX formalisms use the same theoretical principles in constructing the post-Newtonian celestial mechanics of N-body system except for minor mathematical differences. This book elaborates further on the BK-DSX formalism for the purposes of applied relativity in astronomy. More specifically, this book:

- extends the BK-DSX formalism to the class of scalar-tensor theories of gravity;
- gives the post-Newtonian definition of the internal multipole moments in the

scalar-tensor theory of gravity and incorporates them to the PPN coordinate transformations and equations of motion;
- develops the correct matching procedure of the metric tensor and scalar field in the scalar-tensor theory of gravity;
- constructs a set of global and local coordinates;
- describes relativistic celestial dynamics of the N-body system consisting of the rotating bodies possessing any number of the internal multipoles.

We follow closely to the line of study published by Kopeikin and Vlasov [2004].

It should be noticed that Damour and Esposito-Farese [1992] also proposed definition of the post-Newtonian multipole moments for an isolated N-body system in the scalar-tensor theory of gravity. However, they have focused on discussing experimental tests of this theory with pulsar timing and gravitational wave astronomy but did not work out the matching procedure for construction of the astronomical reference frames in the solar system.

Klioner and Soffel [Klioner and Soffel, 2000] have also tried to re-build the standard PPN formalism [Will, 1993] by making use of the DSX matching technique and to construct the PPN local coordinates for a parameterized description of the celestial mechanics of N-body system. They have used the PPN-parameterized definitions of the post-Newtonian multipoles given in [Damour and Esposito-Farese, 1992]. However, the paper [Klioner and Soffel, 2000] did not rely upon a particular class of alternative gravitational theory and abandoned the use of gravitational field equations in the matching procedure. It makes the Klioner-Soffel parametrization of the IAU theory of reference frames doubtful (see [Kopeikin and Vlasov, 2004] for further discussion).

5.2.2
Method of the matched asymptotic expansions in the PPN formalism

Method of the matched asymptotic expansions is the most powerful mathematical tool for solving differential equations depending on small parameters for which a regular perturbation series method fails. It occurs any time when solution of the differential equation can not satisfy to all boundary conditions simultaneously [Eckhaus, 1973; Hinch, 1991]. In this case the solution is looked for in the form of several asymptotic expansions each being valid in a smaller domain of spacetime and satisfying, at least, one of the boundary conditions. Subsequent matching of the asymptotic expansions in a buffer region (boundary layer) where, at least, two of the asymptotic expansions are valid and convergent, allows us to find the law of transformation from one expansion to another and to retrieve coefficients of these expansions. This completes the procedure of solving the differential equation.

In the present book the asymptotic post-Newtonian expansions, which are used in the matching procedure, are solutions of the gravity field equations for the metric tensor and scalar field. They are found in the global and local coordinates and are subject to different boundary conditions imposed respectively at infinity and at the origin of the local coordinates. These solutions are shown in equations (4.68)-(4.72) and (4.135), (4.136), (4.171)-(4.175). Solution for the metric tensor and scalar field in the global coordinates is valid everywhere inside and outside of the N-body system up to infinity including the interior of the bodies. This is because one has stipulated that in global coordinates the gravity field is weak everywhere, regular at infinity, and has no singularities because one works with a smooth distribution of matter having continuous tensor of energy-momentum. Had one worked with the tensor of energy-momentum of massive, point-like particles one would have singularities on the particle's worldlines [Infeld and Plebanski, 1960]. In such case the method of matched asymptotic expansions is the only way to derive the equations of motion of the particles without ambiguities [Blanchet, 2002b].

Because the spacetime singularities are not admitted, it is tempting to think that a single global coordinate chart is sufficient to describe the post-Newtonian celestial dynamics of the N-body system. This idea was indeed dominating in earlier works deriving the relativistic equations of motion [Brumberg, 1972; Fock, 1964; Infeld and Plebanski, 1960; Landau and Lifshitz, 1975; Soffel, 1989]. A single coordinate chart is still implemented in relativistic astrophysics for calculating equations of motion of massive point particles [Damour and Schäfer, 1985; Jaranowski and Schäfer, 1998; Memmesheimer and Schäfer, 2005; Schäfer, 1985] in a model of coalescing binary systems consisting of two black holes. However, the single coordinate chart is not sufficient for purposes of applied relativity and relativistic celestial mechanics of the solar system for two reasons.

First, the local coordinates are required to give physically meaningful definition of the internal multipole moments of each body In addition, one must know how this definition relates to the definition of these multipoles given in the global coordinates. This relationship between the two definitions of the internal multipole moments is a key element in the procedure of derivation of the equations of motion of extended bodies having finite size [Kopeikin and Vlasov, 2004]. We discuss the importance of this issue in section 6.3 in more detail.

Second, the global barycentric coordinates of the solar system are not appropriate reference frame for analysis of gravitational experiments and motion of artificial satellites orbiting the Earth [Brumberg and Kopejkin, 1989a]. They are also inappropriate in data processing of Very Long Baseline Interferometry (VLBI) [Kopeikin, 1990]. This is because the Earth is embedded to the gravitational field of other bodies of the solar system and moves with acceleration with respect to them (orbits the barycenter of the solar system). Simple translation of the origin of the global coordinates to the geocenter (the Galilean transformation) that was frequently used in early publications [Brumberg, 1972; Caporali, 1981; Fock, 1964; Spyrou, 1975; Will, 1993] does not take into account relativistic aspects of the coordinate transformations on the curved spacetime manifold and, hence, can not eliminate a large number of coordinate-dependent (nonphysical) effects, which complicate adequate physical interpretation of astronomical observations [Brumberg and Kopejkin, 1989b; Kopejkin, 1988a; Mashhoon, 1985; Soffel *et al.*, 1986]. Will [1993] understood this problem fairly well but he did not propose any method for construction of the local coordinates in the vicinity of a massive body in the post-Newtonian approximation. The number of the coordinate-dependent effects becomes much smaller if one uses correct relativistic procedure to transform the global coordinates to the local ones. Such post-Newtonian transformation simplifies drastically the analysis of astronomical observations and the description of relativistic dynamics of the lunar motion and artificial satellites [Brumberg, 1991; Brumberg and Kopejkin, 1989a; Damour *et al.*, 1994; Xie and Kopeikin, 2010].

The internal solution for the metric tensor and scalar field in the local coordinates contain the external multipole moments which can not be found as explicit functions of time without matching of the local solution of the gravity field equations to the global one. The matching allows us to express the external multipole moments in terms of the gravitational potentials (4.74) characterizing the global metric tensor and scalar field. At the same time the matching procedure determines the structure and a particular form of the PPN coordinate transformation between the global and local coordinates.

Solution of some problems in cosmology [Deruelle and Mukhanov, 1995] requires to match not only the metric tensor but its first derivatives as well. This is typical in case of the presence of one or more boundary layers separating matter's world tube from empty spacetime. The problem of the boundary layers is beyond of the scope of this book. One stipulates that the metric tensor and scalar field are smoothly differentiable functions in the matching domain and their first derivatives have no jumps. Therefore, equation (5.5) is a consequence of the matching equation (5.4). Matching of the local and global solutions of the metric tensor and scalar field is based on equations (5.3) and (5.4), and consists of the following steps:

Step 1. One re-writes expressions for the local metric tensor and scalar field in the right side of equations (5.3) and (5.4) in terms of the global coordinates $(t, \boldsymbol{x})$. This is achieved by making use of a Taylor expansion of $\hat{\varphi}(u, \boldsymbol{w})$ and $\hat{g}_{\alpha\beta}(u, \boldsymbol{w})$ around the point $x^\alpha = (ct, \boldsymbol{x})$.

Step 2. One calculates the partial derivatives of the local coordinates with respect to the global ones, that is the matrix of transformation of the coordinate bases given in section (5.1.3).

Step 3. One separates the gravitational potentials in the left side of equations (5.3) and (5.4) in two parts - one relating to the body under consideration (Earth) and the other characterizing the external bodies (Moon, Sun, etc.):

$$U(t, \boldsymbol{x}) = U^{(B)}(t, \boldsymbol{x}) + \bar{U}(t, \boldsymbol{x}) \,, \tag{5.28}$$

$$U^i(t, \boldsymbol{x}) = U_i^{(B)}(t, \boldsymbol{x}) + \bar{U}^i(t, \boldsymbol{x}) \,, \tag{5.29}$$

$$\chi(t, \boldsymbol{x}) = \chi^{(B)}(t, \boldsymbol{x}) + \bar{\chi}(t, \boldsymbol{x}) \,, \tag{5.30}$$

$$\Phi_k(t, \boldsymbol{x}) = \Phi_k^{(B)}(t, \boldsymbol{x}) + \bar{\Phi}_k(t, \boldsymbol{x}) \,, \qquad (k = 1, ..., 4) \,, \tag{5.31}$$

where functions with index (B) are given by integrals (4.75)-(4.82) taken only over the volume of the body in question, and the bar over the other functions indicates that the corresponding sum in the definitions (4.74) of these functions excludes the body with the index B, that is the sum takes into account only the external bodies

$$\bar{U} = \sum_{A\neq B} U^{(A)}, \qquad \bar{U}_i = \sum_{A\neq B} U_i^{(A)}, \qquad \bar{\Phi}_k = \sum_{A\neq B} \Phi_k^{(A)}, \qquad \bar{\chi} = \sum_{A\neq B} \chi^{(A)}. \tag{5.32}$$

Step 4. One expands the gravitational potentials of the external masses (that is functions with the bars in equations (5.28)-(5.32)) in a Taylor series in powers of $R_B^i = x^i - x_B^i$ in the vicinity of the origin of the local coordinates, that is the point $x^i = x_B^i$.

Step 5. One equates similar terms of these Taylor expansions from the left side of the matching equations (5.3) and (5.4) with the corresponding Taylor expansions entering their right side.

Step 6. One separates the matching equations in symmetric and anti-symmetric parts that are algebraically independent. Equating left and right sides of these equations determine all functions in the local metric tensor and scalar field, which remained undetermined so far, as well as the form of the expansion coefficients in the coordinate transformations. This fixes the residual gauge freedom and brings about the laws of translational and rotational motion of the local reference frame.

Let us now explain each step of the matching procedure in more detail.

5.2.3
Transformation of gravitational potentials from the local to global coordinates

5.2.3.1 *Transformation of the Internal Potentials*

At the first step of the matching procedure one has to transform the metric tensor and the scalar field in the right side of matching equations (5.3) and (5.4) from the local, $w^\alpha = (cu, \boldsymbol{w})$, to global, $x^\alpha = (ct, \boldsymbol{x})$, coordinates. It is also conceivable to make a reciprocal transform of all functions in the left side of equations (5.3) and (5.4) to the local coordinates $w^\alpha = (cu, \boldsymbol{w})$. However, it is more convenient and simpler to express the metric tensor and scalar field from the right side of equations (5.3) and (5.4) in global coordinates $x^\alpha = (ct, \boldsymbol{x})$ in accordance to the transformations (5.1), (5.2), (5.12)-(5.15), which are already displayed in terms of the global coordinates. One reminds that the internal gravitational potentials associated with the scalar field, equation (4.135), and the metric tensor, equations (4.141)-(4.148), are defined in the local coordinates $w^\alpha = (cu, \boldsymbol{w})$ as integrals over a hypersurface of constant coordinate time u. On the other hand, the corresponding gravitational potentials, equations (4.75)-(4.81), are defined in the global coordinates $x^\alpha = (t, \boldsymbol{x})$ as integrals over a hypersurface of constant coordinate time t. These two hypersurfaces do not coincide in the post-Newtonian theory of gravity and can intersect only at the points that form a two-dimensional sub-hypersurface (see Figure 4.3). For this reason, in order to transform the internal potentials defined in the local coordinates, $w^\alpha = (cu, \boldsymbol{w})$, to those defined in the global coordinates, $x^\alpha = (ct, \boldsymbol{x})$, one needs to make a point-wise transformation given by equations (5.1)-(5.2) along with a Lie transform of integrands of the integrals, which displaces the integrands from the hypersurface of constant time u to that of constant time t (see Fig 5.1). This procedure was worked

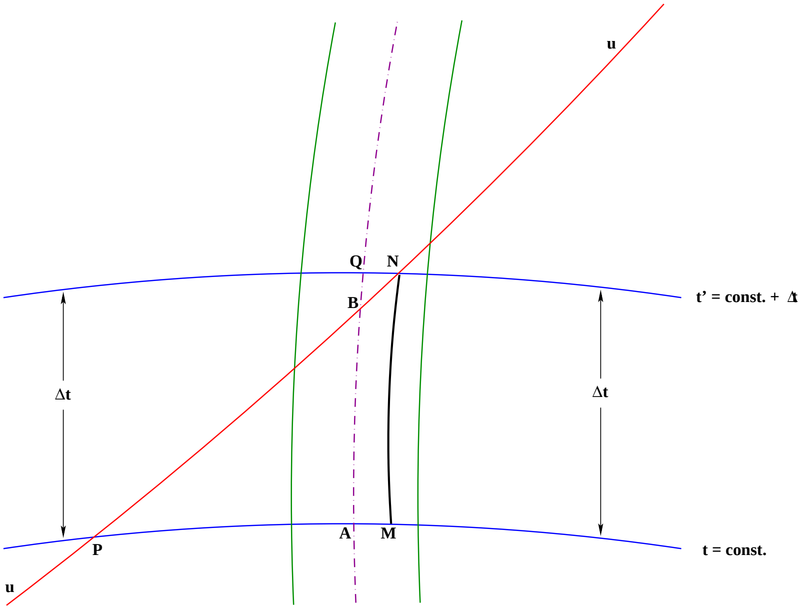

Figure 5.1 Matching the local and global coordinates requires to calculate integrals from various functions, depending on the internal structure of the body, over two different hypersurfaces of constant time, u and t, as shown in Figure 4.3. Relationship between the integrals taken on the two hypersurfaces is established with the help of the Lie transport from one hypersurfaces to another. The integral curves used in the Lie transport are the worldlines of the four-velocity of body's matter. For the sake of simplicity, only one of such lines, MN, is shown in the figure. Dashed line is the worldline of the origin of the local coordinates, which coincides with the center of mass of the body under consideration. In the most general case, the worldline of the center of mass is not tangent to the four-velocity of any material particle inside the body.

out by Kopejkin [1991b] and applied by Brumberg and Kopejkin [1989a] for solving the problem of motion of artificial satellites. Pati and Will [2000] discuss it for derivation of equations of motion of extended bodies via direct integration of the relaxed Einstein equations. We describe the Lie transport below in more detail.

Let us assume that the field point P at which the matching of the internal and external solutions of the metric tensor and the scalar field is done, has global coordinates $x^\alpha(\mathsf{P}) = (ct, \boldsymbol{x})$ and local coordinates $w^\alpha(\mathsf{P}) = (cu, \boldsymbol{w})$ (see Figure 5.1). Notice that the coordinates have different numerical values but the physical point ^{P} is one and the same. These coordinates are related by the pointwise PPN transformation given by equations (5.1)-(5.2). By definition, the matching point P belongs simultaneously to both the hypersurface of constant time u and to that of the constant time t. Let us consider a matter element of the body B located at the point N and lying on the same hypersurface of the constant time u inside body's world tube. One assumes that the point N has local coordinates, $w^\alpha(\mathsf{N}) = (cu, \boldsymbol{w}'(u))$, and global coordinates, $x^\alpha(\mathsf{N}) = (ct', \boldsymbol{x}'(t'))$. We emphasize that the time coordinate u of the points P and N has the same numerical value because they are located on the same hypersurface of the constant time u. However, the value of the time coordinate t' of the point N is different from that of the time coordinate t of the point P, that is $t' \neq t$, because the hypersurfaces of the constant global time passing through the points P and N do not coincide. Let us consider a worldline of the element of the body's matter passing through the point N and intersecting a hypersurface of the constant time t at the point M. This worldline allows us to map the coordinates of the element of the body's matter from the hypersurface t' to the hypersurface t. By the construction, the point M must have the global coordinates $x^\alpha(\mathsf{M}) = (ct, \boldsymbol{x}'(t))$.

One can expand the spatial coordinates of the element of the body's matter in a Taylor series with respect to time

$$x'^i(t') = x'^i(t) + v'^i(t)(t' - t) + O\left(\Delta t^2\right) , \tag{5.33}$$

where v'^i is the spatial velocity of the matter element taken at the point M, and $\Delta t \equiv t' - t$. Let us use letters O and Q to denote positions of the origin of the local coordinates on two hypersurfaces, t and t', respectively (see Figure 5.1). Global coordinates of the origin of the local coordinates, taken on two different hypersurfaces, are related by equation

$$x^i_{\scriptscriptstyle B}(t') = x^i_{\scriptscriptstyle B}(t) + v^i_{\scriptscriptstyle B}(t)(t' - t) + O\left(\Delta t^2\right) . \tag{5.34}$$

The time interval $\Delta t = t' - t$ separating the two hypersurfaces of the constant global coordinate time, t and t', is found under condition that the matching point P is fixed. According to equation (5.1) the relationship between the local time u and the global time t' at the point N is

$$u = t' + \epsilon^2 \xi^0(t', \boldsymbol{x}') + O\left(\epsilon^4\right) . \tag{5.35}$$

Subtracting equation (5.1) from equation (5.35) and accounting for the fact that the spacetime interval between points N and P is small, one obtains

$$\Delta t = t' - t = \epsilon^2 \left[\xi^0(t, \boldsymbol{x}) - \xi^0(t, \boldsymbol{x}')\right] + O\left(\epsilon^4\right) , \tag{5.36}$$

where the notation $\boldsymbol{x}' \equiv \boldsymbol{x}'(t)$ emphasizing that the point $\boldsymbol{x}'$ belongs to the hypersurface of the constant time t.

Local coordinates of the point N are transformed to the global coordinates as follows

$$w'^i = x'^i(t') - x_B(t') + \epsilon^2 \xi^i\left(t', \boldsymbol{x}'(t')\right) . \tag{5.37}$$

Expanding functions in the right side of equation (5.37) in the vicinity of the time instant, t, and taking into account equations (5.33)-(5.36) yield

$$w'^i = R'^i_B + \epsilon^2 \left[\xi'^i - (v'^i - v^i_B)\left(\xi'^0 - \xi^0\right)\right] + O\left(\epsilon^4\right) , \tag{5.38}$$

where $\xi'^i \equiv \xi^i(t, \boldsymbol{x}')$, $\xi'^0 \equiv \xi^0(t, \boldsymbol{x}')$, and $R'^i_B = x'^i(t) - x^i_B(t)$.

Transformation (5.38) is used for deriving a relationship between the absolute values of distances $|\boldsymbol{w}' - \boldsymbol{w}|$ and $|\boldsymbol{x}' - \boldsymbol{x}|$, which enter the denominators of the integrands in the integrals defining the internal gravitational potentials. Subtracting equation (5.2) from equation (5.38) and taking the absolute value of the difference, give us the following equation

$$|\boldsymbol{w}' - \boldsymbol{w}| = |\boldsymbol{x}' - \boldsymbol{x}| + \epsilon^2 \left[n^k\left(\xi'^k - \xi^k\right) - n^k\left(v'^k - v^k_B\right)\left(\xi'^0 - \xi^0\right)\right] + O\left(\epsilon^4\right) , \tag{5.39}$$

where $\xi'^i \equiv \xi^i(t, \boldsymbol{x}')$, $\xi^i \equiv \xi^i(t, \boldsymbol{x})$, and $n^i \equiv (x'^i - x^i)/|\boldsymbol{x}' - \boldsymbol{x}|$.

We must also perform a Lie transform to find a relationship between the volume elements d^3w' and d^3x' taken at the points N and M respectively. We note that the invariant density ρ^* introduced in equation (4.43), possesses one more remarkable property in addition to equation (4.44) pointed out by Kopejkin [1991b]. Specifically, the Lie derivative of the product of the invariant density and the volume element is zero, which means that this product is Lie invariant

$$\rho^*\left(t, \boldsymbol{x}'(t)\right) d^3x'(\mathsf{M}) = \rho^*\left(t', \boldsymbol{x}'(t')\right) d^3x'(\mathsf{N}) = \rho^*\left(u, \boldsymbol{w}'\right) d^3w'(\mathsf{N}) . \tag{5.40}$$

Locally measurable velocity ν'^i of the body's element at the point N is defined with respect to the origin of the local coordinate system. It relates to velocity v'^i of the same element of the body taken at point M in the global coordinates by equation

$$\nu'^i(u) = v'^i(t) - v^i_B(t) + O\left(\epsilon^2\right) , \tag{5.41}$$

which can be derived from equation (5.38) by direct differentiation with respect to time.

Equations (5.39)-(5.41) allow us to obtain the post-Newtonian transformations of the internal gravitational potentials from the local to global coordinates. Calculations are straightforward and are not reproduced here. The transformations of the potentials are given by the following equations

$$\hat{U}^{(B)}(u,\boldsymbol{w}) = U^{(B)}(t,\boldsymbol{x}) + \epsilon^2\mathcal{U}^{(B)}(t,\boldsymbol{x}) + O(\epsilon^4)\,, \tag{5.42}$$

$$\hat{U}^{(B)}_i(u,\boldsymbol{w}) = U^{(B)}_i(t,\boldsymbol{x}) - v^i_{_B}U^{(B)}(t,\boldsymbol{x}) + O(\epsilon^2)\,, \tag{5.43}$$

$$\hat{\chi}^{(B)}(u,\boldsymbol{w}) = \chi^{(B)}(t,\boldsymbol{x}) + O(\epsilon^2)\,, \tag{5.44}$$

$$\hat{\Phi}^{(B)}_1(u,\boldsymbol{w}) = \Phi^{(B)}_1(t,\boldsymbol{x}) + v^2_{_B}U^{(B)}(t,\boldsymbol{x}) - 2v^i_{_B}U^i_{(B)}(t,\boldsymbol{x}) + O(\epsilon^2)\,, \tag{5.45}$$

$$\hat{\Phi}^{(B)}_2(u,\boldsymbol{w}) = \Phi^{(B)}_2(t,\boldsymbol{x}) - G\mathbf{I}^{(B)}_{-1}\left\{\rho^*\bar{U}(t,x)\right\} + O(\epsilon^2)\,, \tag{5.46}$$

$$\hat{\Phi}^{(B)}_3(u,\boldsymbol{w}) = \Phi^{(B)}_3(t,\boldsymbol{x}) + O(\epsilon^2)\,, \tag{5.47}$$

$$\hat{\Phi}^{(B)}_4(u,\boldsymbol{w}) = \Phi^{(B)}_4(t,\boldsymbol{x}) + O(\epsilon^2)\,, \tag{5.48}$$

where the post-Newtonian correction $\mathcal{U}^{(B)}(t,\boldsymbol{x})$ to the Newtonian potential $U^{(B)}(t,\boldsymbol{x})$ reads

$$\mathcal{U}^{(B)}(t,\boldsymbol{x}) = G\mathbf{I}^{(B)}_{-2}\left\{\rho^* n^k\left(v'^k - v^k_{_B}\right)\left(\xi'^0 - \xi^0\right) - \rho^* n^k\left(\xi'^k - \xi^k\right)\right\}\,. \tag{5.49}$$

This correction is the result of the post-Newtonian coordinate transformation (5.1), (5.2) and the Lie transport of the integrand of the Newtonian gravitational potential from one hypersurface of constant time to another. Transformation of all other internal gravitational potentials from the local to global coordinates does not require taking into account relativistic corrections as it exceeds the accuracy of the first post-Newtonian approximation.

The matching procedure also requires to derive explicit transformation of the second time derivative of the potential χ. This transformation can be directly obtained from the definition of the potential given in equation (4.77) and the mapping equation (5.44). After straightforward calculation one gets

$$\frac{\partial\hat{\chi}^{(B)}(u,\boldsymbol{w})}{\partial u^2} = c^2\chi^{(B)}_{,00}(t,\boldsymbol{x}) + a^k_B\chi^{(B)}_{,k}(t,\boldsymbol{x}) + 2cv^k_B\chi^{(B)}_{,0k}(t,\boldsymbol{x}) + v^i_{_B}v^j_B\chi^{(B)}_{,ij}(t,\boldsymbol{x}) + O(\epsilon^2)\,, \tag{5.50}$$

where $\chi^{(B)}_{,0}$ and $\chi^{(B)}_{,i}$ denote the partial derivatives of $\chi^{(B)}$ with respect to the global coordinate time $x^0 = ct$ and space x^i, respectively.

5.2.3.2 *Transformation of the External Potentials*

External potentials in the internal solution of the metric tensor depend in the local coordinates $w^\alpha = (cu, \boldsymbol{w})$ on the external multipole moments $Q_L = Q_L(u)$, $C_L = C_L(u)$, $P_L = P_L(u)$, which are defined at the origin of the local coordinates as functions of time u only. The origin of the local coordinates resides in Figure 5.1 at the point B where the hypersurface of the constant time u intersects with the worldline of the origin. However, functions entering the left side of matching equations (5.3), (5.4), depend on time t. Hence, before performing the post-Newtonian coordinate transformation of the external potentials they must be Lie transported along the worldline of the origin of the local coordinates from the point B to the point A located on the hypersurface of the constant global coordinate time t. Time shift Δt along this worldline is determined by equation (5.36) where one has to associate the point $\boldsymbol{x}'$ with the origin of the local coordinates, that is $\boldsymbol{x}' = \boldsymbol{x}_B$, under condition that the matching point P is taken as defined in previous section. Keeping in mind that the external potentials are scalars with respect to the Lie transport, one obtains

$$Q_L(\mathsf{B}) = Q_L(\mathsf{A}) + \dot{Q}_L(\mathsf{A})\Delta t + O\left(\Delta t^2\right) , \tag{5.51}$$

where the dot over Q_L means differentiation with respect to time t. After making use of equations (5.1) and (5.36), accounting for that the local coordinates of the point B are $w^\alpha(\mathsf{B}) = (cu, 0)$ and the global coordinates of the point A are $x^\alpha(\mathsf{A}) = (ct, \boldsymbol{x}_B(t))$, one gets

$$Q_L(u) = Q_L(t) - \epsilon^2 \dot{Q}_L(t)\left[\xi^0(t, \boldsymbol{x}_B) - \xi^0(t, \boldsymbol{x})\right] + O\left(\epsilon^4\right) . \tag{5.52}$$

Formulas for the multipole moments C_L and P_L are derived similarly.

Now one applies the coordinate transformation of the space coordinates given by equation (5.2) to transform the STF product of the local coordinates. One gets

$$w^{<i_1 i_2 \ldots i_l>} = R_B^{<i_1 i_2 \ldots i_l>} + l\epsilon^2 R_B^{<i_1 i_2 \ldots i_{l-1}} \xi^{i_l>} + O\left(\epsilon^4\right) . \tag{5.53}$$

After combining equations (5.52)-(5.53) together, the post-Newtonian transformation of the Newtonian part of the external potential of the internal solution for the metric tensor assumes the following form

$$\begin{aligned}\sum_{l=0}^{\infty} \frac{1}{l!} Q_L(u) w^L &= \sum_{l=0}^{\infty} \frac{1}{l!} Q_L(t) R_B^L + \epsilon^2\left[\xi^0(t, \boldsymbol{x}) - \xi^0(t, \boldsymbol{x}_B)\right] \sum_{l=0}^{\infty} \frac{1}{l!} \dot{Q}_L(t) R_B^L \\ &+\epsilon^2 \sum_{l=1}^{\infty} \frac{1}{(l-1)!} Q_{kL-1}(t) R_B^{<L-1} \xi^{k>} + O\left(\epsilon^4\right) .\end{aligned} \tag{5.54}$$

This is the most complicated transformation of the external potential that is needed. It takes into account the relativistic nature of the PPN coordinate transformation and is reciprocal to equation (5.42) for the internal Newtonian gravitational potential. All other external potentials present in the local metric (4.171)-(4.173) enter its post-Newtonian part and are transformed like in the Newtonian theory without taking into account the post-Newtonian corrections in the right side of equation (5.53).

5.2.4
Matching for the scalar field

Scalar field appears explicitly only in the post-Newtonian terms. For this reason matching of its asymptotic expansions given in the local and global coordinates is quite straightforward. We operate with the external and internal solutions of the scalar field given by equations (4.68) and (4.135), (4.136) respectively. Matching equation (5.3) reveals that the internal potentials referred to the body B cancel out in its left and right sides due to equation (5.42) while the potential depending on multipoles P_L matches to the Newtonian potential of the external bodies and its derivatives. More specifically, for any index $l \geq 0$ the matching yields

$$P_L = \bar{U}_{,L}(\boldsymbol{x}_B) + O(\epsilon^2)\,, \tag{5.55}$$

where the external Newtonian potential $\bar{U}$ is defined in equation (5.32) and is taken at the origin of the local coordinates, that is at the point $x^i = x^i_B(t)$ and at the instant of time t. Thus, for each index l the STF scalar multipole P_L of the external bodies is defined as the l-th spatial derivative of the Newtonian gravitational potential $\bar{U}$. One reminds that according to equation (4.22 the scalar field was normalized to the factor $\gamma - 1$, where $\omega + 2 = 1/(1-\gamma)$, so that physically observed scalar field $\zeta = (1-\gamma)\varphi$. It vanishes in general relativity, where the multipole moments $P_L = 0$ playing no role.

5.2.5
Matching for the metric tensor

5.2.5.1 *Matching $g_{00}(t,\boldsymbol{x})$ and $\hat{g}_{\alpha\beta}(u,\boldsymbol{w})$ in the Newtonian approximation*

Matching equation (5.4) with $g_{00}(t,\boldsymbol{x})$ component of the metric tensor standing in its left side has a Newtonian limit. In the Newtonian approximation the equation (5.4) reads

$$\hat{N}(u,\boldsymbol{w}) = N(t,\boldsymbol{x}) + 2\mathfrak{B}(t,\boldsymbol{x}) - v_B^2 + O(\epsilon^2)\,. \tag{5.56}$$

Function $\mathfrak{B}(t,\boldsymbol{x})$ is taken from equation (5.22) while components $N(t,\boldsymbol{x})$ and $N(u,\boldsymbol{w})$ of the metric tensors in the global and local coordinates are taken from equations (4.69) and (4.171) respectively. One finds that after making use of equation (5.42) the internal gravitational potentials $\hat{U}^{(B)}(u,\boldsymbol{w})$ and $U^{(B)}(t,\boldsymbol{x})$ are equal and, hence, drop out of the left and right sides of equation (5.56). Expanding the external gravitational potential $\bar{U}(t,\boldsymbol{x})$ in a Taylor series around the origin of the local coordinates, x^i_B, and equating similar terms in polynomials having the same power of R^i_E, yields a set of matching equations for the external multipoles Q_L and functions entering the coordinate transformation. In the Newtonian approximation the set of the matching equations reads

$$Q + \dot{\mathcal{A}} = \frac{1}{2}v_B^2 + \bar{U}(\boldsymbol{x}_B) + O(\epsilon^2)\,, \tag{5.57}$$

$$Q_i = \bar{U}_{,i}(\boldsymbol{x}_B) - a^i_B + O(\epsilon^2)\,, \tag{5.58}$$

$$Q_L = \bar{U}_{,L}(\boldsymbol{x}_B) + O(\epsilon^2), \qquad (l \geq 2)\,. \tag{5.59}$$

Equation (5.57) makes it evident that function $\mathcal{A} = \mathcal{A}(t)$ and the external monopole $Q = Q(t)$ can not be determined from the matching procedure separately

- only their linear combination $Q + \dot{\mathcal{A}}$ can be determined. Hence, either Q or $\mathcal{A}$ can be chosen arbitrary. The most preferable choice is to take $Q = 0$ as it simplifies further equations. This choice was adopted, for example, in papers [Damour *et al.*, 1991; Kopejkin, 1988a; Thorne and Hartle, 1985]. It is also consistent with the recommendation of the IAU [Soffel *et al.*, 2003,?], which makes $\hat{g}_{00}(u, \boldsymbol{w})$ component of the local metric tensor equal to -1 at the origin of the local coordinates if gravitational field of the Earth is neglected. However, if one chooses $Q = 0$ the rate of the coordinate time u can be different from that of the coordinate time t because the average value of functions v_B^2 and $\bar{U}$ is not zero for the circular orbit [Murray, 1983]. Hence, the choice of $Q = 0$ can be inconvenient for astronomical data reduction codes and for calculation of ephemerides in the solar system. Therefore, two time scales, $TDB = k_B t$ and $TDT = k_E u$, have been introduced in such a way that their rate at the origin of the local coordinate system is the same [Brumberg and Kopeikin, 1990; Irwin and Fukushima, 1999; Soffel *et al.*, 2003,?]. This makes function

$$Q(t) = - < \dot{\mathcal{A}} >= a + bt + ct^2 + \dots \,, \tag{5.60}$$

that is a polynomial of time with numerical coefficients calculated by means of numerical integration of equation (5.57) over sufficiently long interval of time [Brumberg and Kopeikin, 1990; Irwin and Fukushima, 1999]. Time-rate adjustment coefficients k_B and k_E relate to each other as [Brumberg and Kopeikin, 1990; Irwin and Fukushima, 1999]

$$\frac{k_B}{k_E} = 1 + c^{-2} Q(t_{\rm epoch}) \,, \tag{5.61}$$

where $t_{\rm epoch}$ is the astronomical epoch, let say J2000, chosen by convention.

In accordance with the interpretation given by Misner *et al.* [1973]; Thorne and Hartle [1985], function Q_i from (5.58) must be understood as an inertial acceleration measured by accelerometer being at rest at the origin of the local coordinates under condition that the internal gravitational field of the body B is neglected. The choice, $Q_i = 0$, (see, e.g., [Ashby and Bertotti, 1984; Thorne and Hartle, 1985]) leads to construction of a freely-falling, local coordinate frame, which origin moves along a geodesic worldline on the background spacetime manifold defined by the gravitational potentials of all celestial bodies of the N-body system but the body B. Such choice of Q_i, however, disengage the worldline of the center of mass of the body B from that of the origin of the local coordinates (see Figure 5.2). Indeed, the center of mass of the body B does not move along the geodesic worldline due to the interaction of its internal multipole moments with the tidal gravitational field of external bodies [Kopejkin, 1988a, 1991a,b]. For this reason, a special procedure must be applied for finding function Q_i such that it will make the origin of the local coordinates following the same worldline as the center of mass of the body B. This procedure is described later in section 6.1.4 in full details.

STF external multipoles Q_L ($l \geq 2$) are defined in the Newtonian approximation by equation (5.59). They describe gravitoelectric [Mashhoon, 1985; Mashhoon and Theiss, 2001] tidal field of external bodies at the origin of the local coordinates. Post-Newtonian corrections to the Newtonian value of the multipoles Q_L can be important

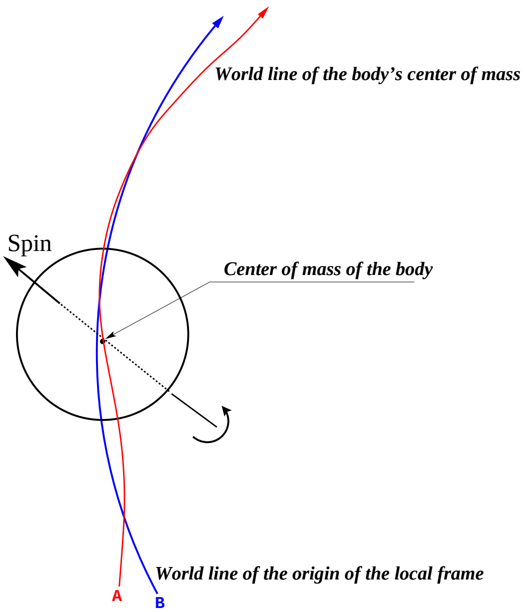

Figure 5.2 Worldlines of the origins of two local coordinates are shown. The curve (A) indicates the worldline of the origin of the local coordinates $\mathcal{A}$ falling freely in the background spacetime described by the external metric tensor $g^{(E)}_{\alpha\beta}$. The curve (B) depicts motion of the origin of the local coordinates $\mathcal{B}$ always located at the center of mass of the body under consideration. In general, the two worldlines do not coincide due to the existence of the inertial force, $\mathcal{M}Q_i$, in the local coordinate system $\mathcal{B}$. This force arises due to the gravitational coupling between the internal multipole moments of the body with the external gravitoelectric and gravitomagnetic tidal fields [Mashhoon, 1985; Mashhoon and Theiss, 2001] of the background metric produced by bodies of the N-body system with the body B excluded.

for lunar laser ranging and other experimental tests of general relativity conducted in the near-Earth space. The post-Newtonian corrections to the external multipoles can be also important in construction of the relativistic theory of accretion disc around a star in a close binary system [Cherepashchuk, 2003]. These corrections are derived later and shown explicitly in equation (5.89).

5.2.5.2 Matching $g_{ij}(t,\boldsymbol{x})$ and $\hat{g}_{\alpha\beta}(u,\boldsymbol{w})$.

One substitutes $g_{ij}(t,\boldsymbol{x})$ component of the metric tensor in the global coordinates to the left side of the matching equation (5.4) and expand this equation by taking into account all post-Newtonian terms of order $O(\epsilon^2)$. One finds that in this approximation the matching equation reads

$$\hat{H}_{ij}(u,\boldsymbol{w}) = H_{ij}(t,\boldsymbol{x}) - \Re^i{}_j(t,\boldsymbol{x}) - \Re^j{}_i(t,\boldsymbol{x}) + v^i_B v^j_B + O(\epsilon^2)\,, \tag{5.62}$$

where the matrix of transformation $\Re^i{}_j(t,\boldsymbol{x})$ from the local, w^α, to global, x^α, coordinates is given in equation (5.26) whereas components of the metric tensor are defined by equation (4.72) in the global coordinates and by equation (4.173) in the local coordinates. Comparing similar terms in both sides of the matching equation, one finds that the internal gravitational potentials of the central body B, $\hat{U}^{(B)}(u,\boldsymbol{w})$ and $U^{(B)}(t,\boldsymbol{x})$, entering the left and right sides of the matching equation, cancel each other. The remaining terms descibe the background gravitational field of external bodies and their matching gives the following set of equations:

$$Y + \mathcal{E} = \frac{1}{6}v^2_B + \gamma\bar{U}(\boldsymbol{x}_B) + O(\epsilon^2)\,, \tag{5.63}$$

$$\mathcal{E}_i = a^i_B + O(\epsilon^2)\,, \tag{5.64}$$

$$\mathcal{E}_L = O(\epsilon^2) \qquad (l \ge 2)\,, \tag{5.65}$$

$$\mathcal{D}_{ij} = \frac{1}{2}v^{<i}_B v^{j>}_B + O(\epsilon^2)\,, \tag{5.66}$$

$$\mathcal{D}_L = O(\epsilon^2) \qquad (l \ge 3)\,, \tag{5.67}$$

$$\mathcal{F}_L = O(\epsilon^2) \qquad (l \ge 2)\,, \tag{5.68}$$

which define the structure of the PPN transformation between spatial coordinates of the global and local frames.

Function $Y = Y(t)$ defines the unit of measurement of spatial distance in the local coordinates. It is convenient from theoretical point of view to chose $Y = 0$ as it was done, for example, in [Kopejkin, 1988a; Thorne and Hartle, 1985]. However, introduction of TDB and TDT time scales in ephemeris astronomy forces us to make $Y \neq 0$ to compensate the effect of function $Q \neq 0$ from equation (5.57) on equations of motion [Brumberg and Kopeikin, 1990; Soffel and Brumberg, 1991]. Specific choice of function Y is a matter of pure practical application of the PPN theory of reference frames in the solar system. Standard decision is to make $Y = Q$ [Brumberg and Kopeikin, 1990; Klioner *et al.*, 2009a] to preserve the canonical form of the Newtonian law of gravity. One does not impose any restriction on function Y at this step of the matching procedure. It will allow us to trace how functions Q and Y affect equations of motion of extended bodies not only in the Newtonian but also in the post-Newtonian approximation.

5.2.5.3 *Matching $g_{0i}(t,x)$ and $\hat{g}_{\alpha\beta}(u,w)$.*

Matching condition (5.4) for $g_{0i}(t,\boldsymbol{x})$ has the following explicit form:

$$\hat{L}_i(u,\boldsymbol{w}) = L_i(t,\boldsymbol{x}) + \mathfrak{B}_i(t,\boldsymbol{x}) - \mathfrak{P}_i(t,\boldsymbol{x}) + v^i_B\Big[\mathfrak{B}(t,\boldsymbol{x}) + N(t,\boldsymbol{x})\Big] \tag{5.69}$$
$$+v^j_B\Big[H_{ij}(t,\boldsymbol{x}) - \mathfrak{R}^i{}_j(t,\boldsymbol{x})\Big] + O(\epsilon^2)\,,$$

where one has employed equations (5.56) and (5.62) for making simplifications. The metric tensor given in the global coordinates by equations (4.69), (4.71) and (4.72), and in the local coordinates by equations (4.140) and (4.161) is substituted to equation (5.69) where one also uses equations (5.29) and (5.43) for vector-potentials $U_i(t,\boldsymbol{x})$ and $\hat{U}^{(B)}_i(u,\boldsymbol{w})$ respectively. Examination of the final form of equation (5.69) reveals that all gravitational potentials depending on the internal structure of the body B (Earth) cancel out. Remaining terms in the matching equation (5.69) depend only on the background values of the gravitational potentials of external bodies. They yield a number of relationships which allow us to express the external multipole moments entering the metric tensor in the local coordinates and functions in the PPN coordinate transformations (5.12)–(5.15) in terms of the gravitational potentials of the external bodies. These relationships are as follows

$$\mathcal{B}_i = 2(1+\gamma)\bar{U}^i(\boldsymbol{x}_B) - (1+2\gamma)v^i_B\bar{U}(\boldsymbol{x}_B) - \frac{1}{2}v^i_B v^2_B - v^i_B Q + O(\epsilon^2)\,, \tag{5.70}$$

$$\mathcal{B}_{<ik>} = 2(1+\gamma)\left[\bar{U}^{<i,k>}(\boldsymbol{x}_B) - v^{<i}_B\bar{U}^{,k>}(\boldsymbol{x}_B)\right] + 2v^{<i}_B a^{k>}_B + Z_{<ik>} + O(\epsilon^2)\,, \tag{5.71}$$

$$\mathcal{B}_{<iL>} = 2(1+\gamma)\left[\bar{U}^{<i,L>}(\boldsymbol{x}_B) - v^{<i}_B\bar{U}^{,L>}(\boldsymbol{x}_B)\right] + Z_{iL} + O(\epsilon^2)\,, \qquad (l\geq 2), \tag{5.72}$$

and

$$\varepsilon_{ipk}\left(C_p + \dot{\mathcal{F}}_p\right) = -2(1+\gamma)\bar{U}^{[i,k]}(\boldsymbol{x}_B) + (1+2\gamma)v^{[i}_B\bar{U}^{,k]}(\boldsymbol{x}_B) + v^{[i}_B Q^{k]} + O(\epsilon^2)\,, \tag{5.73}$$

$$\varepsilon_{ipj}C_{pL-1} = \frac{4l(1+\gamma)}{l+1}\times \tag{5.74}$$
$$\times\left[v^{[i}_B\bar{U}^{,j]L-1}(\boldsymbol{x}_B) - \bar{U}^{[i,j]L-1}(\boldsymbol{x}_B) - \frac{1}{l}\sum_{k=1}^{l-1}\delta^{a_k[i}\dot{\bar{U}}^{,j]L-2}(\boldsymbol{x}_B)\right] + O(\epsilon^2)\,.$$

As one can see the matching equation (5.69) gives rise to two groups of equations, namely, equations (5.70)-(5.72), and equations (5.73 - (5.74), which have different properties of symmetry with respect to their spatial indices. More specifically, the group of equations (5.70)-(5.72) depend only on STF tensors, which are symmetric and trace-free with respect to the entire set of indices $\{i, a_1, a_2, ..., a_l\}$. On the other hand, equations (5.73), (5.74) are symmetric with respect to the set of indices $\{a_1, a_2, ..., a_l\}$ but they are anti-symmetric with respect to any pair of indices composed of index i and any index from the set $\{a_1, a_2, ..., a_l\}$. This separation of the matching equation (5.69) in two groups is due to the fact that this equation has a free (vector-type) index i and can be split in fully symmetric and antisymmetric parts in accordance with the Clebsch-Gordan decomposition of a vector field into its irreducible representations (see [Blanchet and Damour, 1986; Gelfand *et al.*, 1958; Thorne, 1980] for more details).

Let us now discuss a freedom in choosing the angular speed of rotation of spatial axes of the local coordinates. From the discussion in previous sections one knows that there are two possible definitions of the rotation of the local coordinates - either dynamical or kinematical. Dynamically non-rotating local coordinates are defined by the condition that the external dipole moment $C_i = 0$ in equation (4.161). On the other hand, the kinematically non-rotating local coordinates are obtained if one chooses function $\mathcal{F}_i = 0$ in equation (5.15) describing the coordinate transformation of spatial axes. If one does not impose any restriction on the choice of C_i, then, equation (5.73) can be rewritten as

$$\dot{\mathcal{F}}_i = \varepsilon_{ijk}\left[(1+\gamma)\bar{U}^{j,k}(\boldsymbol{x}_B) - (\gamma + \frac{1}{2})v_B^j \bar{U}^{,k}(\boldsymbol{x}_B) - \frac{1}{2}v_B^j Q^k\right] - C_i + O(\epsilon^2). \tag{5.75}$$

The first term in equation (5.75) describes the Lense-Thirring gravitomagnetic precession, the second term describes the de Sitter precession, and the third term describes the Thomas precession of the spatial axes of the local coordinates with respect to the global ones. We emphasize that in the scalar-tensor theory of gravity both the Lense-Thirring and the de Sitter precessions depend on the PPN parameter γ while the Thomas precession does not. The reason is that the Thomas precession is a special relativistic effect [Misner *et al.*, 1973] and can not depend on a particular choice of a specific gravitational theory. If one chooses function $C_i = 0$, the spatial axes of the local coordinates will rotate kinematically with respect to the global coordinates. On the other hand, the choice $\dot{\mathcal{F}}_i = 0$ makes the local coordinates kinematically non-rotating with respect to the global coordinates. The spatial axes of the kinematically non-rotating coordinates are fixed with respect to distant quasars with negligibly small proper motions. These quasars form the International Celestial Reference Frame (ICRF) [Johnston *et al.*, 2000].

Functions $\mathcal{B}_L$ enter the PPN time coordinate transformation as seen from equation (5.14), and couples with the gauge-dependent multipole moments Z_L, which can be chosen arbitrary. There are two, the most preferable options:

1) One simplifies the structure of time transformation described by equation (5.12) as much as possible. In this case the gauge multipoles Z_L must chosen such that

functions $\mathcal{B}_L$ in the time transformation would assume the most simple form

$$\mathcal{B}_{<ik>} = 2v_B^{<i}a_B^{k>} + O(\epsilon^2) \tag{5.76}$$

$$\mathcal{B}_{<iL>} = O(\epsilon^2)\,, \qquad (l \geq 2)\,. \tag{5.77}$$

Here the choice of $\mathcal{B}_{ij}$ in equation (5.76) eliminates all terms depending explicitly on the velocity of the origin of the local coordinates v_B^i. Substituting equations (5.76) and (5.77) into equations (5.71) and (5.72) yield in this case

$$Z_{iL} = -2(1+\gamma)\left[\bar{U}^{<i,L>}(\boldsymbol{x}_B) - v_B^{<i}\bar{U}^{,L>}(\boldsymbol{x}_B)\right] + O(\epsilon^2)\,, \qquad (l \geq 1). \tag{5.78}$$

This makes the metric tensor and the scalar field in the local coordinates determined by four sets of the independent external multipoles P_L, Q_L, C_L, and Z_L. However, the multipole moments Z_L are not physically meaningful and describe the gauge-dependent (unobservable) coordinate effects.

2) One removes from the metric tensor in the local coordinates all physically meaningless multipole moments Z_L. The cost of this choice is a more complicated form of the time transformation, equation (5.12), that contains now an infinite number of coefficients $\mathcal{B}_L$. Assuming that all $Z_{iL} = 0$ for $l \geq 1$ one obtains from equations (5.71) and (5.72)

$$\mathcal{B}_{<ik>} = 2(1+\gamma)\left[\bar{U}^{<i,k>}(\boldsymbol{x}_B) - v_B^{<i}\bar{U}^{,k>}(\boldsymbol{x}_B)\right] + 2v_B^{<i}a_B^{k>} + O(\epsilon^2)\,, \tag{5.79}$$

$$\mathcal{B}_{<iL>} = 2(1+\gamma)\left[\bar{U}^{<i,L>}(\boldsymbol{x}_B) - v_B^{<i}\bar{U}^{,L>}(\boldsymbol{x}_B)\right] + O(\epsilon^2)\,, \qquad (l \geq 2). \tag{5.80}$$

At the present step of the matching procedure one prefers to keep the multipole moments Z_L unspecified. This theoretically-flexible choice gives us freedom in making the residual gauge transformations for various practical purposes.

5.2.5.4 ***Matching $g_{00}(t,\boldsymbol{x})$ and $\hat{g}_{\alpha\beta}(u,\boldsymbol{w})$ in the post-Newtonian approximation.***

Matching of the metric tensor at the post-Newtonian order of $O(\epsilon^4)$ allows us to infer the post-Newtonian equations of motion of the origin of the local coordinates as well as the post-Newtonian corrections to the external multipole moments Q_L and the coefficients of the post-Newtonian part of equation (5.14) describing transformation between the coordinate times u and t of the local and global coordinates. Expansion of the matching equation (5.4) for the case of the metric tensor component, $g_{00}(t,\boldsymbol{x})$, up to the post-Newtonian order generalizes equation (5.56) as follows

$$\begin{aligned}\hat{N}(u,\boldsymbol{w}) + \epsilon^2\hat{L}(u,\boldsymbol{w}) &= N(t,\boldsymbol{x}) + 2\mathfrak{B}(t,\boldsymbol{x}) - v_B^2 \\ &+\epsilon^2 \Big[L(t,\boldsymbol{x}) - 3\mathfrak{B}^2(t,\boldsymbol{x}) + 2\mathfrak{D}(t,\boldsymbol{x}) - 2\mathfrak{B}(t,\boldsymbol{x})N(t,\boldsymbol{x}) \\ &+ 4v_B^2\mathfrak{B}(t,\boldsymbol{x}) + 2v_B^2 N(t,\boldsymbol{x}) + 2v_B^i L_i(t,\boldsymbol{x}) + 2v_B^i\mathfrak{B}_i(t,\boldsymbol{x}) \\ &- \frac{2}{3}v_B^2\mathfrak{R}^k_k(t,\boldsymbol{x}) + 2v_B^i v_B^j\mathfrak{R}^i_j(t,\boldsymbol{x}) - \frac{5}{3}v_B^4 + \frac{1}{3}v_B^2 H(t,\boldsymbol{x})\Big] + O(\epsilon^3)\,,\end{aligned} \tag{5.81}$$

where the gravitational potentials in the right side of this equations are determined by equations (4.69)-(4.72) and those in the left side are given by equations (4.137), (4.138), (4.171), (4.174) and (4.175). Solution of equation (5.81) is done in several steps.

First of all, one substitutes the components of the transformation matrix of the coordinate bases given by equations (5.22)-(5.26), to equation (5.81). Then, one analyzes terms depending separately on the internal gravitational potentials referred to the body B and those referred to the external bodies. The internal gravitational potentials should be transformed from the local to global coordinates by making use of equations (5.42)-(5.50). One notices that the transformation equation (5.49) for the Newtonian gravitational potential of the body B can be written explicitly in terms of the functions coming about from the matching procedure at lower orders. Taking definitions of the gauge functions ξ^0 and ξ^i from section 5.1.2 and substituting to them functions determined at previous steps of the matching procedure, one obtains explicit form of the relativistic correction, $\mathcal{U}^{(B)}(t, \boldsymbol{x})$, from equation (5.49), which describes the post-Newtonian part of the transformation of the Newtonian potential from the local to global coordinates. It reads

$$\begin{aligned}\mathcal{U}^{(B)}(t, \boldsymbol{x}) \quad = \quad & U^{(B)}(t, \boldsymbol{x}) \left(\frac{1}{2} v_B^2 - \gamma \bar{U}(\boldsymbol{x}_B) - a_B^k R_B^k + Y \right) \\ & + \frac{1}{2} v_B^i v_B^j \chi^{(B)}_{,ij}(t, \boldsymbol{x}) + c v_B^k \chi^{(B)}_{,0k}(t, \boldsymbol{x}) \\ & - \frac{1}{2} a_B^k \chi^{(B)}_{,k}(t, \boldsymbol{x}) - v_B^k U^k_{(B)}(t, \boldsymbol{x}) + O(\epsilon^2) \,. \end{aligned} \tag{5.82}$$

At the second step, one substitutes formula (5.82) for the internal Newtonian potential of the body B to the matching equation and use transformation equations for other internal potentials. One gets a rather remarkable result - both the Newtonian and post-Newtonian terms depending on the internal structure of the body B cancel out and disappear from the matching equation (5.81). This property of independence of the matching equation on the internal potentials is called the effacing principle [Kopeikin and Vlasov, 2008]. The effacing property of the matching equation can be explained in terms of the laws of conservation of intrinsic linear and angular momenta of the body B which are valid not only in general relativity but in the scalar-tensor theory of gravity as well [Damour and Esposito-Farese, 1992; Will, 1993]. For this reason, the scalar field can produce neither a net self-force nor a self-torque exerted on the body. Calculation given above, proves that the effacing principle is valid in the first post-Newtonian approximation for extended bodies having arbitrary internal structure and shape. This principle was extrapolated to the 2.5 post-Newtonian approximation for spherically-symmetric and weakly-gravitating bodies in papers [Kopeikin, 1985; Kopejkin, 1988b] by applying the Fock-Papapetrou [Fock, 1964; Papapetrou, 1951a] method of derivation of relativistic equations of motion of extended bodies. Damour [1983] proved the effacing principle for spherically-symmetric neutron stars up to the 2.5 post-Newtonian approximation by making use of the Einstein-Infeld-Hoffmann (EIH) technique [Einstein *et al.*, 1938] along with the method of analytic continuation of generalized functions (distributions) [Gelfand *et al.*, 1958], which are used for definition of a singular energy-momentum tensor of point-like particles.

At the third step, one picks up and equates in equation (5.81) the gravitational potentials generated by all bodies being external with respect to the body B (Earth). This step requires that one knows the law of transformation of function $\sum \frac{1}{l!}Q_L(u)w^L$ from the local to global coordinates within the post-Newtonian accuracy. General formula of transformation of this function is given by equation (5.54) but it should be further specified. This is achieved by substituting to this formula the explicit expressions for functions ξ^0 and ξ^i displayed in equations (5.7), (5.12)-(5.15). One gets

$$
\begin{aligned}
\sum_{l=0}^{\infty}\frac{1}{l!}Q_L(u)w^L &= \sum_{l=0}^{\infty}\frac{1}{l!}Q_L(t)R_B^L\Big[1+l\epsilon^2\left(\gamma\bar{U}(\boldsymbol{x}_B)-Y\right)\Big] \qquad (5.83)\\
&+ \epsilon^2\left[-\sum_{l=1}^{\infty}\frac{1}{(l-1)!}Q_{kL-1}F^{jk}R_B^{<jL-1>}+\dot{Q}v_B^kR_B^k\right]\\
&+ \epsilon^2\sum_{l=1}^{\infty}\frac{1}{(l-1)!}\Big[\frac{1}{2}v_B^jv_B^kQ_{jL-1}R_B^{<kL-1>}+\left(a_B^kQ_L-\frac{1}{l}v_B^k\dot{Q}_L\right)R_B^{<kL>}\Big]\\
&+ \epsilon^2\sum_{l=0}^{\infty}\frac{1}{(2l+3)l!}\Big[\frac{1}{2}v_B^jv_B^kQ_{jkL}-\frac{1}{2}a_B^kQ_{kL}-v_B^k\dot{Q}_{kL}\Big]R_B^LR_B^2+O(\epsilon^4)\,.
\end{aligned}
$$

Matching equation (5.81) requires to calculate function $\bar{\chi}_{,00}(t,\boldsymbol{x})$ explicitly. This function is generated by the external bodies and enters $g_{00}(t,\boldsymbol{x})$ component of the metric tensor in the global coordinates as shown in equations (4.70) and (5.30). Contrary to other potentials $\bar{U}(t,\boldsymbol{x})$, $\bar{U}^i(t,\boldsymbol{x})$, $\bar{\Phi}_k(t,\boldsymbol{x})$, which are solutions of the homogeneous Laplace equation in the vicinity of the body B, function $\bar{\chi}(t,\boldsymbol{x})$ is a particular solution of the Poisson equation [Fock, 1964]

$$
\nabla^2\bar{\chi}(t,\boldsymbol{x}) = -2\bar{U}(t,\boldsymbol{x})\,. \qquad (5.84)
$$

After solving this equation and expanding its solution into STF scalar harmonics (see appendix A) one obtains

$$
\bar{\chi}(t,\boldsymbol{x}) = \sum_{l=0}^{\infty}\frac{1}{l!}\bar{\chi}_{,<L>}(\boldsymbol{x}_B)R_B^L-\sum_{l=0}^{\infty}\frac{1}{(2l+3)l!}\bar{U}_{,L}(\boldsymbol{x}_B)R_B^LR_B^2\,. \qquad (5.85)
$$

Differentiating left and right sides of equation (5.85) two times with respect to the global coordinate time, t, yields

$$
\begin{aligned}
\bar{\chi}_{,tt}(t,\boldsymbol{x}) &= \sum_{l=0}^{\infty}\frac{1}{l!}\bar{\chi}_{,tt<L>}(\boldsymbol{x}_B)R_B^L+\sum_{l=0}^{\infty}\frac{1}{(2l+3)l!}\times \qquad (5.86)\\
&\times\Big[a_B^k\bar{U}_{,kL}(\boldsymbol{x}_B)+2v_B^k\dot{\bar{U}}_{,kL}(\boldsymbol{x}_B)-v_B^jv_B^k\bar{U}_{,jkL}(\boldsymbol{x}_B)-\ddot{\bar{U}}_{,L}(\boldsymbol{x}_B)\Big]R_B^LR_B^2\,.
\end{aligned}
$$

Finally, one expands all functions in both sides of equation (5.81) in the Taylor series with respect to the distance R^i_B from the body B, and reduce all similar terms. One finds that those terms which do not depend on R^i_B (that is, functions of time t only) form ordinary differential equation of a first order for function $\mathcal{B}(t)$ which appears in equation (5.12). This differential equation reads

$$\begin{aligned}\dot{\mathcal{B}}(t) &= -\frac{1}{8}v_B^4 - (\gamma + \frac{1}{2})v_B^2\bar{U}(\boldsymbol{x}_B) + (\beta - \frac{1}{2})\bar{U}^2(\boldsymbol{x}_B) + Q\left[\frac{1}{2}v_B^2 + \frac{3}{2}Q - \bar{U}(\boldsymbol{x}_B)\right] \\ &\quad +2(1+\gamma)v_B^k\bar{U}^k(\boldsymbol{x}_B) - \bar{\Phi}(\boldsymbol{x}_B) + \frac{1}{2}\bar{\chi}_{,tt}(\boldsymbol{x}_B) + O(\epsilon^2)\,.\end{aligned} \tag{5.87}$$

Terms, which are linear with respect to R^i_B, give us the post-Newtonian equation of translational motion of the origin of the local coordinates, $x^i_B(t)$, in the global (barycentric) coordinates. It generalizes the Newtonian equation (5.58). The barycentric acceleration of the origin of the local coordinates with respect to the barycenter of the N-body system is

$$\begin{aligned}a^i_B &= \bar{U}_{,i}(\boldsymbol{x}_B) - Q_i + \epsilon^2\Big[F^{ik}Q_k + \bar{\Phi}_{,i}(\boldsymbol{x}_B) - \frac{1}{2}\bar{\chi}_{,itt}(\boldsymbol{x}_B) + Q_i\big(Y - 2Q\big) \\ &\quad +2(1+\gamma)\dot{\bar{U}}^i(\boldsymbol{x}_B) - 2(1+\gamma)v_B^k\bar{U}^{k,i}(\boldsymbol{x}_B) - (1+2\gamma)v_B^i\dot{\bar{U}}(\boldsymbol{x}_B) \\ &\quad +(2-2\beta-\gamma)\bar{U}(\boldsymbol{x}_B)\bar{U}_{,i}(\boldsymbol{x}_B) + (1+\gamma)v_B^2\bar{U}_{,i}(\boldsymbol{x}_B) - \frac{1}{2}v_B^iv_B^k\bar{U}_{,k}(\boldsymbol{x}_B) \\ &\quad -\frac{1}{2}v_B^iv_B^ka_B^k - v_B^2a_B^i - (2+\gamma)a_B^i\bar{U}(\boldsymbol{x}_B)\Big] + O(\epsilon^4)\,.\end{aligned} \tag{5.88}$$

This equation contains the external dipole moment, Q_i, which must be calculated with the post-Newtonian accuracy in order to complete derivation of the post-Newtonian equation of translational motion of the origin of the local coordinates. A simple choice of $Q_i = 0$ does not allow us to keep the origin of the local coordinates at the center of mass of the body B (Earth) for sufficiently long interval of time. This is because the internal moments of the body B interact with the tidal gravitational field of external bodies (Moon, Sun, etc.) so that the body B does not move along a geodesic worldline while the choice of $Q_i = 0$ makes the origin of the local coordinates moving along a geodesic worldline [Kopejkin, 1988a; Thorne and Hartle, 1985]. Thus, function Q_i must be defined in such a way that the body's center of mass and the origin of the local coordinates could coincide at any instant of time. This problem is equivalent to solution of the problem of motion of the body's center of mass with respect to the origin of the local coordinates and will be discussed in the next section.

Terms which are quadratic, cubic, etc., with respect to the distance, R^i_B, determine the post-Newtonian corrections to the external STF multipole moments $Q_L = Q^N_L + \epsilon^2 Q^{pN}_L$, where the Newtonian term Q^N_L is shown in equation (5.59). The post-

Newtonian corrections are

$$
\begin{aligned}
Q_L^{pN}(t) &= X_{<L>} + \dot{Z}_L + \bar{\Phi}_{,<L>}(\boldsymbol{x}_B) - \frac{1}{2}\bar{\chi}_{,tt<L>}(\boldsymbol{x}_B) + \left(lY - 2Q\right)\bar{U}_{,L}(\boldsymbol{x}_B) \\
&+2(1+\gamma)\dot{\bar{U}}^{<i_l,L-1>}(\boldsymbol{x}_B) - 2(1+\gamma)v_B^k\bar{U}^{k,L}(\boldsymbol{x}_B) \\
&+(l-2\gamma-2)v_B^{<i_l}\dot{\bar{U}}^{,L-1>}(\boldsymbol{x}_B) + (1+\gamma)v_B^2\bar{U}_{,L}(\boldsymbol{x}_B) \\
&-\frac{l}{2}v_B^k v_B^{<i_l}\bar{U}^{,L-1>k}(\boldsymbol{x}_B) + (2-2\beta-l\gamma)\bar{U}(\boldsymbol{x}_B)\bar{U}_{,L}(\boldsymbol{x}_B) \\
&-(l^2-l+2+2\gamma)a_B^{<i_l}\bar{U}^{,L-1>}(\boldsymbol{x}_B) - lF^{k<i_l}\bar{U}^{,L-1>k}(\boldsymbol{x}_B)\,, \qquad (l \geq 2)
\end{aligned} \tag{5.89}
$$

where one has used notations

$$X_{<ij>} = 3a_B^{<i}a_B^{j>}\,, \tag{5.90}$$

$$X_{<L>} = 0\,, \qquad (l \geq 3). \tag{5.91}$$

These equations finalize the description of the STF multipole moments entering external solution of the metric tensor in the local coordinates $w^\alpha = (cu, \boldsymbol{w})$ and the parameterized post-Newtonian transformation between the local and global coordinates.

5.2.6
Final form of the PPN coordinate transformation

For the sake of convenience one summarizes the final form of the parameterized post-Newtonian coordinate transformation from the local to global coordinates. It is given by two equations, which are

$$
\begin{aligned}
u &= t - \epsilon^2\left(\mathcal{A} + v_B^k R_B^k\right) \\
&+ \epsilon^4\left[\mathcal{B} + \left(\frac{1}{3}v_B^k a_B^k - \frac{1}{6}\dot{\bar{U}}(\boldsymbol{x}_B) + \frac{1}{6}\dot{Q}\right)R_B^2 - \frac{1}{10}\dot{a}_B^k R_B^k R_B^2 + \sum_{l=1}^{\infty}\frac{1}{l!}\mathcal{B}_{<L>}R_B^L\right] + O(\epsilon^5)\,,
\end{aligned} \tag{5.92}
$$

$$w^i = R_B^i + \epsilon^2\left[\left(\frac{1}{2}v_B^i v_B^k + \gamma\delta^{ik}\bar{U}(\boldsymbol{x}_B)\delta^{ik}Y + F^{ik}\right)R_B^k + a_B^k R_B^i R_B^k - \frac{1}{2}a_B^i R_B^2\right] + O(\epsilon^4) \tag{5.93}$$

Here functions $\mathcal{A}$ and $\mathcal{B}$ are solutions of the ordinary differential equations

$$\dot{\mathcal{A}} = \frac{1}{2}v_B^2 + \bar{U}(\boldsymbol{x}_B) - Q\,, \tag{5.94}$$

$$
\begin{aligned}
\dot{\mathcal{B}} &= -\frac{1}{8}v_B^4 - (\gamma + \frac{1}{2})v_B^2\bar{U}(\boldsymbol{x}_B) + Q\Big[-\frac{1}{2}v_B^2 + \frac{3}{2}Q - \bar{U}(\boldsymbol{x}_B)\Big] \\
&+ (\beta - \frac{1}{2})\bar{U}^2(\boldsymbol{x}_B) + 2(1+\gamma)v_B^k\bar{U}^k(\boldsymbol{x}_B) - \bar{\Phi}(\boldsymbol{x}_B) + \frac{1}{2}\bar{\chi}_{,tt}(\boldsymbol{x}_B)\,,
\end{aligned} \tag{5.95}
$$

while the other functions are defined as follows

$$\mathcal{B}_i = 2(1+\gamma)\bar{U}^i(\boldsymbol{x}_B) - (1+2\gamma)v_B^i\bar{U}(\boldsymbol{x}_B) - \frac{1}{2}v_B^i v_B^2 - v_B^i Q\,, \tag{5.96}$$

$$\mathcal{B}_{<ik>} = Z_{ik} + 2(1+\gamma)\bar{U}^{<i,k>}(\boldsymbol{x}_B) - 2(1+\gamma)v_B^{<i}\bar{U}^{,k>}(\boldsymbol{x}_B) + 2v_B^{<i}a_B^{k>}\,, \tag{5.97}$$

$$\mathcal{B}_{<iL>} = Z_{iL} + 2(1+\gamma)\bar{U}^{<i,L>}(\boldsymbol{x}_B) - 2(1+\gamma)v_B^{<i}\bar{U}^{,L>}(\boldsymbol{x}_B)\,, \qquad (l \geq 2), \tag{5.98}$$

$$\dot{F}^{ik} = (1+2\gamma)v_B^{[i}\bar{U}^{,k]}(\boldsymbol{x}_B) - 2(1+\gamma)\bar{U}^{[i,k]}(\boldsymbol{x}_B) + v_B^{[i}Q^{k]}\,. \tag{5.99}$$

These equations will be used in subsequent sections for derivation of the post-Newtonian equations of motion of extended bodies. They are also convenient for comparison with the relativistic transformations derived by other researchers [Ashby and Bertotti, 1986; Brumberg and Kopejkin, 1989a; Damour *et al.*, 1991, 1992; Fukushima, 1988; Klioner and Soffel, 2000; Kopejkin, 1988a; Soffel *et al.*, 2003].

6
Relativistic Celestial Mechanics

6.1
Post-Newtonian Equations of Orbital Motion

6.1.1
Introduction

In the Newtonian theory of gravity the definitions of mass and the center of mass of an extended body that is a member of N-body system, are fairly straightforward concepts. These definitions can be extrapolated without any change to the relativistic theory of gravity as proposed by Fock [Fock, 1964]. He used the Newtonian definitions with a corresponding replacement of matter's density ρ with the invariant density ρ^* which obeys the exact, Newtonian-like equation of continuity (4.44) at any order of the post-Newtonian approximations. The invariant density is defined as the ratio of the unit mass of baryons per unit of the proper volume. Integration of the invariant density over the total volume of the body gives its total baryonic mass, which is constant asfollows from equation (1.35). The invariant density is also useful to introduce the Newtonian definitions of the body's center of mass, and its linear momentum. The baryonic mass, the center of mass, and the linear momentum of body B are given [1] by integrals [Brumberg, 1972; Fock, 1964]

$$\mathcal{M}_* = \int_{V_B} \rho^*(u, \boldsymbol{w}) d^3 w \,, \tag{6.1}$$

$$\mathcal{J}^i_* = \int_{V_B} \rho^*(u, \boldsymbol{w}) w^i d^3 w \,, \tag{6.2}$$

$$\mathcal{P}^i_* = \int_{V_B} \rho^*(u, \boldsymbol{w}) v^i d^3 w \,. \tag{6.3}$$

Notice that the integrals are written down in the local coordinates and the integration is performed over the hypersurface of constant time u. In fact, the value of the integrals does not depend on the choice of the coordinates and the hypersurface of integration because the product of the invariant density and three-dimensional volume of integration is invariant by definition: $\rho^*(u, \boldsymbol{w})d^3w = \rho^*(t, \boldsymbol{x})d^3x$.

Newtonian definitions of the center of mass and the linear momentum of the body were used by Fock [Fock, 1964], Papapetrou [Papapetrou, 1951a], Petrova [Petrova, 1949], Brumberg [Brumberg, 1991], and some other researchers[2] for derivation of the post-Newtonian equations of translational motion of spherically-symmetric bodies. It is reasonable to expect that equations of motion of such bodies would depend only on their masses, as in the case of massive point particles that is called the principle of effacing of the internal structure [Kopeikin and Vlasov, 2008]. This expectation is confirmed in the Newtonian theory, where the motion of spherically-symmetric bodies is indistinguishable from that of point particles. However, the

1) One skips in this section label B for all quantities referred to body B as it does not cause misinterpretation. The bodies will be labeled with indices A, B, C, and so on, whenever confusion could arise.

2) See, e.g., [Caporali, 1981; Contopoulos and Spyrou, 1976; Spyrou, 1975, 1978] and references therein.

earlier works in the relativistic celestial mechanics could not provide proof of the effacing principle beyond the Newtonian approximation because the post-Newtonian equations of motion of the spherically-symmetric bodies, unlike the Newtonian theory, contained not only their baryonic masses, $\mathcal{M}_*$, but also other characteristics such as the internal kinetic and gravitational energy, elastic energy, and even their moments of inertia [Brumberg, 1972, 1991; Caporali, 1981; Contopoulos and Spyrou, 1976; Dallas, 1977; Fock, 1964; Spyrou, 1975, 1978].

Explicit appearance of such terms in the post-Newtonian equations of motion complicates their interpretation and calculation of ephemerides. For this reason, the dependence of the post-Newtonian equations of motion on the parameters characterizing the internal structure of the bodies was silently ignored and researchers used the Einstein-Infeld-Hoffmann (EIH) equations of point-like masses [Einstein *et al.*, 1938], where the internal structure of the bodies is irrelevant. The situation remained unsatisfactory for many years. Some researchers even argued that motion of spherically-symmetric extended bodies must be different from that of compact relativistic objects like neutron stars and black holes even if they have the same masses as the extended bodies. This point of view is clearly incompatible with the Newtonian equations of motion of two black holes as it was shown by Demiański and Grishchuk [1974] who proved, at least in the Newtonian approximation, that equations of motion of black holes are the same as for spherically-symmetric stars made of ordinary matter. D'Eath [1975a,b] extended this result to the post-Newtonian approximation by making use of the method of matched asymptotic expansions.

Einstein envisaged that the internal structure of bodies should not affect their orbital equations of motion. This suggests that motion of two spherically-symmetric bodies having the same mass but different distribution of intrinsic density, pressure, etc., is to be indistinguishable. Physical intuition led Einstein to the development of the EIH method of derivation of the post-Newtonian equations of motion from vacuum Einstein's equations where the bodies are treated as point-like singularities of gravitational field. Since gravity is a pure geometric phenomenon in general relativity, it must tell the spherical bodies how to move, irrespectively of the details of their internal structure. Fock [Fock, 1957] was not convinced with the Einstein's argument because the mathematical nature of the EIH singularities and their relation to real celestial bodies (planets, stars) was not quite clear at that time. In addition, most of astronomical bodies are made of ordinary matter and are represented mathematically by smooth, differentiable functions having a compact support (a finite volume in space). Fock [3] proposed a post-Newtonian method of derivation of relativistic equations of motion by integrating the energy-momentum tensor of continuous matter over the volumes of the extended bodies. Fock's method clarified the functional dependence between the equations of motion of bodies and their internal structure. It does not contradict to EIH equations because all terms depending on the internal structure of the bodies can be absorbed to relativistic mass, center of mass, and other parameters of the bodies after corresponding redefinition of these parameters.

3) Papapetrou [Papapetrou, 1951a,b] developed similar approach independently and approximately at the same time.

Indeed, theoretical study [Grishchuk and Kopeikin, 1986; Kopeikin, 1985] of relativistic orbital motion of two spherically-symmetric compact stars comprising a binary system revealed that the Newtonian definitions of mass, the center of mass and the linear momentum of the bodies given in equations (6.1)-(6.3) are not appropriate for calculating their post- and post-post-Newtonian equations of motion. One has shown [Kopeikin, 1985; Kopejkin, 1988b] that if the Newtonian mass and the center of mass are replaced with the appropriate relativistic definitions, all terms in the equations of motion depending on the internal structure of the bodies are effectively eliminated after applying the procedure of relativistic renormalization of masses. Damour [1983] called this property the *effacing principle* and confirmed it for spherically-symmetric and compact relativistic bodies using the mathematical technique of matched asymptotic expansion.

Newtonian theory predicts that if celestial bodies have no spherically-symmetric distribution of matter, their equations of motion must depend on additional integral parameters, which are the multipole moments of the bodies characterizing irregularities in the distribution of mass inside the bodies. It is natural to expect that the post-Newtonian equations of motion of the bodies with arbitrary internal distribution of density and velocity of matter will depend on more complicated set of the parameters including the mass-type and current-type multipole moments given by equations (4.176) and (4.183) respectively. However, it was not quite clear until recently whether some other parameters would appear in the relativistic equations of motion in addition to these two sets of the internal multipoles. Scrutiny analysis of this question in general relativity elucidated that the post-Newtonian equations of motion of extended bodies with arbitrary internal structure do contain only the mass and current multipoles [Damour *et al.*, 1991, 1992, 1993; Kopeikin and Vlasov, 2004] and does not depend on any other internal characteristic of the bodies.

The principle of effacing of the internal structure of the bodies is not valid in alternative theories of gravity because of the violation of the strong principle of equivalence [Dicke, 1962a; Nordtvedt, 1968a,b; Will, 1993] due to the presence of other long-range gravitational fields. This violation reveals already in the Newtonian equations of motion where two masses - inertial and gravitational - become slightly different. Hence, the motion of spherically-symmetric bodies is characterized by two parameters (inertial mass and gravitational mass) as contrasted to general relativity. One may expect that equations of motion of non-spherically symmetric and rotating bodies in alternative theories of gravity will depend not only on the set of mass-type and current-type multipoles but on some other integral characteristics of the bodies like the moments of inertia. The reader will see that this is indeed true in section 6.3.

Consistent derivation of relativistic equations of motion in alternative theories of gravity follows the general relativistic ideas and require introducing the local coordinates associated with each body. The problems is that the global coordinates of N-body system are inappropriate for giving adequate physical definition of the internal multipole moments of each body in N-body system. Local coordinates should be constructed around each body in order to eliminate the Lorentz and Einstein contractions and other gauge-dependent contributions caused by orbital motion of the body with respect to the global coordinates. Consistent relativistic concept of the local coordinates built around each massive, extended body in N-body system, was developed in papers [Brumberg and Kopejkin, 1989a; Damour *et al.*, 1991; Kopejkin, 1988a; Mashhoon, 1985] and outlined in this book in chapter 4. It was also discussed in papers [Ashby and Bertotti, 1984, 1986; Fukushima, 1988; Shahid-Saless and Ashby, 1988] under restriction that the center of mass of the body moves along a geodesic worldline on the background spacetime formed by the gravitational potentials of the external bodies. This restriction is pretty limited and can not be implemented in the most general case of non-spherical and rotating bodies because their worldlines obey the relativistic Mathisson-Papapetrou-Dixon equation of motion (3.227a) which is not geodesic [Damour *et al.*, 1991; Kopejkin, 1988a].

Another essential part of the post-Newtonian celestial mechanics besides the theory of relativistic reference frames deals with the problem of relativistic definition of the multipole moments of the gravitational field of the whole N-body system and each body separately. Solution of this problem was proposed by Thorne [Thorne, 1980] but his definition of the multipoles suffered from divergencies at higher post-Newtonian approximations. Blanchet and Damour [1989] eliminated the divergencies and introduced a compact-support, post-Newtonian definition of the mass-type multipole moments which are symmetric and trace-free (STF) tensors with respect to the transformation of the rotation group SO(3). Blanchet-Damour (BD) definition was fruitfully employed in a series of papers by Damour, Soffel and Xu (DSX) [Damour *et al.*, 1991, 1992] for deriving the post-Newtonian equations of motion of extended bodies in general relativity. They had also proved that the only parameters entering the general-relativistic equations of motion are the Tolman (relativistic) masses of the bodies along with the STF multipole moments of the bodies.

This section generalizes the DSX equations of motion and derives the parameterized post-Newtonian equations of translational motion of extended bodies in the scalar-tensor theory of gravity. One proves that these equations depend only on the inertial and gravitational masses of the bodies and a set of *active* multipole moments of the bodies defined earlier in section 4.5.3. In case of spherically-symmetric bodies the only parameters present in the equations are the inertial and gravitational masses which are different because of coupling of the scalar field with the intrinsic gravitational potential of the body. This difference between the two masses was discussed by Nordtvedt [1968a,b]. It leads to a specific Nordtvedt effect in the relative motion of the Moon around the Earth [Will, 1993].

The post-Newtonian definition of mass and center of mass of a single body are derived from equations (4.129) and (4.130). They are conserved quantities if the body is isolated. It is reasonable to suggest that the gravitational interaction of the body with other bodies of N-body system violates these conservation laws. Dixon [Dixon, 1979] and Thorne and Hartle [Thorne and Hartle, 1985] showed that the law of conservation of mass is indeed violated if the body is not spherically-symmetric (see next section for more detail). However, the law of conservation of the center of mass and the linear momentum of the body can be retained even in case of the gravitational interaction of the body with the external environment in case of a proper choice of the local coordinates [Kopeikin and Vlasov, 2004; Kopejkin, 1988a].

The scalar-tensor theory of gravity admits three types of multipole moments -

active, conformal and scalar - which can be used for definition of body's mass and its center of mass. These moments were introduced in section 4.5.3 and are defined by equations (4.106), (4.113) and (4.119). The active and scalar dipole moments of the body are not appropriate for definition of its center of mass and derivation of its orbital equations of motion because they violate the third Newton's law and lead to self-acceleration of the body which is physically unacceptable. The conformal dipole moment does not have this drawback and is used in the scalar-tensor theory of gravity for definition of the body's center of mass and derivation of equations of motion of the body. Self-acceleration terms do not appear in the equations, which right side has a remarkably simple structure of the post-Newtonian force that is a function of the *active* multipole moments of the body coupled with the tidal gravitational field [Kopeikin and Vlasov, 2004].

Let us discuss the derivation and specific properties of orbital equations of motion of the bodies in more detail starting from the explicit form of the local (macroscopic) equations of motion of matter in local coordinates.

6.1.2
Macroscopic post-Newtonian equations of motion

Macroscopic post-Newtonian equations of motion of matter consist of three groups: (1) the equation of continuity, (2) the thermodynamic equation relating the elastic energy Π with the tensor of stresses $\pi_{\alpha\beta}$, and (3) the Navier-Stokes equation describing dependence of acceleration of the macroscopic element of matter on external gravitational and internal forces caused by the stresses.

The equation of continuity in the local coordinates $(u, \boldsymbol{w})$ has the most simple form in terms of the invariant density ρ^*, and reads

$$\frac{\partial \rho^*}{\partial u} + \frac{\partial \left(\rho^* \nu^i\right)}{\partial w^i} = 0 . \tag{6.4}$$

This equation is exact, that is takes into account all post-Newtonian corrections as follows from the definition of the invariant density ρ^* and equation (4.44).

The thermodynamic equation relating the internal elastic energy Π and the internal tensor of stresses $\pi_{\alpha\beta}$ is required in the post-Newtonian calculations only in a linear order where the tensor of stresses is completely characterized by its spatial components π_{ij}. Expanding the exact equation (4.12) one obtains

$$\frac{d\Pi}{du} + \frac{\pi_{ij}}{\rho^*}\frac{\partial \nu^i}{\partial w^j} = O(\epsilon^2) , \tag{6.5}$$

where the operator of convective time derivative $d/du \equiv \partial/\partial u + \nu^i \partial/\partial w^i$.

The Navier-Stokes equation follows from the spatial part of the covariant law of conservation of the energy-momentum tensor $\nabla_\nu T^{i\nu} = 0$. It yields

$$\rho^* \frac{d}{du}\left\{v^i + \epsilon^2\left[\left(\frac{1}{2}v^2 + \Pi + \frac{1}{2}\hat{N} + \frac{1}{3}\hat{H}\right)v^i + \hat{L}_i\right]\right\} + \epsilon^2 \frac{\partial\left(\pi_{ij}v^j\right)}{\partial u} = \tag{6.6}$$

$$\frac{1}{2}\rho^* \frac{\partial \hat{N}}{\partial w^i} - \frac{\partial \pi_{ij}}{\partial w^j} + \epsilon^2\left\{\rho^*\left[\frac{1}{2}\frac{\partial \hat{L}}{\partial w^i} + \frac{1}{4}\left(v^2 + 2\Pi + \hat{N}\right)\frac{\partial \hat{N}}{\partial w^i} + \frac{1}{6}v^2\frac{\partial \hat{H}}{\partial w^i} + v^k\frac{\partial \hat{L}_k}{\partial w^i}\right]\right.$$

$$\left. + \frac{1}{6}\pi_{kk}\frac{\partial \hat{H}}{\partial w^i} + \frac{1}{2}\pi_{ik}\left(\frac{\partial \hat{N}}{\partial w^k} - \frac{1}{3}\frac{\partial \hat{H}}{\partial w^k}\right) + \frac{1}{2}\left(\hat{N} - \frac{1}{3}\hat{H}\right)\frac{\partial \pi_{ik}}{\partial w^k}\right\} + O(\epsilon^4) ,$$

where gravitational potentials $\hat{N}$, $\hat{H}$, $\hat{L}$, $\hat{L}_i$ are components of the metric tensor in the local coordinates as defined in section 4.5.2.

6.1.3
Mass and the linear momentum of a self-gravitating body

There are two algebraically independent definitions of the post-Newtonian mass in the scalar-tensor theory - the *active* mass and the *conformal* mass, which come out of equations (4.176) and (4.180) respectively, when index $l = 0$. One has pointed out in section 4.5.3 that if the body under consideration is a member of N-body system, the contribution of gravitational potential of the external bodies should be included to the definition of the STF multipole moments of the body. It allows us to cancel out in equations of motion all terms depending on the internal structure of the body, which appear in the intermediate calculations but can not be absorbed to the post-Newtonian definition of the STF multipoles. Disappearance of such terms in equations of motion extrapolates applicability of the effacing principle from general theory of relativity to the scalar-tensor theory of gravity, at least, in the first post-Newtonian approximation. Thorne and Hartle [Thorne and Hartle, 1985] discussed the contribution of the gravitational field of external bodies in the definition of the body's internal multipoles but they did not come up with a definite answer. Our resolution of this question is based on direct calculation of equations of motion with several possible definitions of the center of mass and the STF multipoles. Cumbersome and tedious calculations [Kopeikin and Vlasov, 2004] reveal that the equations of motion have the most simple form and admit the minimal set of parameters, if and only if, one takes the *conformal* definition of mass and the center of mass for each body and include the gravitational potential of external bodies to the definition of the body's multipole moments.

Conformal multipoles are defined in equation (4.180) and, when index $l = 0$, can be reduced to simpler form yielding the *conformal* mass of the body

$$\tilde{\mathcal{M}} = \mathrm{M} - \epsilon^2\left\{[Y + (1-\gamma)P]\,\mathrm{M} + \sum_{l=1}^{\infty}\frac{l+1}{l!}Q_L\mathcal{I}^L\right\} + O(\epsilon^4) , \tag{6.7}$$

where

$$\mathrm{M} = \int_{V_B}\rho^*\left[1 + \epsilon^2\left(\frac{1}{2}v^2 + \Pi - \frac{1}{2}\hat{U}^{(B)}\right)\right]d^3w + O(\epsilon^4) , \tag{6.8}$$

is *general relativistic* definition of the post-Newtonian mass of the body [Will, 1993]. Gravitationally-isolated, single body has only mass M. Gravitational interaction with other bodies of the N-body system brings about additional contribution to the relativistic mass of the body which is described by the last term in the right side of equation (6.7).

Discussion of the strong principle of equivalence requires definition of the *active* mass of the body as well. It is obtained from equation (4.176), when index $l = 0$, and reads

$$\begin{aligned} \mathcal{M} &= \mathrm{M} + \epsilon^2\Big\{\frac{1}{6}(\gamma-1)\ddot{\mathcal{I}}^{(2)} - \frac{1}{2}\eta\int_{V_B}\rho^*\hat{U}^{(B)}d^3w \\ &\quad -\big[Y + (2\beta-\gamma-1)P\big]\mathrm{M} - \sum_{l=1}^{\infty}\frac{1}{l!}\big[(\gamma l+1)Q_L + 2(\beta-1)P_L\big]\mathcal{I}^L\Big\} + O(\epsilon^4)\,, \end{aligned} \tag{6.9}$$

where

$$\mathcal{I}^{(2)} = \int_{V_B}\rho^* w^2 d^3w\,, \tag{6.10}$$

is the second order (rotational) moment of inertia of the body.

It is not difficult to derive a relationship between the *active* and *conformal* post-Newtonian masses by making use of equations (6.7)-(6.9). Direct comparison yields

$$\begin{aligned} \tilde{\mathcal{M}} &= \mathcal{M} + \epsilon^2\Big\{\frac{1}{2}\eta\int_{V_B}\rho^*\hat{U}^{(B)}d^3w - \frac{\gamma-1}{6}\ddot{\mathcal{I}}^{(2)} \\ &\quad +2(\beta-1)\Big(\mathcal{M}P + \sum_{l=1}^{\infty}\frac{1}{l!}P_L\mathcal{I}^L\Big) + (\gamma-1)\sum_{l=1}^{\infty}\frac{1}{(l-1)!}Q_L\mathcal{I}^L\Big\} + O(\epsilon^4)\,, \end{aligned} \tag{6.11}$$

where $\eta = 4\beta - \gamma - 3$ is called the Nordtvedt parameter [Will, 1993]. This parameter characterizes the degree of violation of the strong principle of equivalence for self-gravitating bodies [Nordtvedt, 1968a,b]. Numerical value of this parameter is known with the precision better than 0.02% from the lunar laser ranging (LLR) experiment [Dickey *et al.*, 1994], which lasts continuously already for more than 30 years.

One can see that in the scalar-tensor theory of gravity the *conformal* mass of the body differs from its *active* mass. If the body is completely isolated the difference originates from the *Nordtvedt effect*, that is for $\eta \neq 0$, and from the time-dependence of the body's rotational moment of inertia (for example, due to radial oscillations of the body). In case of the presence of external bodies the additional difference between the masses originates from the coupling of the external gravitational field multipoles, Q_L and P_L, with the internal multipole moments $\mathcal{I}_L$ of the body.

General relativistic post-Newtonian mass M of an individual body in N-body system is not conserved. Indeed, taking a time derivative from both sides of equation (6.8) and making use of the macroscopic equations of motion given in section 6.1.2, one gets [Caporali, 1981]

$$\dot{\mathrm{M}} = \epsilon^2\sum_{l=1}^{\infty}\frac{1}{l!}Q_L\dot{\mathcal{I}}^L + O(\epsilon^4)\,, \tag{6.12}$$

where the dot above M denotes the time derivative with respect to the local coordinate time u. This equation reveals that *general relativistic* mass of the body is constant, if and only if, the mass-type multipole moments $\mathcal{I}_L$ of the body do not depend on time and/or there is no external tidal field, that is $Q_L = 0$. It is worth noticing that the *conformal* and *active* post-Newtonian masses are not constant in the presence of the tidal field even if the body's multipole moments $\mathcal{I}_L$ are constant. This is because the external multipole moments Q_L enter definitions (6.7) and (6.9) of these masses explicitly and, in general, they depend on time. Hence, differentiation of equations (6.7) and (6.9) shows that the time derivatives of the *conformal* and *active* masses are not zero.

Direct calculation of the equations of motion elucidates that definition of the *conformal* mass dipole moment derived from equation (4.180) for $l = 1$, gives the most physically-justified choice of the post-Newtonian center of mass for each body. This is because after differentiation with respect to time only the *conformal* dipole moment leads to the law of conservation of the body's linear momentum, when one neglects the influence of other external bodies. The post-Newtonian *scalar* or *active* dipole moments do not possess such property. Thus, the post-Newtonian center of mass of the body, $\mathcal{J}^i \equiv \tilde{\mathcal{I}}^i$, is derived from equation (4.180) for $l = 1$, and reads

$$\begin{aligned}
\mathcal{J}^i &= \int_{V_B} \rho^* w^i \left[1 + \epsilon^2 \left(\frac{1}{2}\nu^2 + \Pi - \frac{1}{2}\hat{U}^{(B)}\right)\right] d^3w \\
&\quad -\epsilon^2 \left\{ [Y + (1-\gamma)P]\,\mathcal{J}^i_* + \sum_{l=1}^{\infty} \frac{l+1}{l!} Q_L \mathcal{I}^{iL} \right. \\
&\quad \left. + \frac{1}{2}\sum_{l=1}^{\infty} \frac{1}{(2l+1)(l-1)!} Q_{iL-1} \mathcal{N}^{L-1} \right\} + O(\epsilon^4) ,
\end{aligned} \tag{6.13}$$

where here and everywhere else

$$\mathcal{N}^L = \int_{V_B} \rho^* w^2 w^{<L>} d^3w , \qquad (l \geq 0) \tag{6.14}$$

denotes a new STF tensor [Klioner and Soffel, 2000]. One calls attention of the reader that for $l = 0$ the scalar function $\mathcal{N} \equiv \mathcal{I}^{(2)}$, where $\mathcal{I}^{(2)}$ is the rotational moment of inertia that has been defined in equation (6.10).

It is worth emphasizing that the post-Newtonian definitions of mass and the center of mass of the body depend not only on the internal distribution of matter's density, velocity, and stresses inside the body but also on terms describing the coupling of the internal gravitational field of the body with external masses. Inclusion of these coupling terms to definitions (6.7), (6.9) and (6.13) drastically simplifies translational equations of motion and bring them to the form, which can be reduced to the EIH equations of motion in the limiting case of spherically symmetric bodies. This settles the question about whether the coupling of internal and external gravitational fields should be included in the definitions of the post-Newtonian mass and the center of mass.

The post-Newtonian linear momentum of the body, $\mathcal{P}^i$, is defined as the first time derivative of the dipole moment given by equation (6.13), that is $\mathcal{P}^i = \dot{\mathcal{J}}^i$, where the dot indicates a derivative with respect to local coordinate time u. After taking the derivative, one obtains

$$\begin{aligned}
\mathcal{P}^i &= \int_{V_B} \rho^* \nu^i \left[1 + \epsilon^2 \left(\frac{1}{2}\nu^2 + \Pi - \frac{1}{2}\hat{U}^{(B)}\right)\right] d^3w \\
&+\epsilon^2 \int_{V_B} \left[\pi_{ik}\nu^k - \frac{1}{2}\rho^* \hat{W}_i^{(B)}\right] d^3w \\
&-\epsilon^2 \frac{d}{du}\left\{[Y + (1-\gamma)P]\mathcal{J}_*^i + \sum_{l=1}^{\infty} \frac{l+1}{l!} Q_L \mathcal{I}^{iL}\right. \\
&\left.+\frac{1}{2}\sum_{l=1}^{\infty} \frac{1}{(2l+1)(l-1)!} Q_{iL-1}\mathcal{N}^{L-1}\right\} \\
&+\epsilon^2 \sum_{l=1}^{\infty} \frac{1}{l!}\left[Q_L \dot{\mathcal{I}}^{iL} + \frac{l}{2l+1} Q_{iL-1}\dot{\mathcal{N}}^{L-1}\right] - \epsilon^2 \sum_{l=1}^{\infty} \frac{1}{l!} Q_L \int_{V_B} \rho^* \nu^i w^L d^3w\,,
\end{aligned} \tag{6.15}$$

where function

$$\hat{W}_i^{(B)}(u, \boldsymbol{w}) = G \int_{V_B} \frac{\rho^*(u, \boldsymbol{w}')\nu'^k(w^k - w'^k)(w^i - w'^i)}{|\boldsymbol{w} - \boldsymbol{w}'|^3} d^3w'\,. \tag{6.16}$$

Until now the point $\boldsymbol{x}_B(t)$ in the global coordinates represented location of the origin of the local coordinates taken at time t. In general, the origin of the local coordinates does not coincide with the center of mass of the body, which can move with respect to the local coordinates with some velocity and acceleration. In order to keep the center of mass of the body at the origin of its local coordinates one must prove that for any instant of time the dipole moment defined by equation (6.13) and its time derivative (that is, the linear momentum of the body) given by equation (6.15) can be made equal to zero. This requirement can be achieved, if and only if, the second time derivative of the dipole moment with respect to the local coordinate time u is identically equal to zero, that is

$$\dot{\mathcal{P}}^i = 0\,. \tag{6.17}$$

It is remarkable that this equation can be satisfied after making an appropriate choice of the external dipole moment Q_i that characterizes a locally measurable acceleration of the origin of the local coordinates with respect to a geodesic worldline of the background spacetime. This statement has been proven in [Kopejkin, 1988a] in the Newtonian approximation and, then, extended up to the post-Newtonian approximation in [Damour *et al.*, 1991]. The present book investigates the consequences of equation (6.17) in the post-Newtonian approximation of the scalar-tensor theory of gravity characterized by two PPN parameters, β and γ.

6.1.4
Translational equation of motion in the local coordinates

Translational equation of motion of a body in its own local coordinates, $w^\alpha = (cu, \boldsymbol{w})$, is derived by making use of the definition of the *conformal* linear momentum, $\mathcal{P}^i$, of the body that is defined in equation (6.15). Taking a derivative of equation (6.15) with respect to the local coordinate time u, making use of macroscopic equations of motion (6.4)-(6.6) and integrating by parts, one obtains

$$
\begin{aligned}
\dot{\mathcal{P}}^i \quad = \quad & \mathcal{M}Q_i(u) + \sum_{l=1}^{\infty}\frac{1}{l!}Q_{iL}(u)\mathcal{I}^L(u) + \epsilon^2\Delta\dot{\mathcal{P}}^i \qquad (6.18)\\
-\epsilon^2 \quad & \Bigg\{\sum_{l=2}^{\infty}\frac{1}{(l+1)!}\Big[(l^2+l+4)Q_L - 2(1-\gamma)P_L\Big]\ddot{\mathcal{I}}^{iL}\\
& + \sum_{l=2}^{\infty}\frac{2l+1}{(l+1)(l+1)!}\Big[(l^2+2l+5)\dot{Q}_L - 2(1-\gamma)\dot{P}_L\Big]\dot{\mathcal{I}}^{iL}\\
& + \sum_{l=2}^{\infty}\frac{2l+1}{(2l+3)(l+1)!}\Big[(l^2+3l+6)\ddot{Q}_L - 2(1-\gamma)\ddot{P}_L\Big]\mathcal{I}^{iL}\\
& + \Big[3Q_k - (1-\gamma)P_k\Big]\ddot{\mathcal{I}}^{ik} + \frac{3}{2}\Big[4\dot{Q}_k - (1-\gamma)\dot{P}_k\Big]\dot{\mathcal{I}}^{ik}\\
& + \frac{3}{5}\Big[5\ddot{Q}_k - (1-\gamma)\ddot{P}_k\Big]\mathcal{I}^{ik} + \sum_{l=2}^{\infty}\frac{1}{l!}\dot{Z}_{iL}\mathcal{I}^L\\
& + \sum_{l=1}^{\infty}\frac{1}{(l+1)!}\varepsilon_{ipq}\Big[\dot{C}_{pL}\mathcal{I}^{<qL>} + \frac{l+2}{l+1}C_{pL}\dot{\mathcal{I}}^{qL}\Big]\\
& - 2\sum_{l=1}^{\infty}\frac{l+1}{(l+2)!}\varepsilon_{ipq}\Big[\Big(2Q_{pL} - (1-\gamma)P_{pL}\Big)\dot{\mathcal{S}}^{qL}\\
& + \frac{l+1}{l+2}\Big(2\dot{Q}_{pL} - (1-\gamma)\dot{P}_{pL}\Big)\mathcal{S}^{qL}\Big] - \sum_{l=1}^{\infty}\frac{l(l+2)}{(l+1)(l+1)!}C_{iL}\mathcal{S}^L\\
& - \frac{1}{2}\varepsilon_{ikq}\Big[\Big(4Q_k - 2(1-\gamma)P_k\Big)\dot{\mathcal{S}}^q + \Big(2\dot{Q}_k - (1-\gamma)\dot{P}_k\Big)\mathcal{S}^q\Big]\\
& + \Big(P_i - Q_i\Big)\Big[\frac{1-\gamma}{6}\ddot{\mathcal{I}}^{(2)} + \frac{1}{2}\eta\int_{V_B}\rho^*\hat{U}^{(B)}d^3w\\
& + \sum_{l=2}^{\infty}\frac{1}{l!}\Big(2(\beta-1)P_L - (1-\gamma)lQ_L\Big)\mathcal{I}^L\Big]\Bigg\} + O(\epsilon^4)\,,
\end{aligned}
$$

where one has shown explicitly all terms proportional to the external dipole moment Q_i and again used notation $\eta = 4\beta - \gamma - 3$ for the Nordtvedt parameter. The post-Newtonian correction $\Delta\dot{\mathcal{P}}^i$ in the first line of equation (6.18) is given by a rather long expression

$$
\begin{aligned}
\Delta\dot{\mathcal{P}}^i &= \sum_{l=1}^{\infty}\frac{1}{l!}\Bigg[\frac{2(1-\gamma)(2l+1)}{(2l+3)(l+1)}\frac{d}{du}\int_{V_B}\rho^* v^k w^{<kL>} d^3w \\
&+(\gamma-1)\int_{V_B}(\rho^* v^2+\hat{\sigma}^{kk})w^L d^3w \\
&+2(1-\beta)\int_{V_B}\rho^*\left(\hat{U}^{(B)}+\sum_{n=0}^{\infty}\frac{1}{n!}P_N w^N\right)w^L d^3w\Bigg]\Big(P_{iL}-Q_{iL}\Big) \\
&+(Q_i-P_i)\Big[2(\beta-1)P_k+(\gamma-1)Q_k\Big]\mathcal{I}^k-\frac{1}{3}\Big[6\ddot{Y}+4(1-\gamma)\ddot{P}\Big]\mathcal{I}^i \\
&-\dot{Z}_{ik}\mathcal{I}^k-\varepsilon_{ipq}\Big(\dot{C}^p\mathcal{I}^q+2C^p\dot{\mathcal{I}}^q\Big) \\
&-\Big[\dot{Q}+4\dot{Y}+2(1-\gamma)\dot{P}\Big]\dot{\mathcal{I}}^i-2\Big[Y-(\beta-1)P\Big]\ddot{\mathcal{I}}^i\,.
\end{aligned}
\tag{6.19}
$$

One emphasizes that $\mathcal{M}$ and $\mathcal{I}^L$ ($l \geq 1$) in equations (6.18), (6.19) are respectively the *active* mass and the *active* multipole moments of the body, both depending on time u taken at the origin of the local coordinates.

External dipole moment Q_i defines acceleration of the origin of the local coordinates with respect to a geodesic worldline passing through the origin of the local coordinates at the instant of time u. This acceleration is not restricted by the gravitational field equations and can be chosen arbitrary. Its choice determines the worldline of the origin of the local coordinates. If one chooses $Q_i = 0$ the origin of the local frame, $\boldsymbol{x}_B$, will move along a geodesic worldline defined in the global frame by equation (5.88) while the center of mass of the body under consideration will move in accordance with the equation of motion (6.18) with respect to this geodesic.

Practical applications, however, demand [Kopejkin, 1988a; Soffel *et al.*, 2003] to chose the origin of the local frame being always located at the center of mass of the body. This can be accomplished by assuming that the forces exerted on the body obeys the third Newton's law in the local coordinates. It can be satisfied by equating the net force to zero, that is imposing condition (6.17) with $\dot{\mathcal{P}}^i \equiv \ddot{\mathcal{J}}^i = 0$. It leads to appearance in the local coordinates of the force of inertia associated with the inertial acceleration Q_i, which compensates the force of the gravitational interaction between the internal multipole moments of the body and the external multipoles of the tidal gravitational field. The local coordinates become non-inertial but it makes equation (6.17) justified for any instant of time. Subsequent integration of this equation with respect to time allows us to make the origin of the local coordinates moving along the worldline of the center of mass of the body. This is achieved after the *conformal* dipole moment $\mathcal{J}^i$ is subject to the following conditions: $\dot{\mathcal{J}}^i = \mathcal{J}^i = 0$. Similar conditions are obtained for the *active* dipole moment, $\ddot{\mathcal{I}}^i = \dot{\mathcal{I}}^i = \mathcal{I}^i = 0$ but only in the Newtonian approximation. In the post-Newtonian approximation the *active* dipole moment of the body $\mathcal{I}^i \neq 0$, even if the *conformal* multipole moment of the body $\mathcal{J}^i = 0$, because these two moments differ in the post-Newtonian approximation as shown in section 4.5.3, and more explicitly in equation (6.28).

Placing the origin of the local coordinates at the body's center of mass with ($\ddot{\mathcal{J}}^i =$

$\dot{\mathcal{J}}^i = \mathcal{J}^i = 0$) and noticing that $Q_L = P_L$ for any index $l \geq 2$ makes the post-Newtonian function $\Delta\dot{\mathcal{P}}^i = 0$. Relationship (6.18) becomes algebraic equation for function Q_i, which can be immediately solved. Separating Q_i from all other terms in equation (6.18), taking all terms with Q_i to the left side, and solving for Q_i yields the following post-Newtonian equation

$$\tilde{\mathcal{M}}_{ij}Q_j = \mathbb{F}^i_N + \epsilon^2\left(\mathbb{F}^i_{pN} + \Delta\mathbb{F}^i_{pN}\right) + O(\epsilon^4)\,, \tag{6.20}$$

where the *conformal* anisotropic tensor of mass

$$\tilde{\mathcal{M}}_{ij} \;=\; \tilde{\mathcal{M}}\delta_{ij} - \epsilon^2\left[3\ddot{\mathcal{I}}^{ij} - 2\sum_{l=1}^{\infty}\frac{1}{l!}\left(Q_{jL}\mathcal{I}^{iL} - Q_{iL}\mathcal{I}^{jL}\right)\right], \tag{6.21}$$

and the tidal gravitational forces

$$\mathbb{F}^i_N \;=\; -\sum_{l=1}^{\infty}\frac{1}{l!}\,Q_{iL}(u)\mathcal{I}^L(u)\,, \tag{6.22}$$

$$\begin{aligned}
\mathbb{F}^i_{pN} \;=\;& \sum_{l=2}^{\infty}\frac{l^2+l+2(1+\gamma)}{(l+1)!}\,Q_L\ddot{\mathcal{I}}^{iL} + 6\dot{Q}_k\dot{\mathcal{I}}^{ik} + 3\ddot{Q}_k\mathcal{I}^{ik} \\
&+\sum_{l=2}^{\infty}\frac{(2l+1)}{(l+1)!}\left(\frac{l^2+2l+2\gamma+3}{l+1}\,\dot{Q}_L\dot{\mathcal{I}}^{iL} + \frac{l^2+3l+2\gamma+4}{2l+3}\,\ddot{Q}_L\mathcal{I}^{iL}\right) \\
&+\sum_{l=2}^{\infty}\frac{1}{l!}\varepsilon_{ipq}\left(\dot{C}_{pL-1}\mathcal{I}^{qL-1} + \frac{l+1}{l}\,C_{pL-1}\dot{\mathcal{I}}^{qL-1}\right) \\
&-2(1+\gamma)\sum_{l=2}^{\infty}\frac{l}{(l+1)!}\varepsilon_{ipq}\left(Q_{pL-1}\dot{\mathcal{S}}^{qL-1} + \frac{l}{l+1}\,\dot{Q}_{pL-1}\mathcal{S}^{qL-1}\right), \\
&-\sum_{l=1}^{\infty}\frac{l(l+2)}{(l+1)(l+1)!}\,C_{iL}\mathcal{S}^L + \sum_{l=2}^{\infty}\frac{1}{l!}\,\dot{Z}_{iL}\mathcal{I}^L - \varepsilon_{ipq}\dot{Q}_p\mathcal{S}^q
\end{aligned} \tag{6.23}$$

$$\begin{aligned}
\Delta\mathbb{F}^i_{pN} \;=\;& (1-\gamma)\left(\frac{1}{2}\varepsilon_{ikq}\dot{P}_k\mathcal{S}^q - \frac{3}{2}\dot{P}_k\dot{\mathcal{I}}^{ik} - \frac{3}{5}\ddot{P}_k\mathcal{I}^{ik}\right) \\
&+\;(1-\gamma)\left[\sum_{l=1}^{\infty}\frac{1}{l!}\left(Q_{kL}\mathcal{I}^{iL} - Q_{iL}\mathcal{I}^{kL}\right) - \ddot{\mathcal{I}}^{ik}\right]P_k \\
&+\;\left[\frac{1}{2}\eta\int_{V_B}\rho^*\hat{U}^{(B)}d^3w + \frac{1-\gamma}{6}\,\ddot{\mathcal{I}}^{(2)} + \sum_{l=2}^{\infty}\frac{2(\beta-1)+(\gamma-1)l}{l!}\,Q_L\mathcal{I}^L\right]P_i\,.
\end{aligned} \tag{6.24}$$

Equations (6.20)-(6.24) describe the inertial acceleration Q_i of the body in the local coordinates in the presence of external bodies, which create the inertial force standing in the right side of equation (6.20) and displace the motion of the body's center of mass from a geodesic worldline. Newtonian, $\mathbb{F}^i_N$, and the post-Newtonian, $\mathbb{F}^i_{pN}$, tidal forces are caused by gravitational coupling of the body's internal (*active*) multipole moments, $\mathcal{I}_L$ and $\mathcal{S}_L$, with the external multipole moments, Q_L and C_L. The post-Newtonian tidal force, $\mathbb{F}^i_{pN}$, is reduced in the limit of $\gamma = 1$ to general relativistic expression derived previously by Damour *et al.* [1991]. The spin-dipole term,

$\sim C_{ij}\mathcal{S}^j$, in equation (6.23) has been derived by Mathisson [Mathisson, 1937] and Papapetrou [Papapetrou, 1951a] with a different mathematical technique. The derivation has been reproduced in the paper by Dixon [Dixon, 1979], who has worked out a powerful geometric approach to the problem of motion of extended bodies based on the invariant concept of the 'skeletonized' tensor of energy-momentum of matter. Dixon also obtained equation (6.22) with the internal multipole moments $\mathcal{I}^L$ defined in terms of the 'skeletonized' tensor of matter. Our equations (6.20)-(6.24) represent a comprehensive post-Newtonian generalization of the Mathisson-Papapetrou-Dixon equations of motion for rotating bodies with arbitrary internal structure.

It is worthwhile to emphasize that summation with respect to index l in equation (6.22) begins from $l = 1$. The point is that one has defined the center of mass of the body B in terms of the *conformal* dipole moment $\mathcal{J}^i$ by the condition $\mathcal{J}^i = 0$. However, the force $\mathbb{F}^i_N$ depends on the *active* multipole moments of the body and the *active* dipole moment $\mathcal{I}^i \neq \mathcal{J}^i$. Hence, $\mathcal{I}^i \neq 0$, and one should retain the dipole term with $l = 1$ in equation (6.22). The contribution to the force $\mathbb{F}^i_N$ coming out of the non-zero *active* dipole of the body has a post-Newtonian order of magnitude and can be written down more explicitly as

$$\left(\mathbb{F}^i_N\right)_{\rm dipole} = -Q_{ij}\mathcal{I}^j \;, \tag{6.25}$$

where the *active* dipole moment $\mathcal{I}^i$ is

$$\begin{aligned}
\mathcal{I}^j &= \epsilon^2\Bigg\{-\frac{1}{2}\eta\int_{V_B}\rho^*\hat{U}_{(B)}w^j d^3w + \frac{1}{5}(\gamma-1)\left[3\dot{\mathcal{R}}^j - \frac{1}{2}\ddot{\mathcal{N}}^j\right] \\
&+ \sum_{l=0}^{\infty}\frac{1}{l!}\Big[(1-\gamma)l\,Q_L + 2(1-\beta)P_L\Big]\mathcal{I}^{jL} \\
&+ \frac{1}{2}\sum_{l=0}^{\infty}\frac{1}{(2l+3)l!}\Big[(\gamma-1)Q_{jL} + 4(1-\beta)P_{jL}\Big]\mathcal{N}^L\Bigg\} + O\left(\epsilon^4\right) ,
\end{aligned} \tag{6.26}$$

where

$$\mathcal{R}^i = \int_{V_B}\rho^* v^k w^{<k}w^{i>}d^3w \;. \tag{6.27}$$

If one takes into account the explicit relationship between the multipole moments P_L of the scalar field and gravitational potential of the external bodies, then equation (6.26) can be slightly simplified

$$\begin{aligned}
\mathcal{I}^j &= \epsilon^2\Bigg\{-\frac{1}{2}\eta\int_{V_B}\rho^*\hat{U}_{(B)}w^j d^3w + \frac{1}{5}(\gamma-1)\left[3\dot{\mathcal{R}}^j - \frac{1}{2}\ddot{\mathcal{N}}^j\right] \\
&+ 2(1-\beta)\left[\bar{U}(\boldsymbol{x}_B)\mathcal{I}^j + a^k_B\mathcal{I}^{jk} + \frac{1}{3}a^j_B\mathcal{N}\right] \\
&- \frac{\eta}{2}\sum_{l=0}^{\infty}\frac{1}{(2l+3)l!}Q_{jL}\mathcal{N}^L + \sum_{l=1}^{\infty}\frac{(1-\gamma)l + 2(1-\beta)}{l!}Q_L\mathcal{I}^{jL}\Bigg\} + O\left(\epsilon^4\right) .
\end{aligned} \tag{6.28}$$

It is clear that the *active* dipole moment of the body can contribute to the equations of motion of the body only in the scalar-tensor theory of gravity because in general relativity $\beta = \gamma = 1$ and there is no difference between the *conformal* and *active* internal dipoles. One notices the presence of the self-gravitational energy coupled with the Nordtvedt parameter $\eta = 4\beta - \gamma - 3$ in the definition of the *active* dipole moment of the body. Similar terms appear in the definition of the *active* mass of the body leading to inequality of inertial and gravitational masses of the body. Presence of the Nordtvedt parameter-dependent term in the *active* dipole moment is not important for the bodies, whose shape is close to spherically-symmetric. However, it may play a role in motion of sub-systems, like Earth and Moon, which possess a large deviation from spherical symmetry. More detailed study of this problem is desirable.

Gravitational force $\Delta\mathbb{F}^i_{pN}$ presents only in the scalar-tensor theory of gravity. Its magnitude is proportional to the external dipole moment of the scalar field, P_i, and its time derivatives. The dipole P_i couples with the self-gravitational energy $\sim \int \rho^* \hat{U}^{(B)}$ of the body, with the energy of external gravitational field, and with the kinetic energy of the body's internal motion given by the term $\ddot{\mathcal{I}}^{(2)}$. Next section will demonstrate that it is this coupling that is responsible for the inequality of inertial and gravitational masses of the body (the Nordtvedt effect) leading to violation of the strong principle of equivalence.

6.1.5
Orbital equation of motion in the global coordinates

Equation of orbital motion of the center of mass of body B in the global coordinates $x^\alpha = (ct, \boldsymbol{x})$ are obtained from the equation of motion (5.88) of the origin of the local coordinates, $\boldsymbol{x}_B$, where the inertial acceleration Q_i obeys the translational equation of motion (6.20). One uses equation (5.58) to replace, in forces $\mathbb{F}^i_{pN}$, and $\Delta\mathbb{F}^i_{pN}$ all terms depending explicitly on Q_i, with a linear combination of the barycentric acceleration a^i_B and the gradient $\bar{U}_{,i}(\boldsymbol{x}_B)$ of the gravitational potential of external bodies, $Q_i = \bar{U}_{,i}(\boldsymbol{x}_B) - a^i_B$. Moreover, one expresses the scalar external multipoles in terms of the external gravitational potential: $P_L = \bar{U}_{,L}(\boldsymbol{x}_B)$, that is equation (5.55). Then, we take all terms depending on the gradient $\bar{U}_{,i}(\boldsymbol{x}_B)$ of the external gravitational potential, to the right side of equation (5.88) and put those depending on the barycentric acceleration of the body's center of mass a^i_B, to its left side. It brings equation of motion (5.88) to the following form

$$
\begin{aligned}
\tilde{\mathcal{M}}_B a_B^i &= \mathcal{M}_B\left\{\bar{U}_{,i}(\boldsymbol{x}_B) + \epsilon^2\left[\bar{\Phi}_{,i}(\boldsymbol{x}_B) - \frac{1}{2}\bar{\chi}_{,itt}(\boldsymbol{x}_B)\right]\right\} - \mathbb{F}^i_N - \epsilon^2\mathbb{F}^i_{pN} \qquad (6.29)\\
&+ \epsilon^2\mathcal{M}_B\left\{\left[\gamma\delta_{ik}v_B^2 - v_B^i v_B^k - 2(\gamma+\beta)\delta_{ik}\bar{U}(\boldsymbol{x}_B)\right]\bar{U}_{,k}(\boldsymbol{x}_B) - \mathcal{A}\dot{Q}_i\right.\\
&\left.+2(1+\gamma)\dot{\bar{U}}^i(\boldsymbol{x}_B) - 2(1+\gamma)v_B^k\bar{U}^{k,i}(\boldsymbol{x}_B) - (1+2\gamma)v_B^i\dot{\bar{U}}(\boldsymbol{x}_B)\right\}\\
&+ \epsilon^2\left\{\left[2Q - Y - v_B^2 - (2+\gamma)\bar{U}(\boldsymbol{x}_B)\right]\delta_{ik} - \frac{1}{2}v_B^i v_B^k - F^{ik} + \frac{3}{\mathcal{M}_B}\ddot{\mathcal{I}}^{ik}\right.\\
&\left.+\frac{1}{\mathcal{M}_B}\sum_{n=1}^{\infty}\frac{1}{n!}\left[Q_{iN}\mathcal{I}^{kN} - Q_{kN}\mathcal{I}^{iN}\right]\right\}\sum_{l=2}^{\infty}\frac{1}{l!}Q_{kL}\mathcal{I}^L\\
&+ \epsilon^2(1-\gamma)\left\{\ddot{\mathcal{I}}^{ik}\bar{U}_{,k}(\boldsymbol{x}_B) + \frac{3}{2}\dot{\mathcal{I}}^{ik}\dot{\bar{U}}_{,k}(\boldsymbol{x}_B) + \frac{3}{5}\mathcal{I}^{ik}\ddot{\bar{U}}_{,k}(\boldsymbol{x}_B)\right.\\
&\left.+\frac{1}{2}\varepsilon_{ipk}\mathcal{S}^p\dot{\bar{U}}_{,k}(\boldsymbol{x}_B) + \sum_{l=1}^{\infty}\frac{1}{l!}\left(Q_{iL}\mathcal{I}^{kL} - Q_{kL}\mathcal{I}^{iL}\right)\bar{U}_{,k}(\boldsymbol{x}_B)\right\} + O(\epsilon^4)\,,
\end{aligned}
$$

where the external potentials $\bar{U}(\boldsymbol{x}_B)$, $\bar{U}^i(\boldsymbol{x}_B)$, $\bar{\Phi}(\boldsymbol{x}_B)$ and $\bar{\chi}(\boldsymbol{x}_B)$ are taken on the world-line of the center of mass of the body B at the instant of time t as defined in equation (5.32). The external potential of any other body A (A≠B) can be further expanded in a multipolar series around its own center of mass $\boldsymbol{x}_A$ so that the orbital equation of motion (6.29) will depend on the *active* multipole moments of the body B and the multipole moments of other bodies of the N-body system. Some research in this direction has been pursued in papers by Xu *et al.* [1997] and Racine and Flanagan [2005]. Complete calculation of the orbital equations of motion of the bodies with all possible multipole moments taken into account, have been recently executed by Y. Xie and S. Kopeikin. This general result is too complicated and goes beyond the scope of the present book. One considers more simple case of the post-Newtonian equations of spherically-symmetric and rotating bodies in section 6.3 as it has practical application in calculation of ephemerides of the solar system bodies. Moreover, it allows us to reproduce the most important physical aspects of the post-Newtonian celestial mechanics.

One has to notice that the inertial mass $\tilde{\mathcal{M}}_B$ of the body B in the left side of equation (6.29) is its *conformal* mass. It is not equal to the gravitational mass $\mathcal{M}_B$ of the body appearing in the right side of this equation as its *active* mass. The difference between the two masses is given by equation (6.11) and leads to violation of the strong principle of equivalence for massive extended bodies in the scalar-tensor theory of gravity. The fact that inertial and gravitational masses can not be equal in alternative theories of gravity was pointed out by Dicke [Brans and Dicke, 1961; Dicke, 1962a,b] and Nordtvedt [Dickey *et al.*, 1994; Nordtvedt, 1973, 1970]. The difference between the masses originates from the coupling of the external scalar multipole $P_i = \bar{U}_{,i}(\boldsymbol{x}_B)$ with gravitational and kinetic internal energies of the body as shown in equation (6.24). This coupling contributes to the gravitational mass of the body but does not affect its inertial mass.

Forces $\mathbb{F}^i_N$ and $\mathbb{F}^i_{pN}$ are given by equations (6.22) and (6.23). Terms in the first and second curled brackets of equation (6.29), which are proportional to mass $\mathcal{M}_B$, are the post-Newtonian corrections to the Newtonian force acting on the body B considered as a monopole massive particle. The group of terms in the third and forth curled brackets in equation (6.29) represents the post-Newtonian correction to the Newtonian tidal force $\mathbb{F}^i_N$ and takes into account higher-order internal multipoles of the body B. In particular, these terms contain the time-dependent scaling functions, Q and Y, which define the unit of time and length in the local coordinates [Klioner *et al.*, 2009a; Kopeikin, 2010a]. This correction also contains matrix of the relativistic precession F^{ik} given by equation (5.99).

Equation (6.29) describes a generic case of orbital equation of motion of extended bodies having arbitrary structure and rotation. They must be supplemented by the post-Newtonian equations of rotational motion of the bodies, which will be derived in the next section.

6.2
Rotational Equations of Motion of Extended Bodies

6.2.1
The angular momentum of a self-gravitating body

Rotational equations of motion of a body define orientation of its angular momentum (spin) at each instant of time with respect to the local coordinates $w^\alpha = (cu, \boldsymbol{w})$, which spatial axes are not dynamically-rotating. It means that the spatial axes are transported in spacetime along the worldline of the origin of the local coordinates in accordance with the Fermi-Walker law [Misner *et al.*, 1973] given by equation (5.99) describing orientation of the spatial axes of the local coordinates with respect to the global coordinates at each instant of time.

First, one needs to introduce the post-Newtonian definition of the angular momentum of an extended body that is a member of N-body system. It seems that the most straightforward procedure to get it is to use the multipolar expansion of $\hat{g}_{0i}(u, \boldsymbol{w})$ component of the metric tensor calculated up to the post-post-Newtonian approximation [Misner *et al.*, 1973; Thorne and Hartle, 1985]. Damour and Iyer [1991a,b] have indeed applied this procedure for deriving the post-Newtonian definition of the angular momentum of a single, isolated body. Unfortunately, it is not known if this procedure can be applied to defining the post-Newtonian angular momentum of a body from N-body system as it involves a lot of mathematical subtleties, yet unresolved. Therefore, another approach, proposed by Damour *et al.* [1993], will be used to bypass this difficulty.

Let us introduce a *bare* spin of a single body B in the local coordinates by making use of the following post-Newtonian definition [Misner *et al.*, 1973]

$$\mathcal{S}^i = \epsilon \int \varepsilon_{ijk} w^j \hat{\Theta}^{0k} d^3 w \,, \tag{6.30}$$

where a hat above any quantity means that it is expressed in the local coordinates, and

$$\hat{\Theta}^{0k} = (-\hat{g}) \frac{\phi}{\phi_0} \left(c^2 \hat{T}^{0k} + \hat{t}^{0k} \right) \,, \tag{6.31}$$

is a linear combination of the tensor of energy-momentum of matter $T^{\mu\nu}$ and the pseudo-tensor of gravitational field $\hat{t}^{\mu\nu}$. Definition of the bare spin of the body B by equation (6.30) corresponds to the conserved spin of an isolated astronomical system of N bodies as defined in equation 4.126. One also stipulates that the center of mass of the body B is chosen in the local coordinates such that its conformal dipole moment $\mathcal{J}^i$ is zero.

Integration in equation (6.30) is formally performed in the local coordinates over entire space. Tensor of energy-momentum $\hat{T}^{0k}$ includes the matter of the body B only but one assumes that it depends on the complete metric tensor $\hat{g}_{\mu\nu}$ in the local coordinates. On the other hand, one postulates that the pseudo-tensor $\hat{t}^{0k}$ depends only on the internal part of the metric tensor in the local coordinates of the body B.

Integration by parts allows us to represent the bare spin of the body B in equation (6.30) in the following form

$$\begin{aligned}\mathcal{S}^i &= \epsilon^2 \int_{V_B} \varepsilon_{ijk} w^j \Big\{ \rho^* v^k \Big[c^2 + \frac{1}{2} v^2 + \Pi + (2\gamma + 1)\hat{U}_{(B)} \\ &+ \sum_{l=1}^{\infty} \frac{1}{l!} \Big(3Q_L + 2(\gamma - 1)P_L \Big) w^L + 3Y + (1-\gamma)P \Big] \\ &+ \pi^{kn} v^n - \frac{1}{2} \rho^* \Big[\hat{W}_k^{(B)} + (3 + 4\gamma)\hat{U}^k_{(B)} \Big] \Big\} + O\left(\epsilon^4\right) ,\end{aligned} \tag{6.32}$$

where the integration is now only over the volume of the body B and potential $\hat{W}_k^{(B)}$ is defined by equation (6.16). Equation (6.32) will be used to derive the rotational equations of motion of the angular momentum of the body.

6.2.2
Equations of rotational motion in the local coordinates

Rotational equations of motion for body's spin are derived by differentiation of equation (6.32) with respect to the local coordinate time u. After taking the time derivative and making use of the macroscopic equations of motion in local coordinates given in section 6.1.2, one makes several transformations of the integrand to reduce similar terms and to simplify the final result. After tedious but straightforward calculations, one obtains equations of the rotational motion of the body B

$$\frac{d\mathcal{S}^i}{du} = \mathcal{T}^i + \epsilon^2 \left(\Delta\mathcal{T}^i - \frac{d}{du} \Delta\mathcal{S}^i \right) + O\left(\epsilon^4\right) , \tag{6.33}$$

where $\mathcal{T}^i$ is a linear superposition of generic Newtonian and post-Newtonian torques, $\Delta\mathcal{T}^i$ is its post-Newtonian correction in the scalar-tensor theory of gravity due to the explicit presence of the scalar field, and $\Delta\mathcal{S}^i$ can be considered as a supplementary post-Newtonian contribution to the bare spin $\mathcal{S}^i$. The torque and other terms in the right side of equation (6.33) read as follows:

$$\mathcal{T}^i = \sum_{l=0}^{\infty}\frac{1}{l!}\varepsilon_{ijk}\left[\mathcal{I}^{jL}\left(Q_{kL}-\epsilon^2\dot{Z}_{kL}\right)+\epsilon^2\mathcal{S}^{jL}C_{kL}\right] + \left[Y+Q+(2\beta-\gamma-1)\bar{U}(\boldsymbol{x}_B)\right]\sum_{l=0}^{\infty}\frac{1}{l!}\varepsilon_{ijk}\mathcal{I}^{jL}Q_{kL}\,, \tag{6.34}$$

$$\Delta\mathcal{T}^i = \varepsilon_{ijk}a^j_B\left[\frac{3(1-\gamma)}{5}\dot{\mathcal{R}}^k+\frac{\gamma-1}{10}\ddot{\mathcal{N}}^k+\frac{\eta}{2}\int_{V_B}\rho^*\hat{U}_{(B)}w^k d^3w + \frac{\eta}{2}\sum_{l=0}^{\infty}\frac{1}{(2l+3)l!}Q_{kL}\mathcal{N}^L+\sum_{l=1}^{\infty}\frac{(\gamma-1)l+2(\beta-1)}{l!}Q_L\mathcal{I}^{kL} + 2(\beta-1)\left(a^n_B\mathcal{I}^{kn}+\frac{1}{3}a^k_B\mathcal{N}\right)\right], \tag{6.35}$$

$$\Delta\mathcal{S}^i = -\sum_{l=1}^{\infty}\frac{1}{l!}\mathcal{I}^{iL}C_L+\sum_{l=0}^{\infty}\frac{l+2}{(2l+3)(l+1)!}\mathcal{N}^LC_{iL} + \sum_{l=0}^{\infty}\frac{1}{(2l+5)l!}\varepsilon_{ijk}\left[\frac{1}{2}\dot{\mathcal{N}}^{jL}Q_{kL}-\frac{l+2(2\gamma+3)}{2(l+2)}\mathcal{N}^{jL}\dot{Q}_{kL}-\frac{2(1+\gamma)(2l+3)}{l+2}\mathcal{R}^{jL}Q_{kL}\right] + \frac{1-\gamma}{5}\varepsilon_{ijk}\left[3\mathcal{R}^ja^k_B+\mathcal{N}^j\dot{a}^k_B\right]+\left[Q-Y+(\gamma-1)\bar{U}(\boldsymbol{x}_B)\right]\mathcal{S}^i\,, \tag{6.36}$$

where

$$\mathcal{R}^L = \int_{V_B}\rho^* v^k w^{<kL>}d^3w\,, \tag{6.37}$$

is additional set of multipole moments which has been used already in definition of the multipole moments in section 4.4.

Gravitational torque $\mathcal{T}^i$ depends on the multipole moments Z_L, which define the residual gauge freedom. The gauge multipoles Z_L can be chosen to eliminate some terms in the torque $\mathcal{T}^i$ depending on the external multipoles Q_L. The post-Newtonian correction $\Delta\mathcal{T}^i = \varepsilon_{ijk}a^j_B\left(\mathcal{J}^i-\mathcal{I}^i\right)$, where $\mathcal{I}^i$ and $\mathcal{J}^i$ are the *active* and *conformal* dipole moments of the body B respectively. The difference between the two dipole moments taken under condition that $\mathcal{J}^i = 0$ is given by equation (6.28), and re-appears in equation (6.35). One has taken into account in equation (6.34) the external monopole moments Q and Y defining the units of measurement of the local time and spatial coordinates respectively in accordance with the IAU 2000 Resolutions [Soffel *et al.*, 2003]. Their contribution to the rotational equations of motion is extremely small and can be omitted in practical calculations. One re-defines the spin of the body as

$$\mathcal{S}^i_+ = \mathcal{S}^i+\epsilon^2\Delta\mathcal{S}^i\,, \tag{6.38}$$

so that equations of rotational motion acquire their final form [Kopeikin and Vlasov, 2004]

$$\frac{d\mathcal{S}^i_+}{du} = \mathcal{T}^i+\epsilon^2\Delta\mathcal{T}^i+O\left(\epsilon^4\right). \tag{6.39}$$

Derivation of this equation reveals that the bare value of the spin of the body B is not directly observable. Equations of the rotational motion of the body contain additional terms, which should be interpreted as a part of the total spin of the body. It is the total spin $\mathcal{S}^i_+$, which is conserved, not its bare value $\mathcal{S}^i$.

Equations of rotational motion (6.38) should be complemented by a relativistic relationship giving connection between the total spin $\mathcal{S}^i_+$ of the body and its angular velocity of rotation ω^i. This relationship is well-known in the Newtonian theory and given by equation

$$\mathcal{S}^i_+ = I^{ij}\omega^j \,, \tag{6.40}$$

where I^{ij} is tensor of the moment of inertia of the body [Landau and Lifshitz, 1969]. Extrapolation of equation (6.40) to the post-Newtonian approximation has not yet been achieved. Some research towards solution of this problem has been undertaken by Klioner [Klioner, 1996].

One has compared equations (6.39) with analogous equations derived by Klioner and Soffel [Klioner and Soffel, 2000, equations (9.42)-(9.47)] in their own approach to the parameterized post-Newtonian formalism with the local frames. First of all, one notices that definition of the multipole moments C_L differs by a numerical factor $(l+1)/l$ from that, $C_L^{\text{Klioner-Soffel}}$, used by Klioner and Soffel, that is

$$C_L = -2(1+\gamma)\,\frac{l+1}{l}\,C_L^{\text{Klioner-Soffel}} \,. \tag{6.41}$$

Comparison of spin $\mathcal{S}^i_+$ with that S'^i given by Klioner and Soffel, shows that they are equal if the sign minus in front of three last terms in equation (9.45) of Klioner-Soffel's paper [Klioner and Soffel, 2000], is replaced with sign plus. The general-relativistic torque $\mathcal{T}^i$ in equation (6.34) coincides naturally with that derived in [Klioner and Soffel, 2000]. The biggest difference occurs between the post-Newtonian correction $\Delta\mathcal{T}^i$ to the torque, equation (6.35), and a corresponding quantity given in equations (9.43), (9.46) in paper [Klioner and Soffel, 2000]. First three terms in equation (6.35) completely coincides with equation (9.43) derived by Klioner and Soffel, thus, confirming the presence of the Nordtvedt effect for the rotational motion of the bodies. However, we obtained different terms in the second and third lines of equation (6.35) as contrasted with equation (9.46) by Klioner and Soffel [2000]. It is likely that the difference originates from slightly different gauge conditions used in this book and in [Klioner and Soffel, 2000]. Additional origin of the difference is that the present book uses definition of the center of mass of the body B, which is not reduced to that used by Klioner and Soffel.

6.3
Motion of Spherically-Symmetric and Rigidly-Rotating Bodies

Major astronomical bodies in the solar system (planets, Sun) are almost spherically-symmetric. Their rotation is also very close to that of a rigid body. Newtonian theory of equilibrium of self-gravitating bodies [Moritz, 1989; Zharkov and Trubitsyn, 1978] describes the existing deviations from the spherically-symmetric shape and the rigid rotation of the Earth and planets with a remarkable precision. These deviations are so small that their account in the post-Newtonian approximation yields corrections that are comparable with the post-post Newtonian terms, which are considered as negligible in this book. For this reason, one will focus on derivation of the post-Newtonian equations of motion of spherically-symmetric and rigidly rotating bodies. These equations have been discussed by many previous authors (see, for example, books [Brumberg, 1991; Fock, 1964; Infeld and Plebanski, 1960; Soffel, 1989; Will, 1993] and references therein). Nevertheless, this section gives a new insight to the old problem from the point of view of the matching technique, the global and local coordinates, investigation of the effacing principle and the strong principle of equivalence, which have not yet received an adequate treatment.

6.3.1
Definition of a spherically-symmetric and rigidly rotating body

First, one needs to give an exact definition of spherically-symmetric and rotating body from N body system, which goes beyond the Newtonian theory. The main point is that the concept of spherical symmetry and rigid rotation for an extended body is not invariant but coordinate-dependent [Kopejkin, 1988b]. This is seen already in special relativity, where the spatial coordinate grid of a moving inertial frame is linearly deformed with respect to that of inertial frame at rest and the magnitude of this Lorentz deformation depends on velocity of the moving frame (see Section 2.3.5). This observation assumes that if one considers a spherically-symmetric body in a static frame it will be not spherically-symmetric in the moving frame. Deformation of the body's shape can be calculated by applying the Lorentz transformation to the equation describing the shape of the body in the static frame [Batygin and Toptygin, 1978]. However, physics of this deformation is more involved. Moreover, the Lorentz deformation of the moving, spherically-symmetric body is invisible [Batygin and Toptygin, 1978, problem 583]. The reason is that the Lorentz deformation is solely coordinate effect, which can not generate real physical stresses (tensions) inside the moving body. Nonetheless, the Lorentz deformation of the body's shape has to be taken into account for making correct calculation of observed physical effects associated with the translational motion of the body. Poincaré and Lorentz were first, who took into account the special relativistic deformation of a moving electron for calculation of the electromagnetic radiation-reaction force exerted by the electron on itself due to the emission of electromagnetic radiation [Jackson, 1998].

In general relativity, gravitational field causes the Einstein deformation of the coordinate grid of a global static frame with respect to that of a local inertial frame of observer, where the gravitational field is absent as a consequence of the principle of equivalence [Novikov and Frolov, 1989, §4.3]. This represents a pure mathematical comparison between the tetrad bases of the two frames [Misner *et al.*, 1973]. It should not be interpreted physically as shielding of gravitational field, which does not exist either in the Newtonian theory or in general relativity [4)]. Hence, the Einstein distortion of the coordinate grid causes neither physical deformation of the body nor internal stresses in body's matter. However, the Einstein deformation of the coordinate's grid must be taken into account in mathematical calculation of the orbital equations of motion of the body with respect to the global coordinate frame. It is worth mentioning that one has to distinguish the mathematical deformation of the frame caused by the presence of gravitational field from real physical deformation of the body caused by the tidal gravitational field [Xu *et al.*, 2003]. Tidal deformations are due to the curvature of space (second derivatives of the metric tensor) while the Einstein deformation is associated with the affine connection (first derivatives of the metric tensor) [Novikov and Frolov, 1989]. Lorentz and Einstein deformations of the frames can be calculated easily as soon as a precise relativistic theory of reference frames is employed. Alternative theories of gravity operates with additional fields besides the metric tensor making the calculation of the frame deformations more complicated.

Newtonian definition of the internal multipole moments of a body is unique and unambiguous. It follows from the Taylor expansion of the Newtonian gravitational potential in a series of spherical harmonics outside the body [Moritz, 1980; Moritz and Mueller, 1987]. On the other hand, the post-Newtonian definition of the internal multipole moments of the body is coordinate-dependent [Blanchet, 1998; Blanchet and Damour, 1989; Thorne, 1980]. Therefore, the explicit structure of the multipolar expansion of the body's gravitational field crucially depends on the choice of coordinates. The post-Newtonian transformation from the local to global coordinates will change the mathematical description of the multipole moments and the number of multipoles in the expansion. One has to be careful in finding the most adequate formulation of the spherical symmetry of the body to avoid introduction of non-physical multipole moments, which are merely reflecting the gauge freedom of the gravitational field. Any misunderstanding in this question will lead to erroneous derivation of the orbital equations of motion of the bodies from N-body system in the first and higher order post-Newtonian approximations [Futamase and Itoh, 2007; Kopeikin, 1985] and appearance of spurious terms having no physical meaning.

When several bodies form a self-gravitating system they interact to each other and disturb the interior distribution of matter via their tidal field. In the Newtonian physics, this disturbance induces deviation from spherically-symmetric distribution of matter, which leads to appearance of tide-induced multipole moments of the gravitational field of the body. This makes the equations of motion of the bodies different from those of the point-like masses. One can postpone the appearance of the tide-induced multipoles by making the characteristic distance between the bodies large

4) Experiments for detection of the effect of shielding of the gravitational field exclude it with convincing evidence [Caputo, 2006].

enough, thus, reducing the effect of these multipole moments to negligible order [Kopeikin, 1985]. Indeed, the tide-induced quadrupole moment of the body interacts with the tidal field and exerts on the body the orbital force [Alexander, 1973]

$$F_{\text{tide}} \simeq \kappa_{\text{tide}} \left(\frac{v_{\text{e}}}{v_{\text{s}}}\right)^2 \left(\frac{L}{R}\right)^5 F_N \,, \tag{6.42}$$

where $F_N = GM^2/R^2$ is the Newtonian gravity force for a point-like mass, M and L are characteristic mass and size of the bodies, R is the average distance between the bodies, G is the universal gravitational constant, v_{e} is the body's escape velocity, v_{s} is the speed of sound inside the body's interior, and κ_{tide} is a numerical factor depending on the internal distribution of density (equivalent to the tide Love number [Moritz, 1989; Zharkov and Trubitsyn, 1978]). Decreasing the ratio L/R can make the force F_{tide} much smaller than the post-Newtonian force [Damour, 1983; Kopeikin and Vlasov, 2008; Kopeikin, 1985]. The rotational deformation leads to the rotation-induced quadrupole moment of the body, which interacts with the tidal field of external bodies and exerts on the body under consideration the orbital force [Alexander, 1973]

$$F_{\text{rot}} \simeq \kappa_{\text{rot}} \left(\frac{v_{\text{r}}}{v_{\text{s}}}\right)^2 \left(\frac{L}{R}\right)^5 F_N \,, \tag{6.43}$$

where $v_{\text{r}} \simeq \Omega_{B}L$ is the characteristic linear velocity of the body's rotation, Ω_{B} is the angular rotational frequency of the body, and κ_{rot} is a numerical factor depending on the internal distribution of density (equivalent to the rotation Love number [Moritz, 1989; Zharkov and Trubitsyn, 1978]). Making L/R sufficiently small one can neglect F_{rot} [Damour, 1983; Kopeikin and Vlasov, 2008; Kopeikin, 1985].

One stipulates that for each body of N-body system the geometrical center of the body's spherical symmetry is located at the center of mass of the body that coincides with the origin of the local coordinates associated with this body. One also assumes that the Newtonian tides caused by external bodies and the rotational deformation of the body are negligibly small so that all functions characterizing internal structure of the body have spherically-symmetric distribution in the local coordinates. These functions are: the invariant density ρ^*, the internal energy Π, and the tensor of stresses π_{ij}. Spherical symmetry in the local coordinates means that these functions depend only on the local radial coordinate $r = |\boldsymbol{w}|$:

$$\rho^*(u, \boldsymbol{w}) = \rho^*(r) \,, \qquad \Pi(u, \boldsymbol{w}) = \Pi(r) \,, \qquad \pi^{ij}(u, \boldsymbol{w}) = \delta^{ij} p(r) \,. \tag{6.44}$$

Moreover, one assumes that the internal distribution of matter does not depend on the local coordinate time u that excludes radial pulsations of the body from consideration. Radial pulsations can be easily included in this scheme of calculation of the equations of motion as well. Notice that the radial pulsations of matter do not affect the gravitational field (the metric tensor) of a spherically-symmetric distribution of the matter in general relativity due to Birkhoff's theorem [Landau and Lifshitz, 1975; Misner *et al.*, 1973; Novikov and Frolov, 1989; Wald, 1984]. However, our calculations are done in the framework of the scalar-tensor theory of gravity, where the Birkhoff's theorem holds only for time-independent scalar field [Dutta Choudhury and Bhattacharya, 1980; Krori and Nandy, 1977]. Radial pulsations of matter density makes the scalar field time dependent as follows from the post-Newtonian equation (4.58) for scalar field. This should bring extra terms to the post-Newtonian equations of motion of extended, spherically-symmetric bodies.

Spherically-symmetric distribution of matter generates a spherically-symmetric gravitational field both in general relativity and in the scalar-tensor theory of gravity. Therefore, the multipolar expansion of the Newtonian gravitational potential of a spherically-symmetric, extended body must have in the local coordinates only a monopole term

$$\hat{U}_B(u, \boldsymbol{w}) = G \int_{V_B} \frac{\rho^*(u, \boldsymbol{w}') d^3 w'}{|\boldsymbol{w} - \boldsymbol{w}'|} = \frac{G\mathcal{M}_{*B}}{r} \,, \tag{6.45}$$

where the baryon (Newtonian) mass $\mathcal{M}_{*B}$ is defined in equation (6.1). Strictly speaking, this monopole expansion will be violated at some order of approximation because the external tidal force of the background gravitational field acts on the body and deforms its spherically-symmetric distribution of matter. However, equations (6.42) demonstrates that the tidal deformation can be neglected in the first post-Newtonian approximation for sufficiently small ratio of L/R.

One will consider the case of rigidly rotating bodies for which the internal velocity of matter (as defined in the local coordinates) is a vector product of the angular velocity Ω^i_B, referred to the local frame, and the radius-vector w^i, that is

$$v^i = \varepsilon^i{}_{jk} \Omega^j_B w^k \tag{6.46}$$

$$v^2 = \frac{2}{3} \Omega^2_B r^2 - \Omega^j_B \Omega^k_B w^{<jk>} \,, \tag{6.47}$$

in the local frame of the body B. The rotation causes rotational deformation of the body and distorts its spherical symmetry. However, the rotational deformation is proportional to the Love number $\kappa_{\rm rot}$, and by assuming that the body is rigid enough and rotates sufficiently slow, one can make the impact of the rotational deformation on the body's motion negligibly small.

Spherical symmetry of each body assumes that one can use the following (pure mathematical) properties in performing volume's integration of any function $f(r)$, depending on radial coordinate r only [Thorne, 1980]:

$$\int_{V_B} f(r) w^{i_1 i_2 \ldots i_{2l}} d^3 w = \frac{1}{2l+1} \delta_{(a_1 a_2} \ldots \delta_{a_{2l-1} a_{2l})} \int_{V_B} f(r) r^{2l} d^3 w \,, \tag{6.48}$$

$$\int_{V_B} f(r) w^{i_1 i_2 \ldots i_{2l+1}} d^3 w = 0 \,, \tag{6.49}$$

where the integral is taken over the volume of the body B, and $\delta_{(a_1 a_2} \ldots \delta_{a_{2l-1} a_{2l})}$ is the fully symmetric linear combination of the Kronecker delta symbols [Thorne, 1980].

In particular, for any $l \geq 1$ one has

$$\int_{V_B} f(r) w^{<i_1 i_2 \dots i_l>} d^3 w = 0 , \tag{6.50}$$

where the angular brackets around indices denote symmetric and trace-free (STF) part of the multi-index geometric object [Thorne, 1980]. One will also need several other equations for performing integration over the body's volume in the local coordinates. They are as follows:

$$A_{<iL>} B_{<N>} \int_{V_B} \rho^* w^{<L>} w^{<N>} d^3 w = \begin{cases} \dfrac{l!}{(2l+1)!!} A_{<iL>} B_{<L>} \mathcal{I}_B^{(2l)} , & (n = l) \\ 0 , & (n \neq l) \end{cases} \tag{6.51}$$

$$A_{<iL>} \int_{V_B} \rho^* v^2 w^{<L>} d^3 w = \begin{cases} -\dfrac{2}{15} A_{<ijk>} \Omega_B^j \Omega_B^k \mathcal{I}_B^{(4)} , & (l = 2) \\ 0 , & (l > 2) \end{cases} \tag{6.52}$$

where one has used equation (6.47), $A_{<L>}$ and $B_{<L>}$ are arbitrary STF tensors, and

$$\mathcal{I}_B^{(2l)} = \int_{V_B} \rho^* r^{2l} d^3 w, \tag{6.53}$$

is $2l$-th order rotational moment of inertia of the body B [5]. Equations (6.48)-(6.53) will be used for calculation of multipolar expansions of various gravitational potentials entering orbital equations of motion of the bodies.

6.3.2
Coordinate transformation of the multipole moments

Multipolar expansion of the Newtonian potential in the global coordinates, $x^\alpha = (ct, \boldsymbol{x})$, introduces *barycentric* multipole moments of a body defined as integrals over the body's volume taken on hypersurface of constant global coordinate time t, that is

$$\mathbb{I}_B^L = \int_{V_B} \rho^*(t, \boldsymbol{x}) R_B^{i_1} R_B^{i_2} \dots R_B^{i_l} d^3 x , \tag{6.54}$$

where $R_B^i = x^i - x_B^i$, and x_B^i is the origin of the local coordinates coinciding with the center of mass of the body. One has postulated that the density and other structure-dependent functions inside the body have a spherical-symmetric distribution in the local coordinates $w^\alpha = (cu, \boldsymbol{w})$, so that according to equation (6.50) the following relationship holds for any $l \geq 1$

$$\int_{V_B} \rho^*(r') w'^{<L>} d^3 w' = 0 , \qquad (l \geq 1) \tag{6.55}$$

5) Notice that all odd rotational moments $\mathcal{I}^{(2l+1)} = 0$ as follows from the angular integration over a spherical volume of the body.

where the integration across the volume of the body is over a hypersurface of constant local coordinate time u.

Equation (6.55) does not assume that the multipole moments, $\mathbb{I}^L_B$, of the body B defined in the global coordinates are equal to zero, and indeed, $\mathbb{I}^L_B \neq 0$ for $l \geq 1$. One can calculate $\mathbb{I}^L_B$ directly from equation (6.55) after making use of transformation formula, equation (5.38), from the local to global coordinates written down for an element of matter inside the body [6]

$$w'^i = R'^i_B + \epsilon^2 \left[\left(\frac{1}{2} v^i_B v^j_B + F_{ij} + D_{ij} \right) R'^j_B + D_{ijk} R'^j_B R'^k_B \right] + \epsilon^2 \left(v'^i - v^i_B \right) \left(R'^j_B - R^j_B \right) v^i_B + O\left(\epsilon^4\right) , \tag{6.56}$$

where $R'^i_B = x'^i - x^i_B$, $R^i_B = x^i - x^i_B$, $v'^i = dx'^i/dt$, $v^i = dx^i/dt$,

$$F_{ij} = -\varepsilon_{ijk} \mathcal{F}^k , \tag{6.57}$$

$$D_{ij} = \gamma \delta_{ij} \bar{U}(\boldsymbol{x}_B) - \delta_{ij} Y , \tag{6.58}$$

$$D_{ijk} = \frac{1}{2} \left(a^j_B \delta^{ik} + a^k_B \delta^{ij} - a^i_B \delta^{jk} \right) , \tag{6.59}$$

and function $\mathcal{F}^k$ is defined by an ordinary differential equation (5.75).

Equation (6.56) must be used for transforming integrals shown in equation (6.55) from the local to global coordinates. It takes into account that the integration in equation (6.55) is performed over the hypersurface of constant (local) time coordinate u while similar integrals in the multipolar decomposition of the Newtonian gravitational potential in the global coordinates are defined on the hypersurface of constant (global) time coordinate t. Transformation of space coordinates from the spacelike hypersurface of constant time u to that of time t depends on space coordinates, $\boldsymbol{x}$, of the point, at which matching of the local and global coordinates is done so that both coordinates: $\boldsymbol{x}'$ - the point of integration and, $\boldsymbol{x}$ - the matching point, appear in equation (6.56) together and belong to the same hypersurface of constant time t.

Substitution of equation (6.56) into equation (6.55) yields [Kopejkin, 1988b, 1991b]

$$\begin{aligned} \int_{V_B} \rho^*(u, \boldsymbol{w}') w'^{<L>} d^3 w' &= \mathbb{I}^{<L>}_B + \epsilon^2 \left(\frac{l}{2} v^j_B v^{<i_l}_B \mathbb{I}^{L-1>j}_B - l F^{j<i_l} \mathbb{I}^{L-1>j}_B \right. \\ &+ l D^{j<i_l} \mathbb{I}^{L-1>j}_B + l \mathbb{I}^{jk<L-1}_B D^{i_l>jk} + v^j_B \dot{\mathbb{I}}^{j<L>}_B \\ &- \left. \left(v^j_B R^j_B \right) \dot{\mathbb{I}}^{<L>}_B - v^j_B \int_{V_B} \rho^*(u, \boldsymbol{w}') v'^j w'^{<L>} d^3 w' \right) + O\left(\epsilon^4\right) . \end{aligned} \tag{6.60}$$

Taking into account equation (6.55) one concludes that only the dipole, $\mathbb{I}^i_B$, and the quadrupole, $\mathbb{I}^{ij}_B$, barycentric moments differ from zero in the first post-Newtonian

6) Taking into account transformation of the multipole moments from the local to global coordinates is vitally important for correct calculation of equations of motion of extended bodies in the third post-Newtonian approximation as well [Futamase and Itoh, 2007].

approximation. More specifically,

$$\mathbb{I}^i_B = \frac{\epsilon^2}{3}\mathcal{I}^{(2)}_B\left(\varepsilon_{ijk}v^j_B\Omega^k_B + \frac{1}{2}a^i_B\right) + O\left(\epsilon^4\right) , \tag{6.61}$$

$$\mathbb{I}^{<ij>}_B = -\frac{\epsilon^2}{3}\mathcal{I}^{(2)}_B v^{<i}_B v^{j>}_B + O\left(\epsilon^4\right) , \tag{6.62}$$

$$\mathbb{I}^{<L>}_B = O\left(\epsilon^4\right) , \qquad (l \geq 3) \tag{6.63}$$

The same expressions for the multipole moments can be obtained in a different way by making use of multipolar expansions of the Newtonian potential of body B in the local and global coordinates and their subsequent comparison with the help of transformation formula shown in equation (5.82). One has checked that both derivations are self-consistent and yield identical expressions for the barycentric multipole moments given in equations (6.61)-(6.63). Transformation of multipole moments of body's gravitational field from the global to local coordinates were used in papers [Brumberg and Kopeikin, 1989; Brumberg and Kopejkin, 1989a] in order to derive the post-Newtonian equations of motion of Earth's artificial satellite in the geocentric frame with taking into account relativistic corrections due to the presence of Earth's quadrupole field.

Earlier derivations of the post-Newtonian equations of motion of non-rotating extended bodies stipulated that the bodies have spherically-symmetric distribution of mass in the global coordinates (see, for example, [Brumberg, 1972; Dallas, 1977; Fock, 1964; Spyrou, 1975]). This assumption is non-physical because it leads to non-spherical mass distribution inside the body in the local coordinates, which must be maintained by some internal stresses depending on the characteristics of the orbital motion of the body. However, this effect is nothing else but the result of the post-Newtonian coordinate transformation, which can not be associated with physics of the interior structure of the bodies. That is why the earlier assumptions underlying the post-Newtonian calculations of the equations of motion has been abandoned to avoid the apperance of the spurious, coordinate-dependent effects.

Another important point is the center of mass of the body defined in the local coordinates differs from that defined in the global frame as follows from equation (6.61). Earlier works on the post-Newtonian approximations [Brumberg, 1972; Dallas, 1977; Fock, 1964; Spyrou, 1975] including the PPN formalism [Will, 1993] defined the center of mass of each body in the global coordinates, that is equated the dipole moment $\mathbb{I}^i_B$ to zero: $\mathbb{I}^i_B = 0$. This is plausible but brings about in equations of motion of extended bodies terms, which are proportional to the rotational moment of inertia $\mathcal{I}^{(2)}_B$ even if the bodies have spherically-symmetric form. Such terms violate the principle of effacing of the internal structure of the bodies. If they really existed the motion of extended bodies would differ from that of point-like masses (neutron stars or black holes). Definition of the center of mass of the body in terms of its local coordinates allows us to eliminate all non-physical terms depending on the rotational moments of inertia, from the orbital equations of motion [Kopeikin and Vlasov, 2004] and to prove that the effacing principle is valid in the first post-

Newtonian approximation of general relativity (see discussion after equation (6.88)).

6.3.3
Gravitational multipoles in the global coordinates

In order to derive the orbital equations of motion of spherically-symmetric bodies of finite size one needs to know the multipolar decomposition of the gravitational potentials $U_{(B)}(t,\boldsymbol{x})$, $U^i_{(B)}(t,\boldsymbol{x})$, $\Phi_{(B)}(t,\boldsymbol{x})$, and $\chi_{(B)}(t,\boldsymbol{x})$ in the global coordinates. The potentials under discussion are defined in equations (4.75)-(4.81). Multipolar decomposition of the potentials in the global coordinates x^α will have all multipoles different from zero because the spherical symmetry of the bodies is defined with respect to the local coordinates w^α, and the bodies have no spherically-symmetric shape in the global coordinates.

For the Newtonian potential $U_B(t,\boldsymbol{x})$ one has

$$
\begin{aligned}
U_B(t,\boldsymbol{x}) &= \frac{G\mathcal{M}_{*B}}{R_B} + \sum_{l=1}^{\infty}\frac{(-1)^l}{l!}G\mathbb{I}^{<L>}\frac{\partial^L}{\partial x^L}\left(\frac{1}{R_B}\right) \\
&= \frac{G\mathcal{M}_*}{R_B} - G\mathbb{I}^i\frac{\partial}{\partial x^i}\left(\frac{1}{R_B}\right) + \frac{1}{2}G\mathbb{I}^{<ij>}\frac{\partial^2}{\partial x^i\partial x^j}\left(\frac{1}{R_B}\right) + O\left(\epsilon^4\right) ,
\end{aligned}
\tag{6.64}
$$

where dipole, $\mathbb{I}^i$, and quadrupole, $\mathbb{I}^{<ij>}$, moments are given by equations (6.61) and (6.62).

Vector-potential $U^i_{(B)}(t,\boldsymbol{x})$ is decomposed as follows

$$
U^i_{(B)}(t,\boldsymbol{x}) = \frac{G\mathcal{M}_{*B}v^i_B}{R_B} - \frac{1}{3}G\mathcal{I}^{(2)}_B\varepsilon^i{}_{jk}\Omega^j_B\frac{\partial}{\partial x^k}\left(\frac{1}{R_B}\right) + O\left(\epsilon^2\right) , \tag{6.65}
$$

where one has used the fact that inside the body velocity of matter is linearly decomposed in two terms: $v^i = v^i_B + \nu^i + O\left(\epsilon^2\right)$ with v^i_B being velocity of the center of mass of the body and the internal velocity, ν^i, is defined in equation (6.46).

Potential $\chi_{(B)}(t,\boldsymbol{x})$ (sometimes called super-potential [Chandrasekhar, 1965; Chandrasekhar and Lebovitz, 1962; Will, 1993]) has the following multipolar decomposition

$$
\chi_{(B)}(t,\boldsymbol{x}) = -G\mathcal{M}_{*B}R_B - \frac{1}{3}\frac{G\mathcal{I}^{(2)}_B}{R_B} + O\left(\epsilon^2\right) . \tag{6.66}
$$

Post-Newtonian potential $\Phi_{(B)}(t,\boldsymbol{x})$ consists of a linear combination of four functions as shown in equation (4.73). Each of these functions is decomposed in multipoles as follows

$$\Phi_1^{(B)}(t,\boldsymbol{x}) = G\int_{V_B}\frac{\rho^*(t,\boldsymbol{x}')v'^2d^3x'}{|\boldsymbol{x}-\boldsymbol{x}'|} = \frac{G\mathcal{M}_{*B}v_B^2}{R_B} + \frac{G}{R_B}\int_{V_B}\rho^*(r)v^2d^3w + \frac{2G}{3}\frac{\varepsilon_{ijk}R_B^iv_B^j\Omega_B^k\mathcal{I}_B^{(2)}}{R_B^3} - \frac{G}{5}\frac{\Omega^{<i}\Omega^{j>}R_B^iR_B^j\mathcal{I}_B^{(4)}}{R_B^5} + O\left(\epsilon^2\right), \tag{6.67}$$

$$\Phi_2^{(B)}(t,\boldsymbol{x}) = G\int_{V_B}\frac{\rho^*(t,\boldsymbol{x}')U(t,\boldsymbol{x}')d^3x'}{|\boldsymbol{x}-\boldsymbol{x}'|} = \frac{G}{R_B}\int_{V_B}\rho^*(r)\hat{U}_B(r)d^3w + G^2\sum_{A\neq B}\sum_{l=0}^{\infty}\frac{(-1)^l}{(2l+1)l!}\frac{\mathcal{M}_{*A}\mathcal{I}_B^{(2l)}R_{BA}^{<L>}}{R_{BA}^{2l+1}}\frac{\partial^L}{\partial x^L}\left(\frac{1}{R_B}\right) + O\left(\epsilon^2\right), \tag{6.68}$$

$$\Phi_3^{(B)}(t,\boldsymbol{x}) = G\int_{V_B}\frac{\rho^*(t,\boldsymbol{x}')\Pi(t,\boldsymbol{x}')d^3x'}{|\boldsymbol{x}-\boldsymbol{x}'|} = \frac{G}{R_B}\int_{V_B}\rho^*(r)\Pi(r)d^3w + O\left(\epsilon^2\right) \tag{6.69}$$

$$\Phi_4^{(B)}(t,\boldsymbol{x}) = G\int_{V_B}\frac{\pi^{kk}(t,\boldsymbol{x}')d^3x'}{|\boldsymbol{x}-\boldsymbol{x}'|} = \frac{3G}{R_B}\int_{V_B}p(r)d^3w + O\left(\epsilon^2\right). \tag{6.70}$$

This concludes the set of equations describing the multipolar decomposition of the gravitational potentials in the global coordinates.

One noticess that the multipolar decomposition of all the potentials contain dependence on the rotational moments of inertia $\mathcal{I}_B^{(2)}$, $\mathcal{I}_B^{(4)}$, etc. Appearance of the moments of inertia in the multipolar decompositions may indicate violation of the principle of effacing of the internal structure of the bodies. However, the moments of inertia of the second order $\mathcal{I}_B^{(2)}$ vanish in the final form of the orbital equations of motion in general relativity. The only contribution to these equations comes from the moments of inertia of the forth order $\mathcal{I}_B^{(4)}$. Physically it happens because the Einstein field equations are non-linear and in case of N-body system the non-linearity leads to the coupling of the finite volume of the body with the external (tidal) gravitational field of the other bodies of the N-body system [Kopeikin and Vlasov, 2004; Nordtvedt, 1994].

6.3.4
Orbital post-Newtonian equations of motion

Both conditions of the spherical symmetry and rigid rotation, equations (6.44) and (6.46), allow us to simplify equation (6.29) of the orbital motion of the center of mass of body B drastically. For example, equation (6.50) assumes that STF mass-type multipole moments of the body $\mathcal{I}^L = O\left(\epsilon^2\right)$ for any index $l \geq 1$, and STF current-type multipoles $\mathcal{S}^L = O\left(\epsilon^2\right)$ for any index $l \geq 2$. Therefore, calculation of the tidal Newtonian force $\mathbb{F}^i_N$ from equation (6.29), yields

$$\mathbb{F}^i_N = \epsilon^2 \left\{ \frac{2\gamma+1}{30} Q_{ijk}\Omega^j_B\Omega^k_B \mathcal{I}^{(4)}_B - \sum_{l=1}^{\infty} \frac{2(1-\beta)P_L - Q_L}{l!(2l+1)!!} Q_{iL}\mathcal{I}^{(2l)}_B \right\} , \tag{6.71}$$

which has the post-Newtonian, ($\sim \epsilon^2$), order of magnitude.

The post-Newtonian gravitomagnetic tidal force $\mathbb{F}^i_{pN}$ from equation (6.29) is reduced to a simple expression

$$\mathbb{F}^i_{pN} = -\frac{3}{4} C_{ij} \mathcal{S}^j_B , \tag{6.72}$$

where the external (gravitomagnetic-type) quadrupole

$$\begin{aligned} C_{ij} = & - \frac{20(1+\gamma)G}{3} \sum_{C\neq B} \mathcal{I}^{(2)}_C \frac{\Omega^p_C R^{<ijp>}_{BC}}{R^7_{BC}} \\ & + 2(1+\gamma)G \sum_{C\neq B} \frac{\mathcal{M}_C (v^p_C - v^p_B)}{R^5_{BC}} \left(\varepsilon_{ipq} R^{<jq>}_{BC} + \varepsilon_{jpq} R^{<iq>}_{BC} \right) , \end{aligned} \tag{6.73}$$

$\mathcal{M}_B$ is the mass, and

$$\mathcal{S}^i_B = \frac{2}{3} \mathcal{I}^{(2)}_B \Omega^i_B , \tag{6.74}$$

is the spin of the body B.

All other post-Newtonian terms in equation (6.29) depending explicitly on $\mathcal{I}^L$, are equal to zero for spherically-symmetric and rigidly rotating bodies. Hence, equation of motion (6.29) is drastically simplified and reduces to

$$\begin{aligned} \tilde{\mathcal{M}}_B a^i_B = & \mathcal{M}_B \bar{V}_{,i}(\boldsymbol{x}_B) - \mathbb{F}^i_N \\ + \epsilon^2 & \mathcal{M}_B \Big\{ \Big[\gamma\delta_{ik} v^2_B - v^i_B v^k_B - 2(\gamma+\beta)\delta_{ik}\bar{U}(\boldsymbol{x}_B) \Big] \bar{U}_{,k}(\boldsymbol{x}_B) \\ & + 2(1+\gamma)\dot{\bar{U}}^i(\boldsymbol{x}_B) - 2(1+\gamma) v^k_B \bar{U}^{k,i}(\boldsymbol{x}_B) - (1+2\gamma) v^i_B \dot{\bar{U}}(\boldsymbol{x}_B) \Big\} \\ & + \epsilon^2 \left[\frac{1}{2}(1-\gamma)\varepsilon_{ipk} \mathcal{S}^p \dot{\bar{U}}_{,k}(\boldsymbol{x}_B) + \frac{3}{4} C_{ij} \mathcal{S}^j \right] + O(\epsilon^4) , \end{aligned} \tag{6.75}$$

where the *conformal*, $\tilde{\mathcal{M}}_B$, and *active*, $\mathcal{M}_B$, masses od the body B are related to each other via equation (6.11), that is

$$\tilde{\mathcal{M}}_B = \mathcal{M}_B + \epsilon^2 \left[\frac{\eta}{2} \int_{V_B} \rho^* \hat{U}_B d^3w + 2(\beta-1) \sum_{C\neq B} \frac{G\mathcal{M}_C \mathcal{M}_B}{R_{CB}} \right] , \tag{6.76}$$

and the gravitational potential

$$\bar{V}(\boldsymbol{x}) = \bar{U}(\boldsymbol{x}) + \epsilon^2 \left[\bar{\Phi}(\boldsymbol{x}) - \frac{1}{2} \bar{\chi}_{,tt}(\boldsymbol{x}) \right] . \tag{6.77}$$

The tidal force $\mathbb{F}^i_N$ is given by equation (6.71), and C_{ij} is shown in equation (6.73).

Let us now calculate all terms in the right side of equation (6.75) explicitly in terms of body's mass, rotational moment of inertia, and spin. Among them, the most complicated is the first one, that is gradient $\bar{V}_{,i}(\boldsymbol{x}_B)$. By making use of equations (6.64)-(6.70) one obtains

$$\begin{aligned}\bar{V}(t,\boldsymbol{x}) &= \sum_{C\neq B}\frac{G\mathcal{M}_C}{R_C}\left\{1+\epsilon^2\left[(\gamma+1)v_C^2-\frac{1}{2}a_C^kR_C^k-\frac{\left(v_C^kR_C^k\right)^2}{2R_C^2}-\gamma\sum_{D\neq C}\frac{GM_D}{R_{CD}}\right]\right\} \quad (6.78)\\ &+\epsilon^2 \quad G\sum_{C\neq B}\left\{\frac{1}{3}\mathcal{I}_C^{(2)}\frac{R_C^k}{R_C^3}\left[2(1+\gamma)\varepsilon_{kpq}v_C^p\Omega_C^q+a_C^k\right]-\frac{1+2\gamma}{10}\frac{R_C^jR_C^k}{R_C^5}\Omega_C^{<j}\Omega_C^{k>}\mathcal{I}_C^{(4)}\right.\\ &\left.+(1-2\beta)\sum_{l=1}^{\infty}\frac{(2l-1)!!}{(2l+1)l!}\mathcal{I}_C^{(2l)}\frac{R_C^{<L>}}{R_C^{2l+1}}\sum_{D\neq C}GM_D\frac{R_{CD}^{<L>}}{R_{CD}^{2l+1}}\right\}+O(\epsilon^4),\end{aligned}$$

where

$$\mathcal{M}_C = M_C-\epsilon^2\left[\frac{\eta}{2}\int_{V_C}\rho^*\hat{U}_Cd^3w+(2\beta-\gamma-1)\sum_{D\neq C}\frac{GM_CM_D}{R_{CD}}\right], \quad (6.79)$$

is the *active* mass of body C, M_C is the general relativistic mass of the body C defined by equation (6.8), where for the sake of simplicity one assumesd $Y=0$, and $\eta = 4\beta-\gamma-3$ is the Nordtvedt parameter.

After calculating derivatives from potentials $\bar{V}(t,\boldsymbol{x})$, $\bar{U}(t,\boldsymbol{x})$, and substituting them into equation (6.75) one obtains the following expression for acceleration of the center of mass of body B:

$$M_Ba_B^i = F_N^i+\epsilon^2\left\{F_{EIH}^i++F_S^i+F_{\mathcal{I}GR}^i+F_{\mathcal{I}ST}^i\right\}+O(\epsilon^4)\,, \quad (6.80)$$

where F_N^i is the Newtonian force, and F_{EIH}^i, F_Ω^i, $F_{\mathcal{I}GR}^i$, $F_{\mathcal{I}ST}^i$ are the post-Newtonian forces of gravity. They are given by the following expressions [Kopeikin and Vlasov, 2004]

$$F^i_{N} = -\sum_{C\neq B}\frac{G\mathfrak{M}_B\mathfrak{M}_C R^i_{BC}}{R^3_{BC}}\,, \tag{6.81}$$

$$\begin{aligned}F^i_{EIH} = &-\sum_{C\neq B}\frac{G\mathrm{M}_B\mathrm{M}_C R^i_{BC}}{R^3_{BC}}\Big\{\gamma v_B^2 - 2(1+\gamma)(\boldsymbol{v}_B\cdot\boldsymbol{v}_C) + (1+\gamma)v_C^2\\
&-\frac{3}{2}\left(\frac{\boldsymbol{R}_{BC}\cdot\boldsymbol{v}_C}{R_{BC}}\right)^2 - (1+2\gamma+2\beta)\frac{G\mathrm{M}_B}{R_{BC}} - 2(\gamma+\beta)\frac{G\mathrm{M}_C}{R_{BC}}\\
&+\sum_{D\neq B,C}\left[(1-2\beta)\frac{G\mathrm{M}_D}{R_{CD}} - 2(\gamma+\beta)\frac{G\mathrm{M}_D}{R_{BD}} + \frac{G\mathrm{M}_D(\boldsymbol{R}_{BC}\cdot\boldsymbol{R}_{CD})}{2R^3_{CD}}\right]\Big\}\\
&+\sum_{C\neq B}\Big\{\frac{G\mathrm{M}_B\mathrm{M}_C(v^i_B - v^i_C)}{R^3_{BC}}\Big[2(1+\gamma)(\boldsymbol{v}_B\cdot\boldsymbol{R}_{BC}) - (1+2\gamma)(\boldsymbol{v}_C\cdot\boldsymbol{R}_{BC})\Big]\\
&-\frac{3+4\gamma}{2}\frac{G\mathrm{M}_B\mathrm{M}_C}{R_{BC}}\sum_{D\neq B,C}\frac{G\mathrm{M}_D R^i_{CD}}{R^3_{CD}}\Big\}\,,\end{aligned} \tag{6.82}$$

$$\begin{aligned}F^i_{S} = &\;G\sum_{C\neq B}\Big\{\frac{\mathrm{M}_C\mathcal{S}^p_B(v^k_B - v^k_C)}{2R^5_{BC}} + \Big[3(1+\gamma)\left(\varepsilon_{kpq}R^{<iq>}_{BC} - \varepsilon_{ikq}R^{<pq>}_{BC}\right)\\
&+3(1-\gamma)\varepsilon_{ipq}R^{<kq>}_{BC}\Big] - 3(1+\gamma)\frac{\mathrm{M}_B\mathcal{S}^p_C(v^k_B - v^k_C)}{R^5_{BC}}\Big[\varepsilon_{ipq}R^{<kq>}_{BC} - \varepsilon_{kpq}R^{<iq>}_{BC}\Big]\\
&+\frac{15(1+\gamma)}{2}\frac{\mathcal{S}^j_B\mathcal{S}^k_C R^{<ijk>}_{BC}}{R^7_{BC}} + \left(\gamma+\frac{1}{2}\right)\frac{R^{<ijk>}_{BC}}{R^7_{BC}}\Big[\mathrm{M}_B\mathcal{I}^{(4)}_C\Omega^j_C\Omega^k_C + \mathrm{M}_C\mathcal{I}^{(4)}_B\Omega^j_B\Omega^k_B\Big]\Big\}\,,\end{aligned} \tag{6.83}$$

$$\begin{aligned}F^i_{\mathcal{I}GR} = &\;G^2\sum_{C\neq B}\sum_{l=2}^{\infty}\frac{(2l-1)!!}{l!}\Big[(-1)^l\mathrm{M}_B\mathcal{I}^{(2l)}_C\frac{R^{<iL>}_{BC}}{R^{2l+3}_{BC}}\sum_{D\neq C}\frac{\mathrm{M}_D R^{<L>}_{CD}}{R^{2l+1}_{CD}}\\
&+\mathrm{M}_C\mathcal{I}^{(2l)}_B\frac{R^{<L>}_{BC}}{R^{2l+1}_{BC}}\sum_{D\neq B}\frac{\mathrm{M}_D R^{<iL>}_{BD}}{R^{2l+3}_{BD}}\Big]\,,\end{aligned} \tag{6.84}$$

$$\begin{aligned}F^i_{\mathcal{I}ST} = &\;2(\beta-1)G^2\sum_{C\neq B}\Big\{\mathrm{M}_C\mathcal{I}^{(2)}_B\frac{R^k_{BC}}{R^3_{BC}}\sum_{D\neq B}\frac{\mathrm{M}_D R^{<ik>}_{BD}}{R^5_{BD}}\\
&-\mathrm{M}_B\mathcal{I}^{(2)}_C\frac{R^{<ik>}_{BC}}{R^5_{BC}}\sum_{D\neq C}\frac{\mathrm{M}_D R^k_{CD}}{R^3_{CD}} + \sum_{l=2}^{\infty}\frac{(2l-1)!!}{l!}\Big[(-1)^l\mathrm{M}_B\mathcal{I}^{(2l)}_C\frac{R^{<iL>}_{BC}}{R^{2l+3}_{BC}}\sum_{D\neq C}\frac{\mathrm{M}_D R^{<L>}_{CD}}{R^{2l+1}_{CD}}\\
&+\mathrm{M}_C\mathcal{I}^{(2l)}_B\frac{R^{<L>}_{BC}}{R^{2l+1}_{BC}}\sum_{D\neq B}\frac{\mathrm{M}_D R^{<iL>}_{BD}}{R^{2l+3}_{BD}}\Big]\Big\}\,,\end{aligned} \tag{6.85}$$

where one has defined the coordinate distance between two bodies $R^i_{BC} = x^i_B - x^i_C$,

$R_{BC} = |\boldsymbol{x}_B - \boldsymbol{x}_C|$, and spin $\mathcal{S}^i_B$ of body B relates to the angular speed of its own rotation Ω^i_B by equation (6.74). Equation (6.80) elucidates that *inertial* mass M_B of body B is equal to its general-relativistic mass given by equation (6.8). This mass is conserved (constant) for spherically-symmetric bodies as follows from equation (6.12) because the internal multipole moments $\mathcal{I}_L = O(\epsilon^2)$ so that the right side of the equation is the post-post-Newtonian order of magnitude, which is neglected. The *gravitational* mass $\mathfrak{M}_B$ of the body B depends on the gravitational self-energy of the body coupled with the Nordtvedt parameter $\eta = 4\gamma - \beta - 3$

$$\mathfrak{M}_B = \mathrm{M}_B - \frac{1}{2}\epsilon^2\eta \int_{V_B} \rho^* \hat{U}_B d^3w \ . \tag{6.86}$$

The Newtonian gravitational force F^i_N given by equation (6.81), depends in the scalar-tensor theory only on the *gravitational* masses of the bodies. Will [1993] distinguishes the "active" and "passive" *gravitational* masses, which depend in the PPN formalism on the whole bunch of the PPN parameters. The present book operates only with two PPN parameters, β and γ but, unlike the PPN formalism, the definitions adopted and derivations given here are fully covariant. In the case of the two PPN parameters the "active" and "passive" *gravitational* masses are indistinguishable and are given by one and the same expression (6.86). Nevertheless, the *inertial* and *gravitational* masses of each body are not equal in the scalar-tensor theory of gravity due to violation of the strong principle of equivalence for massive bodies [Dicke, 1962a; Nordtvedt, 1968a,b]. This violation can be explained as due to the interaction of the gravitational self-energy of the body under consideration with the monopole moment of the scalar field generated by external bodies. This interaction leads to a local gravitational force, which brings about a non-zero value of the time derivative of the body's linear momentum in the local coordinates. Indeed, assuming that the body under consideration has finite size, does not rotate, and is spherically-symmetric, one obtains from equation (6.18)

$$\begin{aligned} \dot{\mathcal{P}}^i &= \mathcal{M}Q_i\left(1 + \frac{1}{2\mathcal{M}}\epsilon^2\eta \int_{V_B} \rho^* \hat{U}^{(B)} d^3w\right) - \\ &\quad \frac{1}{2}\epsilon^2\eta P_i \int_{V_B} \rho^* \hat{U}^{(B)} d^3w - \mathbf{F}^i_N + O(\epsilon^4) \ , \end{aligned} \tag{6.87}$$

where P_i is the gradient of the scalar potential generated by the external bodies. If one keeps the body's center of mass at the origin of the local coordinate system ($\dot{\mathcal{P}}^i = 0$), then, the local frame is not inertial and the body's center of mass experiences acceleration $Q_i \neq 0$. This acceleration is due to the interaction of the gravitational self-energy ("gravitational charge") of the body under consideration with the gradient of the external scalar potential. The coupling constant characterizing this scalar interaction is the dimensionless Nordtvedt parameter η. The post-Newtonian gravity forces defined by equations (6.82)-(6.85), depend in this approximation only on the Newtonian mass given by equation (6.1). In the higher post-Newtonian approximations the masses will definitely include relativistic corrections but their exact structure in the scalar-tensor theory of gravity is not yet known. As for general relativity, the masses entering the post-Newtonian force of gravity are the general-relativistic masses. This was proved in papers [Kopeikin, 1985; Kopejkin, 1988b] where the two-body problem accounting for conservative post-Newtonian forces and the gravitational radiation-reaction force was analyzed. An alternative proof was proposed by Damour [1987].

The post-Newtonian force (6.82) is known as the Einstein-Infeld-Hoffmann (EIH) force [Soffel, 1989; Will, 1993]. It is the main post-Newtonian force used for calculation of JPL ephemerides of the solar system bodies [Seidelmann and Urban, 2010]. It was derived in general relativity by Einstein *et al.* [1938] for the case of point-like massive particles viewed as singularities of the metric tensor on the spacetime manifold. Lorentz and Droste [1937] obtained almost the same result two decades earlier by doing comprehensive calculations for homogeneous and spherically-symmetric bodies of constant density. Petrova [1949] and Fock [1964] re-derived EIH equations for extended bodies composed of a perfect fluid without making any restrictions on their internal structure besides the assumptions of the slow-motion of matter and the weak gravity field inside the bodies. Infeld and Plebanski [1960] and Landau and Lifshitz [1975] derived EIH equations from the variational principle with the energy-momentum tensor of matter in the form of delta functions. In scalar-tensor theory of gravity the analogue of EIH force was derived by Estabrook [Estabrook, 1969] in case of $\beta = 1, \gamma \neq 1$, and by Dallas [Dallas, 1977] in the case $\beta \neq 1, \gamma \neq 1$ (see also [Vincent, 1986]). These derivations had stipulated that the bodies have negligible ratio of their radii to the characteristic distance between them (a point-like approximation) as well as that they are moving along geodesic worldlines in "effective" spacetime manifold.

Various post-Newtonian corrections to the EIH force are given by equations (6.83) -(6.85). The force F^i_S, given by equation (6.83), describes the force due to the coupling of body's spin with the orbital angular momentum and with spins of other bodies. It depends on the PPN parameter γ only. If one takes $\gamma = 1$ in equation (6.83), the force F^i_S is reduced exactly to its general relativistic expression obtained earlier by other researchers [Barker and O'Connell, 1976; Brumberg, 1972; Damour *et al.*, 1992; Xu *et al.*, 1997]. One notices that equation (6.83) for the PPN force F^i_S coincides with that derived by Klioner and Soffel [Klioner and Soffel, 2000].

The force (6.84) describes general-relativistic correction to the EIH force due to the finite size of the bodies. It starts from the forth-order rotational moments of inertia of the bodies, $\mathcal{I}^{(4)}$, while all terms, which are proportional to the second-order body's rotational moments of inertia, $\mathcal{I}^{(2)}$, canceled out and do not appear in the equation. Nordtvedt [Nordtvedt, 1994] considered the problem of translational motion of extended bodies in the general class of scalar-tensor theories of gravity. He argued that covariant formulation of the variational principle admits to have terms in the gravitational Lagrangian being proportional to the second-order moment of inertia of extended body coupled with the Ricci tensor $R_{\alpha\beta}$ of the background gravitational field. However, such terms must disappear in general relativity by virtue of the Einstein equations in vacuum: $R_{\alpha\beta} = 0$. Nonetheless, the body's moments of inertia of higher order must couple with the full Riemann tensor $R_{\alpha\beta\gamma\delta}$ and its derivatives, which do not vanish in vacuum. For this reason they can be present in general-relativistic equations of motion of spherically-symmetric bodies as, indeed, demonstrated in equation (6.84). One notices that vanishing of all terms depending on the second-order rotational moment of inertia in general relativity is in disagreement with calculations by Brumberg [1972]; Dallas [1977]; Spyrou [1975] and Vincent [1986], who came to the conclusion that the general relativistic Lagrangian for the system of N spherically-symmetric bodies must depend on the second-order moments of inertia of these bodies, $\mathcal{I}^{(2)}$. Brumberg's expression for the force due to the finite size of the bodies is [Brumberg, 1972]

$$
\begin{aligned}
F^i_{\text{Brumberg}} &= \epsilon^2 G \sum_{C\neq B}\left\{\frac{5\left(\mathrm{M}_B\mathcal{I}^{(2)}_C+\mathrm{M}_C\mathcal{I}^{(2)}_B\right)v^j_C v^k_C R^{<ijk>}_{BC}}{2R^7_{BC}}\right.\\
&+\frac{G\left(\mathrm{M}_B\mathrm{M}_C\mathcal{I}^{(2)}_B-\mathrm{M}^2_B\mathcal{I}^{(2)}_C-2\mathrm{M}^2_C\mathcal{I}^{(2)}_B\right)R^i_{BC}}{3R^6_{BC}}\\
&-\frac{G^2}{2}\sum_{D\neq B,C}\mathrm{M}_C\left[\frac{\mathrm{M}_B\mathcal{I}^{(2)}_D R^k_{CD}R^{<ik>}_{BD}}{R^3_{CD}R^5_{BD}}\right.\\
&\left.\left.+\frac{\mathrm{M}_D\mathcal{I}^{(2)}_B}{R^3_{BC}}\left(\frac{R^k_{CD}R^{<ik>}_{BC}}{R^3_{CD}R^2_{BC}}+\frac{R^k_{BD}R^{<ik>}_{BC}}{R^3_{BD}R^2_{BC}}+\frac{R^k_{BC}R^{<ik>}_{BD}}{R^5_{BD}}\right)\right]\right\}\,,
\end{aligned}
\tag{6.88}
$$

The result (6.84) apparently disagrees with Brumberg's calculations. The reason is that Brumberg [1972] followed Fock's method [Fock, 1964] and used definitions of the multipole moments of extended bodies given in the global (barycentric) coordinates of N-body system. Spherical symmetry of the bodies was also defined by Brumberg in the global coordinates. This definition of the spherical symmetry does not take into account the Lorentz and Einstein contraction of the coordinate volumes of the bodies that is introduced by the law of transformation between the local and global coordinates. A moving body can maintain spherically-symmetric distribution of mass in the global coordinates, if and only if, there are internal stresses inside the bodies that can compensate for the coordinate-dependent Lorentz and Einstein contractions of body's shape [Kopejkin, 1988b]. Existence of such coordinate-dependent internal stresses inside the moving body is unnatural as they are governed by the choice of coordinates which can be defined arbitrary. It indicates that the force F^i_{Brumberg} is spurious and should vanish after appropriate redefinition of the center of mass and the quadrupole moment of the bodies.

In order to prove that the force F^i_{Brumberg} has no physical origin, one has considered translational equation of motion for body B defined in the global frame as follows [Brumberg, 1972; Fock, 1964]

$$
\int_{V_B}\rho^*\frac{dv^i}{dt}d^3x = \int_{V_B}\left(\frac{\partial\pi_{ij}}{\partial x^j}-\rho^*\frac{\partial U}{\partial x^i}\right)d^3x+\epsilon^2F^i_{EIH}+\epsilon^2F^i_{\text{Brumberg}}\,,
\tag{6.89}
$$

where the post-Newtonian EIH and Brumberg forces were taken into account. Velocity v^i of the body's matter in the global coordinates is obtained by differentiation with respect to time of both sides of equation (6.56) with respect to time u. The result can be decomposed in the orbital velocity of the body's center of mass, v^i_B, and the internal velocity of the matter in the local coordinates

$$v^i = v^i_B(t) + v^i(u, \mathbf{w}) + \epsilon^2 \Delta v^i(u, \mathbf{w}) , \tag{6.90}$$

where Δv^i is the post-Newtonian correction, which is a quadratic function of the local coordinates w^i of the body [Kopeikin and Vlasov, 2004]. Subsequent calculation of the time derivative of v^i and calculation of the integral in the left side of equation (6.89) bring about terms, which depend on the moment of inertia of body B. This moment of inertia is reduced to the rotational moment of inertia $\mathcal{I}^{(2)}_B$ in case of a spherically-symmetric body. Calculation of the integral from the Newtonian potential in the right side of equation (6.89) is done by splitting the potential in two parts - internal and external (see equation (5.28)), and applying equations (6.64) and (6.61), (6.62) for calculation of the integrals from the external potential $\bar{U}$. This again gives a number of terms depending on the rotational moment of inertia $\mathcal{I}^{(2)}_B$ of the body B. Summing up all those terms, one obtains exactly the same expression as in equation (6.88). It means that these terms cancel out with the force $F^i_{\rm Brumberg}$ in equation (6.89) so that the rotational moments of inertia of the second order do not enter the post-newtonian equations of motion in general relativity.

This completely agrees with calculation of the force $F^i_{\mathcal{I}GR}$, which does not depend on the body's rotational moments of inertia of the second order either. One concludes that the origin of the coordinate-dependent force $F^i_{\rm Brumberg}$ is directly associated with an inappropriate choice of the body's center of mass and its spherical symmetry, which must be defined not in the global but in local coordinates comoving with the body. Brumberg attempted to make more physical calculation of the force $F^i_{\mathcal{I}GR}$ in [Brumberg, 1991] but he did not arrive to any definite conclusion regarding whether the force $F^i_{\mathcal{I}GR}$ depends on the body's rotational moments of inertia of the second order or not. Our calculation in this book resolves the problem and demonstrates that the force $F^i_{\rm Brumberg}$ is coordinate-dependent and has no physical impact on the orbital motion of the bodies having finite size. General-relativistic correction due to the finite size of the moving bodies is proportional to the forth- and higher-order rotational moments of inertia of the bodies. These corrections are extremely small for the bodies comprising the solar system and can be neglected for any kind gravitational experiments. However, the finite-size effects can become important during the final stage of coalescence of binary neutron stars so that they should be included in the precise calculation of templates of gravitational waves.

The force $F^i_{\mathcal{I}ST}$ describes relativistic correction due to the finite size of the bodies in the scalar-tensor theory of gravity. This force is proportional to the parameter $\beta - 1$ only and, in contrast to general relativity, depends on the second order rotational moments of inertia, $\mathcal{I}^{(2)}_B$. This dependence was noticed by Nordtvedt [Nordtvedt, 1971, 1994] who has found that in the case of weakly, self-gravitating bodies the finite-size effects are proportional to $\eta = 4\beta - \gamma - 3$. This seems to be in disagreement with our calculation of the force $\delta F^i_{\mathcal{I}GR}$ but one can easily reconcile the two formulations. The matter is that Nordtvedt [1971] worked in harmonic coordinates defined by the condition $\partial_\alpha(\sqrt{-g} g^{\alpha\beta}) = 0$ while this book employs the quasi-harmonic coordinates defined by the Nutku condition $\partial_\alpha(\phi\sqrt{-g} g^{\alpha\beta}) = 0$ (see equation (4.45) in the present book). This leads to two different forms of the transformation between the spatial global and local coordinates. In harmonic coordinates this transformation reads [Klioner and Soffel, 2000; Nordtvedt, 1971]

$$\begin{aligned} w^i_{\rm harmonic} &= R^i_B + \epsilon^2\Bigg[\left(\frac{1}{2}v^i_B v^k_B + \gamma\delta^{ik}\bar{U}(\boldsymbol{x}_B) + F^{ik}\right)R^k_B \\ &\quad + \gamma a^k_B R^i_B R^k_B - \frac{\gamma}{2}a^i_B R^2_B\Bigg] + O(\epsilon^4)\,, \end{aligned} \tag{6.91}$$

while in the quasi-harmonic coordinates used in the present book, one has

$$\begin{aligned} w^i &= R^i_B + \epsilon^2\Bigg[\left(\frac{1}{2}v^i_B v^k_B + \gamma\delta^{ik}\bar{U}(\boldsymbol{x}_B) + F^{ik}\right)R^k_B \\ &\quad + a^k_B R^i_B R^k_B - \frac{1}{2}a^i_B R^2_B\Bigg] + O(\epsilon^4)\,. \end{aligned} \tag{6.92}$$

The two transformations have different dependence on γ in terms being proportional to the acceleration so that the difference between the two types of coordinates is

$$w^i_{\rm harmonic} = w^i + (\gamma-1)\epsilon^2\left(a^k_B R^i_B R^k_B - \frac{1}{2}a^i_B R^2_B\right)\,. \tag{6.93}$$

It is due to this difference that PPN parameter γ had appeared in Nordtvedt's calculations and led to the appearance of parameter η in Nordtvedt's final equation for the finite-size effects. However, the dependence of the magnitude of the finite-size effects on parameter γ in Nordtvedt's calculations is a pure coordinate effect, which has no physical meaning. Parameter γ can be eliminated from the force $\delta F^i_{\mathcal{IGR}}$ if one works in the quasi-harmonic coordinates defined by the Nutku condition (4.45) of this paper. Only parameter β remains relevant.

6.3.5
Rotational equations of motion

Derivation of rotational equations of motion for spherically-symmetric bodies requires calculation of the multipole moments $\mathcal{R}_L$ (see equation (6.37)) of the body under consideration. One has

$$\mathcal{R}_L = O\left(\epsilon^2\right)\,, \tag{6.94}$$

which means that $\mathcal{R}_L$ is of the post-Newtonian order of magnitude for spherically-symmetric bodies. All other multipole moments have been calculated in the previous section. Performing calculations of the torques and body's spin, given by equations (6.34)-(6.36), one obtains

$$\mathcal{T}^i = \epsilon^2\left[\frac{2\gamma+1}{15}\varepsilon_{ijk}Q_{jn}\Omega^k_{(B)}\Omega^n_{(B)}\mathcal{I}^{(4)}_{(B)} + \varepsilon_{ijk}\mathcal{S}^j C_k\right]\,, \tag{6.95}$$

$$\Delta\mathcal{T}^i = 0\,, \tag{6.96}$$

$$\Delta\mathcal{S}^i = \frac{2}{3}\mathcal{I}^{(2)}_{(B)}C_i + \left[Q - Y + (\gamma-1)\bar{U}(\boldsymbol{x}_B)\right]\mathcal{S}^i\,. \tag{6.97}$$

Consequently, the rotational equation of motion for the body's spin is

$$\frac{d\mathcal{S}^i_+}{du} = \epsilon^2\left(\frac{2\gamma+1}{15}\varepsilon_{ijk}Q_{jn}\Omega^k_{(B)}\Omega^n_{(B)}\mathcal{I}^{(4)}_{(B)} + \varepsilon_{ijk}\mathcal{S}^j C_k\right) + O(\epsilon^4)\,, \tag{6.98}$$

where C_k is angular velocity of rotation of the local coordinate frame with respect to that which axes are subject to the Fermi-Walker transport.

Equation (6.98) has one extra term comparatively with the corresponding equation (9.75) from the paper [Klioner and Soffel, 2000] by Klioner and Soffel. This term depends on the forth-order rotational moment of inertia, $\mathcal{I}^{(4)}_{B}$, of the body B and has pure general-relativistic origin. This term was not taken into account by Klioner and Soffel because they neglected finite size of the rotating body. Contribution of the forth-order rotational moment of inertia, $\mathcal{I}^{(4)}_{B}$ to the rotational torque is negligibly small for the theory of Earth's rotation. However, it may become significant during last several orbits of a coalescing binary neutron star. It would be interesting to study the impact of this term on the forms of gravitational waves emitted by such binaries.

6.4 Post-Newtonian Two-Body Problem

6.4.1 Introduction

Post-Newtonian equations of motion (6.80)-(6.85) in N-body problem are too complicated in general case for analytic integration and one has to resort to numerical methods for their solution. This is especially important for construction of ephemerides of the solar system bodies - planets and Sun. Numerical ephemerides are discussed in Chapter 9. This section is dedicated to a two-body (Kepler) problem at the first post-Newtonian approximation under assumption that the bodies are spherically-symmetric (tides are ignored). One will also ignore the finite-size effects induced by the rotational moments of inertia and spin, which effectively reduce equations (6.80)-(6.85) to the EIH equations of motion of two point-like masses possessing analytically-tractable solution. Spin and finite size of the post-Newtonian effects in the orbital motion of two bodies can be included in the consideration by making use of the Hamiltonian theory [Barker *et al.*, 1981; Barker and O'Connell, 1976, 1987] and the method proposed by Gergely *et al.* [2006]. This, however, makes analysis too complicated going beyond the scope of the present book.

In many practical situations one body is assumed to have a mass that is negligible compared to the mass of the other body. This is a good approximation for the case of a planet revolving around the Sun, or a photon passing by a star. In such cases, one may assume that only the heavier body contributes to the curvature of spacetime and that it is fixed in space. This curved spacetime is described in general relativity by the Schwarzschild solution of the vacuum Einstein equations [Landau and Lifshitz, 1975; Misner *et al.*, 1973; Wald, 1984]. The motion of the lighter body is described by the spacetime geodesic of the Schwarzschild solution. These geodesic solution accounts for the anomalous precession of the planet Mercury, which is a key piece of experimental evidence supporting the theory of general relativity [Brumberg, 1972; Soffel, 1989; Will, 1993]. The null geodesic also describes the deflection of light in a gravitational field - another prediction, famously used as evidence for general relativity. However, the orbital decay of a gravitationally-bound binary system due to emission of gravitational radiation is not described by geodesics of the Schwarzschild solution and the mass of each body should be taken into account [Damour, 1987, 2000].

Solution of the post-Newtonian Kepler problem in case of two comparable masses had become vitally important for relativistic astrophysics after discovery of a binary pulsar B1913+16 [Hulse and Taylor, 1975] and, especially, a double pulsar J0737-3039 [Possenti *et al.*, 2004]. This is because the timing of binary and double pulsars plays at present the crucial role in testing alternative theories of gravitation in the strong-field limit of the gravitational field of compact relativistic objects - neutron stars or black holes [Damour and Taylor, 1992; Kramer *et al.*, 2006; Lorimer and Kramer, 2004]. Four basic timing models have been used for processing the observational data [Taylor and Weisberg, 1989]. These models have been originally developed by Blandford and Teukolsky [1976]; Epstein [1977]; Haugan [1985], and improved later on by Damour and Deruelle [1985, 1986]. Parametrization of binary pulsar orbit considered by Blandford & Teukolsky (BT) is the Keplerian motion with the only relativistic effect describing a linear (with respect to time) drift of pericenter of the orbit. Parametrization of Epstein, as corrected by Haugan (EH), accounts for all post-Newtonian orbital effects. The models of Epstein [1977] and Haugan [1985] differ only by the representation of apsidal motion. Epstein (insufficiently) used a linear function of time introduced by Blandford and Teukolsky [1976], while Haugan used a more adequate linear function of true anomaly. Damour & Deruelle (DD) suggested a new, original parametrization of a post-Newtonian orbit of a binary system which allows one to simplify considerably the analysis of pulsar timing observations. The DD model was foreseen by Brumberg [1972, pages 167-168] but it went unnoticed since the Brumberg book was never translated into English.

Some relations between parameterizations have been obtained in [Brumberg, 1991; Soffel, 1989]. Clear and exhaustive statement of relations between the different orbital parameterizations of the post-newtonian two-body problem has been given by Klioner and Kopeikin [1994]. The present book reproduces these relations in full detail and compare them with the post-Keplerian solution of two-body problem in osculating elements, which is the most frequently used in celestial mechanics. This allows us to provide a link between all of the parameterizations and to give a robust consistency check of the models used in binary pulsar timing [Lorimer and Kramer, 2004]. One will analyze the two-body problem in the framework of a general approach which admits a freedom of coordinate transformations and is valid for a wide class of Lorenz-invariant alternative theories of gravity that includes the case of the scalar-tensor theory. The Lorentz-invariant theories of gravity admits the law of conservation of the linear momentum of the two-body system [Will, 1993]. Indeed, equations of motion (6.80)-(6.82) for two spherically-symmetric bodies are

$$
\begin{aligned}
\mathrm{M}_1 \dot{v}_1^i = & -\frac{G\mathfrak{M}_1\mathfrak{M}_2 R_{12}^i}{R_{12}^3} \\
& -\epsilon^2 \left\{ \frac{G\mathrm{M}_1\mathrm{M}_2 R_{12}^i}{R_{12}^3} \left[\gamma v_1^2 - 2(1+\gamma)(\boldsymbol{v}_1\cdot\boldsymbol{v}_2) + (1+\gamma)v_2^2 \right. \right. \\
& \left. -\frac{3}{2}\left(\frac{\boldsymbol{R}_{12}\cdot\boldsymbol{v}_2}{R_{12}}\right)^2 - (1+2\gamma+2\beta)\frac{G\mathrm{M}_1}{R_{12}} - 2(\gamma+\beta)\frac{G\mathrm{M}_2}{R_{12}} \right] \\
& \left. -\frac{G\mathrm{M}_1\mathrm{M}_2(v_1^i - v_2^i)}{R_{12}^3}\left[2(1+\gamma)(\boldsymbol{v}_1\cdot\boldsymbol{R}_{12}) - (1+2\gamma)(\boldsymbol{v}_2\cdot\boldsymbol{R}_{12})\right] \right\},
\end{aligned}
\tag{6.99}
$$

$$
\begin{aligned}
\mathrm{M}_2\dot{v}_2^i = & \frac{G\mathfrak{M}_1\mathfrak{M}_2 R_{12}^i}{R_{12}^3} \\
& +\epsilon^2\Bigg\{\frac{G\mathrm{M}_1\mathrm{M}_2 R_{12}^i}{R_{12}^3}\Big[\gamma v_2^2 - 2(1+\gamma)(\boldsymbol{v}_1\cdot\boldsymbol{v}_2) + (1+\gamma)v_1^2 \\
& -\frac{3}{2}\left(\frac{\boldsymbol{R}_{12}\cdot\boldsymbol{v}_1}{R_{12}}\right)^2 - (1+2\gamma+2\beta)\frac{G\mathrm{M}_2}{R_{12}} - 2(\gamma+\beta)\frac{G\mathrm{M}_1}{R_{12}}\Big] \\
& +\frac{G\mathrm{M}_1\mathrm{M}_2(v_1^i - v_2^i)}{R_{12}^3}\left[2(1+\gamma)(\boldsymbol{v}_2\cdot\boldsymbol{R}_{12}) - (1+2\gamma)(\boldsymbol{v}_1\cdot\boldsymbol{R}_{12})\right]\Bigg\}\ ,
\end{aligned}
\tag{6.100}
$$

where $R_{12}^i = x_1^i - x_2^i$ is the relative distance between the bodies, $v_1^i = dx_1^i/dt$ are $v_1^i = dx_1^i/dt$ their velocities. Summing up these two equations and integrating with respect to time, one yields the integral of the linear momentum $\mathbb{P}^i$ and the integral of the center-of-mass $\mathbb{D}^i$

$$
\begin{aligned}
\mathbb{P}^i = & \left[1+\epsilon^2\left(\frac{1}{2}\boldsymbol{v}_1^2 - \frac{G\mathrm{M}_2}{2R_{12}}\right)\right]\mathrm{M}_1 v_1^i \\
& + \left[1+\epsilon^2\left(\frac{1}{2}\boldsymbol{v}_2^2 - \frac{G\mathrm{M}_1}{2R_{12}}\right)\right]\mathrm{M}_2 v_2^i \\
& - \epsilon^2\frac{G\mathrm{M}_1\mathrm{M}_2 R_{12}^i}{2R_{12}^3}\left(\boldsymbol{R}_{12}\cdot\boldsymbol{v}_1 + \boldsymbol{R}_{12}\cdot\boldsymbol{v}_2\right)\ ,
\end{aligned}
\tag{6.101}
$$

$$
\begin{aligned}
\mathbb{D}^i + \mathbb{P}^i t = & \left[1+\epsilon^2\left(\frac{1}{2}\boldsymbol{v}_1^2 - \frac{G\mathrm{M}_2}{2R_{12}}\right)\right]\mathrm{M}_1 x_1^i \\
& + \left[1+\epsilon^2\left(\frac{1}{2}\boldsymbol{v}_2^2 - \frac{G\mathrm{M}_1}{2R_{12}}\right)\right]\mathrm{M}_2 x_2^i\ .
\end{aligned}
\tag{6.102}
$$

Hence, one can always chose the origin of the global coordinates at the center of mass of the binary system [Barker and O'Connell, 1987] and analyze only the relative motion of the bodies. This is what is assumed in the rest of this section by fixing the integrals $\mathbb{P}^i = \mathbb{D}^i = 0$.

6.4.2
Perturbing post-Newtonian force

Equation of relative motion in the two-body system are derived by subtracting equation (6.100) from equation (6.99) and making use of the replacements

$$
x_1^i = \frac{\mathrm{M}_2}{\mathrm{M}_1+\mathrm{M}_2}r^i\ , \qquad x_2^i = -\frac{\mathrm{M}_1}{\mathrm{M}_1+\mathrm{M}_2}r^i\ ,
\tag{6.103}
$$

which are valid in the barycentric coordinate system of the two-body system, and the relative distance $r^i \equiv R_{12}^i$ - the notation will be used in the subsequent equations. One will consider a more general case of the post-Newtonian equations of relative motion following Brumberg [1972, 1991]. According to Brumberg's work the perturbing force of the relative motion is given by equations

$$\ddot{r}^i = -\frac{\mathcal{G}\mathrm{M}r^i}{r^3} + F^i \,, \tag{6.104}$$

$$F^i = \frac{\mathcal{G}\mathrm{M}}{c^2}\left[\left(2\sigma\frac{\mathcal{G}\mathrm{M}}{r} - 2\epsilon v^2 + 3\alpha\frac{(\boldsymbol{r}\cdot\boldsymbol{v})^2}{r^2}\right)\frac{r^i}{r^3} + 2\mu\frac{(\boldsymbol{r}\cdot\boldsymbol{v})}{r^3}\boldsymbol{v}\right] , \tag{6.105}$$

that was originally introduced as a tool to investigate the motion of a test body in the Schwarzschild gravitational field. However, as has been already recognized by Soffel [Soffel, 1989], this force is general enough to cover the case of the weak-field two-body problem in the framework of the parameterized post-Newtonian (PPN) formalism [Will, 1993, 2006]. This perturbing force covers also the case of two-body problem in the strong-field regime of generic gravitational theories [Gergely *et al.*, 2006]. In equation (6.105) α, ϵ, μ and σ are arbitrary numerical parameters, $\mathrm{M} = \mathrm{M}_1 + \mathrm{M}_2$ is the total mass of the system, and $\mathcal{G}$ is the parameter which may differ from the universal gravitational constant G by a constant factor $\mathcal{G} = \mathcal{A}G$. By comparison of equations (6.104), (6.105) with equation (11.68) of [Will, 1993] one can see that

$$\alpha = \frac{1}{2}\nu \,, \tag{6.106}$$

$$\epsilon = \frac{3}{2}\nu + \frac{1}{4}\mathcal{E} - \frac{1}{4} \,, \tag{6.107}$$

$$\mu = -\nu + \frac{1}{2}\mathcal{E} + \frac{1}{2} \,, \tag{6.108}$$

$$\sigma = \nu + \frac{1}{2}\mathcal{E} + \frac{1}{2}\xi \,, \tag{6.109}$$

where

$$\mathcal{E} = \frac{3\mathcal{B}}{\mathcal{A}} \,, \tag{6.110}$$

$$\nu = \frac{\mathrm{M}_1\mathrm{M}_2}{\mathrm{M}^2} \,, \tag{6.111}$$

$$\xi = \frac{\mathrm{M}_1\mathcal{D}_{211} + \mathrm{M}_2\mathcal{D}_{122}}{\mathcal{A}^2\mathrm{M}} \,, \tag{6.112}$$

are the parameters defining the specific theory of gravitation and depending on the masses and internal structure of the bodies [Will, 1993]. In particular, for general relativity, $\mathcal{E} = 3$ and $\xi = 1$, and, therefore,

$$\alpha = \frac{1}{2}\nu \,, \tag{6.113}$$

$$\epsilon = \frac{3}{2}\nu + \frac{1}{2} \,, \tag{6.114}$$

$$\mu = -\nu + 2 \,, \tag{6.115}$$

$$\sigma = \nu + 2 \,, \tag{6.116}$$

which makes equations (6.104)-(6.105) the usual EIH equations of motion of the two-body problem (see, e.g., [Brumberg, 1972, 1991; Soffel, 1989; Will, 1993]). In the weak-field limit of the scalar-tensor theory of gravity one gets [Will, 1993]

$$\mathcal{E} = 2\gamma + 1\,, \tag{6.117}$$
$$\xi = 2\beta - 1\,. \tag{6.118}$$

Let us note that the equations of motion in the form of equations (6.104)-(6.105) are form-invariant with respect to a certain class of coordinate transformations [Barker *et al.*, 1986; Brumberg, 1972]

$$r^i = r'^i\left(1 - \alpha' \frac{\mathcal{G}\mathrm{M}}{c^2 r'}\right)\,, \tag{6.119}$$

where α' is a numerical parameter. In this case the equations of motion in the two-body problem in new coordinates have the same form (6.104), (6.105) with the parameters [Barker *et al.*, 1986, equation (A.1)]

$$\alpha = \frac{1}{2}\nu + \alpha'\,, \tag{6.120}$$
$$\epsilon = \frac{3}{2}\nu + \frac{1}{4}\mathcal{E} - \frac{1}{4} + \frac{1}{2}\alpha'\,, \tag{6.121}$$
$$\mu = -\nu + \frac{1}{2}\mathcal{E} + \frac{1}{2} - \alpha'\,, \tag{6.122}$$
$$\sigma = \nu + \frac{1}{2}\mathcal{E} + \frac{1}{2}\xi - \alpha'\,, \tag{6.123}$$

Transformation (6.119) evidently does not cover all the freedom of possible coordinate transformations of the relative equations of motion in the post-Newtonian two-body problem. The effect of time transformation has been considered in papers by Barker and O'Connell [1984] and Damour and Schäfer [1985] discussed the effect of spacetime coordinate transformation on the Lagrangian and on the metric functional of a system of N point masses. In what follows, when comparing different orbital representations, one considers that the coordinate system for all representations is the same.

6.4.3
Orbital solution in the two-body problem

6.4.3.1 *Osculating elements parametrization*

In astrodynamics the osculating orbit of a celestial body in space is the fictitious Keplerian orbit corresponding to its actual position and velocity for that given moment of time, which the body would have if gravitational perturbations were not present (see Section 1.4). An osculating orbit and the body's position upon it are fully described by the six standard Keplerian orbital elements, which are easy to calculate as long as one knows the body's position and velocity relative to the companion. Perturbations cause the osculating elements to evolve, sometimes very quickly. In such cases, a more complex set of proper orbital elements may better describe the most important aspects of the orbit.

Brumberg [1972, 1991] described the parametrization of the relative orbit of the two-body problem in the post-Newtonian approximation based on the solution of the equations for osculating elements with the perturbing force expressed by equation (6.105). The solution for elliptic motion of the two bodies has the following form:

$$r(t) = a(t)\left[1 - e(t)\cos E\right] , \tag{6.124}$$

$$n\,(t - T_0) = E - e(t)\sin E - \Delta l(t) , \tag{6.125}$$

$$\theta = f + \omega(t) , \tag{6.126}$$

$$\tan\frac{f}{2} = \left[\frac{1 + e(t)}{1 - e(t)}\right]^{1/2}\tan\frac{E}{2} , \tag{6.127}$$

where the angles f, E, and θ are the true anomaly, eccentric anomaly, and the argument of latitude; n is the constant mean motion; r is the distance between bodies; t is the barycentric coordinate time of the binary system; and T_0 is the moment of pericenter passage. The osculating semi-major axis $a(t)$, eccentricity $e(t)$, angular distance of pericenter from ascending node $\omega(t)$, and the periodic part $\Delta l(t)$ of the perturbed mean anomaly are expressed as follows:

$$\begin{aligned} a(t) &= a_0 - da_0 + \frac{me}{(1-e^2)^2} \\ &\quad\times\left\{\left[4\,(\epsilon - \mu - \sigma) - \left(\frac{9}{2}\alpha - 4\epsilon + 4\mu\right)e^2\right]\cos f + (2\epsilon - 2\mu - \sigma)\,e\cos 2f + \frac{1}{2}\alpha e^2\cos 3f\right\} , \end{aligned} \tag{6.128}$$

$$\begin{aligned} da_0 &= \frac{me}{(1-e^2)^2} \\ &\quad\times\left\{\left[4\,(\epsilon - \mu - \sigma) - \left(\frac{9}{2}\alpha - 4\epsilon + 4\mu\right)e^2\right]\cos f_0 + (2\epsilon - 2\mu - \sigma)\,e\cos 2f_0 + \frac{1}{2}\alpha e^2\cos 3f_0\right\} \\ &= m\left\{\left[2(\sigma - \alpha) + (6\alpha - 6\epsilon + 6\mu + \sigma)e^2\right]\frac{1}{(1-e^2)}\right. \\ &\quad\left. + 2(3\alpha - 2\epsilon + 2\mu)\left(\frac{a}{r_0}\right) - 2(3\alpha - 2\epsilon + 2\mu + \sigma)\left(\frac{a}{r_0}\right)^2 + 2\alpha\left(1 - e^2\right)\left(\frac{a}{r_0}\right)^3\right\} , \end{aligned} \tag{6.129}$$

$$\begin{aligned} e(t) &= e_0 - de_0 + \frac{m}{a\,(1-e^2)} \\ &\quad\times\left\{\left[2\,(\epsilon - \sigma) - \left(\frac{9}{4}\alpha - 2\epsilon + 4\mu\right)e^2\right]\cos f + \left(\epsilon - \mu - \frac{1}{2}\sigma\right)e\cos 2f + \frac{1}{4}\alpha e^2\cos 3f\right\} , \end{aligned} \tag{6.130}$$

$$\begin{aligned} de_0 &= \frac{m}{a\,(1-e^2)} \\ &\quad\times\left\{\left[2\,(\epsilon - \sigma) - \left(\frac{9}{4}\alpha - 2\epsilon + 4\mu\right)e^2\right]\cos f_0 + \left(\epsilon - \mu - \frac{1}{2}\sigma\right)e\cos 2f_0 + \frac{1}{4}\alpha e^2\cos 3f_0\right\} \\ &= \frac{m}{ea}\left\{\left[-\alpha - 2\mu + \sigma + (3\alpha - 3\epsilon + 5\mu + \frac{1}{2})e^2\right]\frac{1}{(1-e^2)}\right. \\ &\quad\left. + (3\alpha - 2\epsilon + 4\mu)\left(1 - e^2\right)\left(\frac{a}{r_0}\right) - (3\alpha - 2\epsilon + 2\mu + \sigma)\left(1 - e^2\right)\left(\frac{a}{r_0}\right)^2 + \alpha\left(1 - e^2\right)^2\left(\frac{a}{r_0}\right)^3\right\} , \end{aligned} \tag{6.131}$$

$$\begin{aligned}\omega(t) &= \omega_0 - d\omega_0 + \frac{m}{a(1-e^2)}\{(2\epsilon+2\mu-\sigma)f \\ &+\left[2\frac{\epsilon-\sigma}{e} - \left(\frac{3}{4}\alpha - 2\epsilon\right)e\right]\sin f + \left(\epsilon-\mu-\frac{1}{2}\sigma\right)\sin 2f + \frac{1}{4}\alpha e\sin 3f\right\}\ ,\end{aligned} \tag{6.132}$$

$$\begin{aligned}d\omega_0 &= \frac{m}{a(1-e^2)}\{(2\epsilon+2\mu-\sigma)f_0 \\ &+\left[2\frac{\epsilon-\sigma}{e} - \left(\frac{3}{4}\alpha - 2\epsilon\right)e\right]\sin f_0 + \left(\epsilon-\mu-\frac{1}{2}\sigma\right)\sin 2f_0 + \frac{1}{4}\alpha e\sin 3f_0\right\}\ ,\end{aligned} \tag{6.133}$$

$$\begin{aligned}\Delta l(t) &= \frac{m}{a\sqrt{1-e^2}}\left\{(-3\alpha+2\epsilon-2\mu)\,e\sqrt{1-e^2}\sin E\right. \\ &+\left[2\frac{\sigma-\epsilon}{e} + \left(\frac{15}{4}\alpha + 4\mu - 2\epsilon\right)e\right]\sin f - \left(\epsilon-\mu-\frac{1}{2}\sigma\right)\sin 2f - \frac{1}{4}\alpha e\sin 3f\right\}\ ,\end{aligned} \tag{6.134}$$

Here $m = \mathcal{G}M/c^2$; a_0, e_0, and ω_0 are the osculating numerical values of the elements at the same arbitrary moment of time t_0 (which generally speaking may differ from T_0); the constant relativistic corrections da_0, de_0, and $d\omega_0$ have been chosen so that $a(t_0) = a_0$, $e(t_0) = e_0$, and $\omega(t_0) = \omega_0$, and, therefore, are equal to the periodic parts of $a(t)$, $e(t)$, and $\omega(t)$ evaluated at $t = t_0$; f_0 is the value of true anomaly at $t = t_0$. In equations (6.128)-(6.134) and hereafter in relativistic terms, the letters a and e denote the constant values of the semi-major axis and the eccentricity of the relative orbit, so that

$$\left(\frac{a}{r_0}\right) = \frac{1+e_0\cos f_0}{\left(1-e_0^2\right)}\ . \tag{6.135}$$

The constant value of the orbital frequency n entering the left side of equation (6.125) is

$$\begin{aligned}n &= n_0\left\{1 + \frac{m}{a}\left[-3\alpha+2\epsilon-2\mu+3(3\alpha-2\epsilon+2\mu)\left(\frac{a}{r_0}\right)\right.\right. \\ &\left.\left.-3(3\alpha-2\epsilon+2\mu+\sigma)\left(\frac{a}{r_0}\right)^2 + 3\alpha\left(1-e^2\right)\left(\frac{a}{r_0}\right)^3\right]\right\}\ ,\end{aligned} \tag{6.136}$$

where n_0 is defined in terms of the following equation

$$a_0^3 n_0^2 = \mathcal{G}M\ , \tag{6.137}$$

which is the (unperturbed) third Kepler's law.

6.4.3.2 *The Damour-Deruelle parametrization*

Damour and Deruelle [1985, 1986] have invented a new (DD) parametrization of the post-Keplerian relativistic orbit of the two-body problem. Being originally proved for general relativity, the Damour-Deruelle approach has been generalized by Damour and Taylor [1992] to the case of generic gravitational theories. According to the DD parametrization, the position of a body on the relative orbit can be found as

$$r(t) = a_R\,(1 - e_R \cos U)\ , \tag{6.138}$$

$$n\,(t - T_0) = U - e_t \sin U\ , \tag{6.139}$$

$$\theta = \theta_0 + (1 + k)A_{e_\theta}\ , \tag{6.140}$$

$$\tan\frac{A_{e_\theta}}{2} = \left(\frac{1 + e_\theta}{1 - e_\theta}\right)^{1/2} \tan\frac{U}{2}\ , \tag{6.141}$$

where the angles A_{e_θ}, U are analogies of the true anomaly f and the eccentric anomaly E, but differ from them by relativistic terms. The functions $r(t)$ and $\theta(t)$, being polar coordinates of one body with respect to other, do not depend on the parametrization used and, therefore, are exactly the same as in the osculating elements parametrization. The coordinate time t also does not depend on the parametrization. Taking these circumstances into account, it is easy to check by comparing the functions $r(t)$ and $\theta(t)$ defined by equations (6.124)-(6.127) and (6.138)-(6.141) that the constants n and T_0 are exactly the same as in the osculating elements parametrization, $\theta_0 = \omega_0 - d\omega_0$, and

$$k = \frac{m}{a\,(1 - e^2)}\,(2\epsilon + 2\mu - \sigma)\ . \tag{6.142}$$

Our aim is to find the relation between the integration constants of the osculating element solution and the DD parametrization of the orbit. At the first step, one equates equation (6.125) and (6.139) to get relations between U and E as a function of difference $e_t - e_0$ between eccentricities. Then, by equating equation (6.124) and (6.138) one gets the differences $a_R - a_0$, $e_R - e_0$, and $e_t - e_0$. At the second step, equating equations (6.126) and (6.140) gives the relation between f and A_{e_θ}. Finally, comparing equations (6.127) and (6.141) one gets another difference $e_\theta - e_0$ between eccentricities. One can see that although the system of equations to be solved in order to determine the relations between constants, is an over-determined system of linear equations, it has a single solution, and the constants maps to each other uniquely. The relations read

$$a_R = a_0 - da_0 + \frac{m}{a(1 - e^2)}\left[-2\epsilon + 2\sigma + (2\alpha - 2\epsilon + 6\mu + \sigma)e^2 + 2(\alpha - \epsilon)e^4\right]\ , \tag{6.143}$$

$$e_R = e_0 - de_0 + \frac{em}{2a(1 - e^2)}\left[-2\epsilon + 6\mu + 3\sigma + 4(\alpha - \epsilon)e^2\right]\ , \tag{6.144}$$

$$e_t = e_0 - de_0 + \frac{em}{2a(1 - e^2)}\left[-2\alpha - 2\epsilon + 2\mu + 3\sigma + 2(3\alpha - 2\epsilon + 2\mu)e^2\right]\ , \tag{6.145}$$

$$e_\theta = e_0 - de_0 + \frac{em}{2a(1 - e^2)}\left[2\alpha - 2\epsilon + 6\mu + 3\sigma + 2(\alpha - 2\epsilon)e^2\right]\ . \tag{6.146}$$

Besides that, one can check that

$$a_R = \left(\frac{\mathcal{G}\mathrm{M}}{n^2}\right)^{1/3}\left[1 - \frac{2}{3}\,(\epsilon + 2\mu)\,\frac{m}{a}\right]\ , \tag{6.147}$$

and relations between corresponding anomalies are

$$
\begin{aligned}
A_{e_\theta} &= (1-k)f + \omega(t) - \omega_0 + d\omega_0 \qquad (6.148)\\
&= \frac{m}{a(1-e^2)}\left\{\left[2\frac{\epsilon-\sigma}{e} - \left(\frac{3}{4}\alpha - 2\epsilon\right)e\right]\sin f + \left(\epsilon - \mu - \frac{1}{2}\sigma\right)\sin 2f + \frac{1}{4}\alpha e \sin 3f\right\}\,,\\
U &= E - \frac{\Delta l + [e(t)-e_t]\sin E}{1-e\cos E} \qquad (6.149)\\
&= E - \frac{m}{a(1-e^2)^{3/2}}\left\{\left[2\frac{\sigma-\epsilon}{e} + \left(\frac{7}{4}\alpha + 2\mu - \sigma\right)e\right]\sin f\right.\\
&\quad \left. + \left(\mu - \epsilon + \frac{1}{2}\sigma + \frac{1}{2}\alpha e^2\right)\sin 2f - \frac{1}{4}\alpha e \sin 3f\right\}\,.
\end{aligned}
$$

The relation (6.143) between a_R and a_0 has been already derived with a different technique in [Brumberg, 1991, see page 95], who used notation a' instead of a_R (see section 6.4.3.4 below). According to Damour and Deruelle [1985, 1986] it is useful to introduce a new eccentricity $e_T = e_t(1+\delta) + e_\theta - e_R$, where a new parameter

$$
\delta = \frac{\mathcal{G}\mathrm{M}_2(\mathrm{M}_1 + 2\mathrm{M}_2)}{c^2 a_R \mathrm{M}}\,, \qquad (6.150)
$$

describes a periodic difference between the proper time of the first body and the barycentric coordinate time. Then,

$$
\begin{aligned}
\delta_R &= \frac{e_R - e_T}{e_T} = \frac{m}{a}2\mu - \delta\,, \qquad (6.151)\\
\delta_\theta &= \frac{e_\theta - e_T}{e_T} = \frac{m}{a}(\alpha + 2\mu) - \delta\,, \qquad (6.152)
\end{aligned}
$$

are parameters, which describe purely periodic relativistic corrections to the Keplerian motion. Their measurement would provide a stringent test on relativistic theories of gravity. Unfortunately, it is almost impossible to observe these parameters with the current pulsar timing technique because they are highly correlated with standard astrometric observables of unperturbed Keplerian orbit [Damour and Deruelle, 1986].

One has analyzed the relative orbit. Transformation, for example, to the orbit of the first body with respect to the barycenter of the binary system can be done by accounting for the difference between the relative orbit and the orbit of the body relative to the barycenter of the system, which can be written as

$$
\begin{aligned}
x_1 &= a_r(1 - e_r\cos U)\,, \qquad (6.153)\\
a_r &= \frac{\mathrm{M}_2}{\mathrm{M}}a_R\,, \qquad (6.154)\\
e_r &= e_R\left[1 + \frac{m}{a}\frac{\mathrm{M}_1(\mathrm{M}_2 - \mathrm{M}_1)}{2\mathrm{M}^2}\right]\,, \qquad (6.155)
\end{aligned}
$$

and substituting the parameters from equations (6.106)-(6.112) one gets exactly the same expressions for $\delta_r = (e_r - e_T)/e_T$, δ_θ, k and the relation between a_R and n as in paper by Damour and Taylor [1992]. Note that for the orbit of the first body with respect to the barycenter, $e_T = e_r(1+\delta) + e_0 - e_r$.

Portilla and Villareal [2004] have generalized the DD parametrization for hyperbolic and parabolic orbits of the two-body problem in the first post-Newtonian approximation. The DD parametrization of the elliptic two-body orbit has been extended to the second post-Newtonian (2PN) approximation in papers [Damour and Schäfer, 1988; Schäfer and Wex, 1993a,b], and even to the third post-Newtonian (3PN) approximation in [Memmesheimer *et al.*, 2004]. These PN extensions are required for predicting templates of gravitational wave forms from the events of coalescence of compact binary systems comprised of neutron stars and/or black holes. This is one of the most important theoretical problems of modern relativistic astrophysics but it goes far beyond the scope of the relativistic celestial mechanics in the solar system. For this reason, one skips the interesting but otherwise cumbersome, details of the 2PN and 3PN parameterizations of the relativistic two-body problem.

6.4.3.3 *The Epstein-Haugan parametrization*

Epstein [1977] and Haugan [1985] proposed another (EH) parametrization of the post-Keplerian orbit of two bodies in the framework of general relativity. The parametrization is based on the results obtained earlier by Wagoner and Will [1976], and can be written down in the following generalized form:

$$r(t) = a'\left(1 - e'\cos Z - h'\cos 2Z\right) , \tag{6.156}$$

$$n\left(t - T_0\right) = Z - g\sin Z - h\sin 2Z , \tag{6.157}$$

$$\theta = \theta_0 + (1+k)\eta , \tag{6.158}$$

$$\tan\frac{\eta}{2} = \left(\frac{1+e''}{1-e''}\right)^{1/2}\tan\frac{Z}{2} , \tag{6.159}$$

where a', e', e'', g, $h' = O\left(\epsilon^2\right)$, $h = O\left(\epsilon^2\right)$, θ_0, n, T_0, and k are the constants of integration. Three eccentricities e', e'', and g differ from each other by the post-Newtonian terms of the order of $O(\epsilon^2)$. Let us note that Haugan [1985] denotes by the letter e the quantity that one has designated by e". Again, it is easy to check that constants θ_0, n, T_0, and k from the EH parametrization coincide with those from the previous section 6.4.3.2. Comparison of equations (6.156)-(6.159) with equations (6.138)-(6.141) gives the set of relations between the orbital constants

$$g = e_t - \frac{2}{e}h , \tag{6.160}$$

$$e' = e_R - eh , \tag{6.161}$$

$$e'' = e_\theta + \frac{2(1-e^2)}{e}h , \tag{6.162}$$

$$a' = a_R + ah , \tag{6.163}$$

$$h' = h , \tag{6.164}$$

$$\eta = A_{e_\theta} , \tag{6.165}$$

$$Z = U - \frac{2h}{e}\sin U . \tag{6.166}$$

One can see, and it was first realized by Klioner and Kopeikin [1994], that the EH parametrization in the form of equations (6.156)-(6.159) contains in fact one free post-Newtonian parameter $h = O\left(\epsilon^2\right)$, which can be chosen arbitrary. This means that the semi-major axis and the eccentricity obtained from fitting of binary pulsar timing data in the framework of the EH timing model depend on the specific value of h in accordance with equations (6.162), (6.163). For example, for $h = 0$ the Epstein-Haugan parametrization coincides with that of Damour and Deruelle. If one chooses h in the same way as Epstein and Haugan did,

$$h = \frac{me^2}{4a(1-e^2)^2}\left[\left(-13+2e^2\right)+\left(3+5e^2\right)\frac{\mathrm{M}_1\mathrm{M}_2}{(\mathrm{M}_1+\mathrm{M}_2)^2}\right], \tag{6.167}$$

one gets the same expressions for the general-relativistic relations between g, e', e'' as described in Haugan [1985].

Soffel [1989] has proved already that the EH parametrization remains valid in more general case of the PPN formalism [Will, 1993]. Our results demonstrates that it is also valid in generic gravitational theories in the strong-field limit. The values of the EH parameters for generic gravitation theories can be calculated from equations (6.160)-(6.164), the values of constants of the DD parameterizations, and the expression

$$h = \frac{me^2}{a\left(1-e^2\right)^2}\left[\frac{1}{2}\epsilon - \frac{3}{4}\sigma - \frac{1}{2}\alpha - \mu - \left(\frac{1}{2}\alpha - \epsilon\right)e^2\right], \tag{6.168}$$

which has been derived in [Klioner and Kopeikin, 1994] for the perturbing force (6.105) according to the original Wagoner & Will [Wagoner and Will, 1976] method.

Since the DD representation can be considered as a special case of the EH representation with $h = 0$, and the original EH model corresponds to h defined by equation (6.167), one can compare the derived theoretical relations between relevant constants with numerical values obtained in paper by Taylor and Weisberg [1989] for PSR 1913+16. In the original representation of Damour and Deruelle the letters "a" and "e" designate $a_R \equiv a_{DD}$ and $e_T \equiv e_{DD}$, respectively, while in the original representation of Epstein and Haugan, ”a” and ”e” designate $a' \equiv a_{EH}$ and $e'' \equiv e_{EH}$. Therefore, in the case of general relativity for PSR 1913+16 our theoretical relations between constants give $e_{EH} - e_{DD} = -8.2 \times 10^{-6}$ and $(a_{EH} - a_{DD})/a_{DD} = (x_{EH} - x_{DD})/x_{DD} = -6.1 \times 10^{-6}$, where $x = a \sin i/c$ is the observable parameter. These values are to be compared with the observational data presented in [Taylor and Weisberg, 1989, Table 4]. The latter gives $e_{EH} - e_{DD} = (-4.4 \pm 5.1) \times 10^{-6}$ and $(x_{EH} - x_{DD})/x_{DD} = (-2.1 \pm 6.4) \times 10^{-5}$. One can see a good agreement of the results that provides a practical consistency check of different pulsar timing models.

6.4.3.4 ***The Brumberg parametrization***

Brumberg [1972, 1991] has proposed independent parametrization of the relativistic two-body problem, which can be written as (see [Brumberg, 1991, equations

(3.1.60)-(3.1.69)])

$$\begin{aligned} r(t) &= a^*\,(1-e^*\cos U)\ , && (6.169)\\ n\,(t-T_0) &= U - e^*\sin U + F_1(U)\ , && (6.170)\\ \theta &= \theta_0 + (1+k)A + F_2(U)\ , && (6.171)\\ \tan\frac{A}{2} &= \left(\frac{1+e^*}{1-e^*}\right)^{1/2}\tan\frac{U}{2}\ , && (6.172) \end{aligned}$$

where $a^* \equiv a_R$, $e^* \equiv e_R$, $A \equiv A_{e_R} = A_{e_\theta} - F_2(U)$, and

$$\begin{aligned} F_1(U) &= (e_R - e_t)\sin U = e\,(\delta+\delta_\theta)\sin U\ , && (6.173)\\ F_2(U) &= \frac{e_\theta - e_R}{1-e^2}\sin A(U) = \frac{e}{1-e^2}\,(\delta_\theta - \delta_R)\sin A(U)\ . && (6.174) \end{aligned}$$

This exact parametrization of the post-Keplerian relative orbit is useful to check how far from the reality is the modified parametrization of Blandford and Teukolsky (BT+) having being used for binary pulsar data analysis by Damour and Taylor [1992]. The latter is defined, in fact, by Brumberg's equations (6.169)-(6.172) with $F_1(U) = F_2(U) = 0$. Neglecting F_1 and F_2 in the orbital parametrization produces a rather large effect in theoretical arrival times of pulsar's pulses. For PSR 1913+16 the effect may amount to 9 μs. However, by adjusting the Keplerian elements of the orbit, one can reduce the post-fit residual difference between the DD and BT+ models to 2 μs as it happened in [Damour and Taylor, 1992]. It explains why the BT+ parametrization can be fitted (at least, in some cases) so well to the actual observational data.

One sees that the Brumberg parametrization is very close to the Damour-Deruelle parametrization. The major semi-axes a^* and the eccentricity e^* of the Brumberg parametrization are exactly equal to a_R and e_R of the DD parametrization. This fact has been recognized by Soffel [1989]. The only transformation needed to get the DD parametrization from the Brumberg's one is to introduce A_{e_θ} instead of A_{e_R}, which allows us to get rid of $F_2(U)$ as in the DD model.

7
Relativistic Astrometry

7.1 Introduction

Until now it are electromagnetic signals coming from various astronomical objects which deliver the most exhaustive and accurate physical information about numerous intriguing phenomena going on in the surrounding universe. Astrometry is the branch of astronomy that relates to precise measurements and explanations of the positions and movements of stars and other celestial bodies. The information obtained by astrometric measurements is very important in contemporary research on the kinematics and physical origin of the solar system, the Milky Way, and the whole universe. The principal function of astrometry is to provide astronomers with an inertial reference frame in the sky that is used to record their observations. Astrometry is also fundamental for fields of celestial mechanics, stellar dynamics and galactic astronomy [Kovalevsky and Seidelmann, 2004].

Astronomers use astrometric techniques for the tracking of near-Earth objects and to identify distant celestial objects by observing their proper motions. Many record-breaking solar system objects were detected by observing their movements relative to the background stars, which remain fixed during the observational session. Examples are - Quaoar and Sedna - two Trans-Neptunian object and potential dwarf planets orbiting the Sun in the Kuiper belt. They were discovered by C. A. Trujillo and M. E. Brown, and other astronomers at Caltech using the Palomar Observatory's Samuel Oschin telescope and the Palomar-Quest large-area CCD camera [Trujillo and Brown, 2003; Williams *et al.*, 2002]. The ability of astronomers to track the positions and movements of such celestial bodies is crucial to the understanding of the origin of the solar system and its past, present, and future.

Astrometry is also instrumental for keeping time as the universal coordinated time (UTC) is basically the atomic time synchronized to Earth's rotation by means of exact observations. Astrometric time-keepers like pulsars may supersede the atomic clocks on long intervals of time. Astrometry is a powerful tool to detect extra-solar planets orbiting other stars by measuring the periodic displacement the planets cause in their parent star's apparent position on the sky, due to their mutual orbit around the center of mass of the planetary system. NASA's planned Space Interferometry Mission (SIM) was designed to utilize astrometric techniques to detect terrestrial planets orbiting around several hundred of the nearest solar-type stars.

Astrometric measurements are used by astrophysicists to constrain certain models of physical processes. By measuring the velocities of pulsars, it is possible to put a limit on the asymmetry of supernova explosions. Also, astrometric results are used to determine the distribution of dark matter in the galaxy. Astrometry is involved in creating the cosmic distance ladder because it is used to establish parallax distance estimates for stars in the Milky Way [Weinberg, 1972]. Furthermore, astrometry always served as an experimental technique for testing fundamental theories of gravity devised by Newton and notably by Einstein [Will, 1993].

Relativistic astrometry can be viewed as a part of relativistic celestial mechanics dealing with motion of relativistic particles of light - photons - in a gravitational field of massive bodies. In astrometry photons are considered as test particles having no influence on the metric tensor generated by the massive bodies. For this reason, their motion is described by the equations of *light* (*null*) geodesics. Photons are particles of vector electromagnetic field having two independent degrees of freedom (polarizations) [Landau and Lifshitz, 1975; Weinberg, 1972]. Polarization plays major role in optical and radio observations of celestial objects. However, relativistic effects associated with the polarization of light are negligibly small [Barbieri and Guadagnini, 2005; Mashhoon, 1974] so that they can be ignored in subsequent discussion.

In vacuum and in the absence of gravitational field photons move along straight lines with constant speed c, which is equal to the fundamental speed of the Minkowski spacetime. For this reason, the convention is to call the fundamental speed c as the “speed of light" irrespectively of the nature of equations where it appears. This linguistic tradition brings sometimes misunderstanding and confusion among researchers [Ellis and Uzan, 2005]. For example, the theoretical value of the speed of gravity in general relativity is the same as the speed of light in Maxwell's theory but the two speeds must be carefully distinguished in experimental gravitational physics [Kopeikin, 2004]. Astrometric study of how gravitational perturbation of a moving massive body affects motion of photons from their rectilinear and uniform motion allows us to understand deeper the dynamic properties of gravitational field. Because photons move with the same speed as gravity, they are much more sensitive to gravitational perturbations caused by space-space, g_{ij}, and spacetime, g_{0i}, components of the metric tensor, than massive test particles which are affected mainly by the time-time component, g_{00}, of the metric. This can be easily seen from the equations of geodesic for massive particles and photons.

Relativistic effects in propagation of light through static gravitational field played a key role in experimental confirmation of general relativity until recently [Will, 1993, 2006]. However, rapidly growing accuracy of astrometric measurements demands more exact solution of the problem of propagation of electromagnetic waves including prediction of relativistic deflection of light and time delay in non-stationary gravitational fields. Present day technology has achieved a level of precision of optical and radio interferometric observations approaching 1 μas [Fomalont and Reid, 2004; Johnston *et al.*, 2000]. Nano-arcsecond astrometric measurements may be available in next few decades [Johnston *et al.*, 2000; Perryman, 2005]. Adequate interpretation of the astrometric data having this precision requires corresponding theoretical treatment of periodic relativistic effects in the propagation of electromagnetic signals in variable gravitational fields of oscillating and precessing stars, steady-state and coalescing binary systems, and colliding galaxies [Kopeikin and Korobkov, 2005; Kopeikin *et al.*, 2006, 1999; Schluessel, 2008]. Furthermore, space astrometric missions like Gaia [Lindegren and Perryman, 1996] and SIM [Peterson and Shao, 1997] will also have precision of about 1-10 μas in measuring positions and parallaxes of stars, and about 1-10 μas per year in measuring their proper motion. At this level of accuracy the gravitational field of the solar system can not be treated as static and spherically symmetric. Rotation and oblateness of the Sun and large planets as well as time variability of the gravitational field caused their orbital motion should be undoubtedly taken into account [Klioner, 2003a; Klioner and Kopeikin, 1992; Kopeikin and Mashhoon, 2002; Kopeikin and Makarov, 2007; Kopeikin *et al.*, 2007;

Kopeikin and Schäfer, 1999; Kopeikin, 2009].

Most of the practical approaches having been developed for integrating equations of propagation of electromagnetic signals in gravitational field of the solar system and other massive celestial bodies were based on the post-Newtonian presentation of the metric tensor. It is well-known [Fock, 1964; Misner *et al.*, 1973] that the post-Newtonian approximation for the metric tensor is valid only within the so-called near zone of an isolated astronomical N-body system that is the region of space around the system with the characteristic size of a wavelength of gravitational radiation emitted by the system. For example, Jupiter orbiting the Sun emits gravitational waves with a wavelength of about 1.7 parsecs, and the binary pulsar PSR B1913+16 radiates gravitational waves with a wavelength of about 28 astronomical units. It is clear that the post-Newtonian metric can be used for calculation of light propagation only from the sources lying inside the near zone of the gravitating system of the massive bodies. However, most of the stars, quasars, and other sources of electromagnetic radiation are lying far beyond the boundary of the near zone of the solar system and other massive bodies and another method of solving the problem of propagation of light from these sources to observer on the Earth should be applied.

In the absence of such advanced mathematical technique researches relied upon the post-Newtonian approximation of the metric tensor assuming implicitly that the perturbations from the time-dependent part of the metric tensor are sufficiently small and may be neglected in the existed data processing algorithms [Brumberg, 1991; Moyer, 2003; Soffel, 1989; Will, 1993]. An attempt to scrutinize this assumption has been undertaken by Brumberg *et al.* [1990]; Klioner and Kopeikin [1992]; Kopeikin [1990] who employed the technique of matching of two asymptotic solutions of the light ray geodesics - in the near zone of the solar system and far away from it, where its gravitational field can be approximated by the static Schwarzschild solution. Nevertheless, a rigorous solution of the equations of light propagation being simultaneously valid both far outside and inside the solar system was not found at that time [1].

One additional problem to be enlightened relates to the approximation of motion of the light-ray deflecting bodies during the time of propagation of light from the point of emission to the point of observation. The post-Newtonian metric of a gravitating N-body system is not static and the bodies move around a common center of mass while light is propagating. It was presupposed that the body a exerts the biggest force on the propagating photon at the time t_a of the closest approach of the light ray to the body. For this reason, most of the papers on the light propagation fix coordinates of each gravitating body $a = 1, 2, 3, ..., N$ in the post-Newtonian metric at the corresponding times t_a (see, for instance, [Brumberg, 1991; Hellings, 1986a,b; Will,

1) Will [2003] used similar technique in order to identify relativistic terms associated with the speed of propagation of gravitational waves in the expression for gravitational time delay of light. Will's analysis has subtle pitfalls caused by the limited power of his parametrization of equations of general relativity with the speed of gravity parameter c_g. General relativistic parametrization of the speed of gravity is achieved after replacement of the fundamental speed $c \to c_g$ in *all* terms containing time derivatives, while Will [2003] parameterized a limited number of terms with the second-order time derivatives. Detailed discussion of this issue is given in [Kopeikin, 2003b, 2004; Kopeikin and Fomalont, 2006].

1993] and interesting discussion of this issue in [Sovers *et al.*, 1998, page 1406]). Nonetheless, until recently it was not fully clear why should one uses precisely the moment t_a for fixing positions of the massive bodies in the solution of the light-ray equations and what magnitude of residual terms will be in the calculation of relativistic time delay and light deflection angle if one chooses a slightly different moment of time. Various researches came up with a conceivable explanation for choosing the moment t_a as the time of the closest approach [Hellings, 1986a,b; Klioner, 2003a; Klioner and Kopeikin, 1992] arguing that such a choice minimizes the residual terms in the solution of the equation of propagation of light rays obtained by the asymptotic matching technique. One notes, however, that the proof of this statement was never completed with sufficient mathematical rigor and can not be actually achieved for the reasons clarified below.

Recently, astronomers have started reconsidering the problem of propagation of light rays in variable gravitational fields of gravitating system of bodies to find out new, promising applications of relativistic astrometry. First of all, a profound, systematic approach to integration of light geodesic equations in arbitrary, time-dependent gravitational field of a localized self-gravitating system possessing a multipolar decomposition [Blanchet, 1998; Blanchet and Damour, 1986; Thorne, 1980] has been worked out [Kopeikin and Korobkov, 2005; Kopeikin *et al.*, 2006; Kopeikin, 1997c; Kopeikin and Makarov, 2007; Kopeikin *et al.*, 1999]. One has used a special technique of integration of the equations of light propagation based on the retarded time argument, which allowed to discover a rigorous solution of the equations everywhere outside the localized system including both the near and radiative zones. This research was oriented towards development of a theoretical tool for detection of gravitational waves by means of astrometric techniques - the topic of active discussion in literature [Dubath *et al.*, 2007; Jenet *et al.*, 2004; Kopeikin and Korobkov, 2005; Kopeikin *et al.*, 1999; Lesovik *et al.*, 2005; Schluessel, 2008].

Theoretical algorithms of a micro-arcsecond relativistic astrometry in the solar system should take into account relativistic effects in propagation of light in the field of point-like massive bodies having arbitrary translational motion in space and rotating. This model of the N-body system closely resembles the solar system and, thus, is practically important for astrometric missions. Solution of this problem was pioneered by Brumberg [Brumberg, 1972, 1991] and further developed in [Brumberg *et al.*, 1990; Klioner, 2003a; Klioner and Kopeikin, 1994]. However, the most powerful approach has been developed in papers [Kopeikin and Mashhoon, 2002; Kopeikin, 2009; Kopeikin and Schäfer, 1999]. It is based on the mathematical technique of the retarded *Liénard-Wiechert* potentials of the gravitational field of the moving masses [Landau and Lifshitz, 1975] which drastically simplifies the integration of the light-ray geodesics and allows us to formulate the final result for the gravitational deflection angle, time delay, and frequency change in terms of the Lorentz-invariant expressions.

The present chapter summarizes results of the previous studies and describes the integration technique for constructing a Lorentz covariant solution of equations of propagation of light rays both outside and inside of a gravitating system of massive point-like particles moving along arbitrary worldlines. In finding the solution one uses the *Liénard-Wiechert* presentation for the metric tensor, which accounts for all possible effects in the description of the gravitational field and is valid everywhere outside the worldlines of the moving bodies. The solution shows that the time t_a of the closest approach of light ray to body a is actually irrelevant and must be replaced by the retarded time, which originates from the fact that the gravitational field propagates on the null cone of the background Minkowski spacetime. The time of the closest approach is a good approximation of the retarded time but it does not enter the final result for observable astrometric quantities. Straightforward calculation yields the complete expressions for the angle of light deflection, relativistic time delay, and gravitational shift of observed electromagnetic frequency of the emitted photons. These expressions are exact at the linear approximation with respect to the universal gravitational constant, G, and at arbitrary order of magnitude with respect to the parameter v_a/c, where v_a is a characteristic velocity of the a-th light-deflecting body, and c is the fundamental speed. One discusses practical applications of the retarded solution of the equations of light propagation including moving gravitational lenses, timing of binary pulsars, the consensus model of very long baseline interferometry, and the relativistic reduction of astrometric observations in the solar system.

The mathematical formalism of the present chapter can be also used in astrometric experiments for testing alternative scalar-tensor theories of gravity after formal replacing in all subsequent formulas the universal gravitational constant G by the product $G(\gamma^* + 1)/2$, where γ^* is the effective light-ray deflection parameter, which is slightly different from its weak-field limiting value γ of the standard parameterized post-Newtonian (PPN) formalism [Will, 1993]. In the weak-field limit the numerical value of γ^* coincides with that of γ as discussed in [Damour and Esposito-Farèse, 1996; Nordtvedt, 1985]. The difference between the parameters γ^* and γ reveals only in static terms of the second post-Newtonian order of magnitude depending on the internal structure of the bodies. This statement is a direct consequence of a conformal invariance of the equations of light rays [Damour and Esposito-Farèse, 1998b]. In what follows, this difference is ignored because the solar system experiments are not sensitive enough to detect the difference between the two parameters. However, it may play a role in the binary pulsars analysis [Nordtvedt, 1985].

The chapter is organized as follows. Section 7.2 presents a short description of the energy-momentum tensor of the light-deflecting bodies and the metric tensor given in the form of the *Liénard-Wiechert* potential. Section 7.3 is devoted to the development of a mathematical technique for integrating equations of propagation of electromagnetic waves in the geometric optics approximation. Formal solution of these equations and relativistic perturbations of a photon trajectory are given in section 7.4. Section 7.5 deals with a general treatment of observable relativistic effects - the integrated time delay, the deflection angle, and gravitational shift of frequency. Particular cases of application astrometry in astrophysics are presented in section 7.6. They include the Shapiro time delay in binary pulsars and moving gravitational lenses. General relativistic astrometry in the solar system is presented in section 7.7. Finally, section 7.8 discusses relativistic description of the Doppler tracking of interplanetary spacecrafts.

7.2
Gravitational Liénard-Wiechert Potentials

Our interpretation of observable quantities in relativistic astrometry is based solely on general relativity and the physical meaning of the retarded *Liénard-Wiechert* potentials used to solve the Einstein equations [Bel *et al.*, 1981; Kopeikin and Schäfer, 1999]. Built directly from Einstein's equations, these potentials describe the complete, relativistically invariant, time-varying gravitational field for a point masses in arbitrary motion. Gravitational radiation in the form of waves can be obtained from these potentials as well. Remarkable property of the *Liénard-Wiechert* potentials is that they describe the propagation of the gravity field even if the mass moves uniformly with constant speed. One emphasizes that although gravitational waves are not generated by a uniformly moving body the null characteristics of the gravitational field are still precisely defined by the *Liénard-Wiechert* potentials through their property of the Lorentz invariance and the principle of physical causality [Kopeikin and Fomalont, 2006]. In the near-field zone of a gravitating N-body system the null characteristics of gravity reveal themselves through the Lorentz invariance of the gravitational force (the aberration of gravity) when one compares observations done in one inertial frame with another, while in the far-field radiative zone the null characteristics of gravity can be traced as freely propagating gravitational waves.

In general relativity light and gravity null rays form bi-characteristic hypersurfaces [Frolov, 1979], that is they propagate in space with the same fundamental speed c. At the first glance it may look that the relativistic astrometry can not discriminate the propagational effects of gravity from those associated with light. However, this point of view is short-sighted. Although light and gravity propagate with the same speed on the hypersurface of one and the same null cone, they generally propagate in different directions in space in each particular gravitational experiment and are associated with different physical effects. Our point of view is that the most natural interpretation of astrometric observations of relativistic effects should rest solidly on general relativity which has passed all other experimental tests in strong compliance with the current theoretical understanding of gravitational physics. In other words, in order to interpret the astrometric observation properly, the spacetime properties, which general relativity postulates, must not be violated. Specifically, spacetime is a differentiable manifold endowed with an affine connection (Christoffel symbols) whose geodesics form a privileged set of worldlines in spacetime, and their knowledge (observation) allows us to extract information about the curvature tensor (tidal gravitational force) without ambiguity. The Christoffel symbols define the gravitational law of motion of test particles and extended bodies and, hence, generalize the concept of the Newtonian gravitational force in general relativity. They are formed from partial derivatives of the metric tensor that defines geometric properties of the spacetime and the causal structure of the null cone. At the same time the metric tensor is associated with the gravitational potentials whose properties are determined by Einstein's equations.

The solution of the Einstein equations, using the *Liénard-Wiechert* potentials has been essentially understood and developed by previous researchers. This iterative procedure is called post-Minkowskian approximations [Damour, 1987]. For the purpose of this chapter the first (linearized gravity) post-Minkowskian approximation is sufficient since the gravitational light-ray deflection experiments are not yet sensitive enough to measure non-linearities of the Einstein equations. The metric tensor in the linearized approximation reads

$$g_{\alpha\beta}(t,\boldsymbol{x}) \quad = \quad \eta_{\alpha\beta} + h_{\alpha\beta}(t,\boldsymbol{x}) \;, \tag{7.1}$$

where t is coordinate time, $\boldsymbol{x} = x^i = (x^1, x^2, x^3)$ denotes spatial coordinates of a field point in space, $\eta_{\alpha\beta} = \mathrm{diag}(-1,+1,+1,+1)$ is the Minkowski metric of the background flat spacetime and the metric perturbation $h_{\alpha\beta}(t,\boldsymbol{x})$ is a function of time and spatial coordinates [2]. It can be found by solving the Einstein field equations in harmonic gauge. One reminds that the harmonic gauge [Fock, 1964; Weinberg, 1972] is fixed in the first post-Minkowskian approximation by the four differential conditions

$$\eta^{\mu\nu}\left(\partial_\mu h_{\alpha\nu} - \frac{1}{2}\partial_\alpha h_{\mu\nu}\right) = 0 \;, \tag{7.2}$$

where $\partial_\mu = \partial/\partial x^\mu$ denotes a partial derivative with respect to a corresponding coordinate. Taking into account the harmonic condition (7.2) reduces the Einstein equations in the first post-Minkowskian approximation to the wave equation for the metric tensor perturbations ([Weinberg, 1972, chapter 10]

$$\Box\, h_{\alpha\beta}(t,\boldsymbol{x}) \quad = \quad -\frac{16\pi G}{c^4} S_{\alpha\beta}(t,\boldsymbol{x}) \;, \tag{7.3}$$

where the *D'Alambertian (wave) operator*

$$\Box = -\frac{1}{c^2}\partial_t^2 + \boldsymbol{\nabla}^2 \;, \tag{7.4}$$

$\partial_t = \partial/\partial t$, the Laplace differential operator $\boldsymbol{\nabla}^2 = \delta^{ij}\partial_i\partial_j$,

$$S_{\alpha\beta}(t,\boldsymbol{x}) \quad = \quad T_{\alpha\beta}(t,\boldsymbol{x}) - \frac{1}{2}\eta_{\alpha\beta}\, T^\lambda{}_\lambda(t,\boldsymbol{x}) \;, \tag{7.5}$$

and $T_{\alpha\beta}(t,\boldsymbol{x})$ is the tensor of energy-momentum of the system of massive bodies. This tensor for the case of a point-like bodies in a covariant form is given, for example, by Landau & Lifshitz [Landau and Lifshitz, 1975]

$$T^{\alpha\beta}(t,\boldsymbol{x}) \quad = \quad \sum_{a=1}^{N} T_a^{\alpha\beta}(t)\delta\left(\boldsymbol{x} - \boldsymbol{x}_a(t)\right) \;, \tag{7.6}$$

$$T_a^{\alpha\beta}(t) \quad = \quad m_a c^2 \gamma_a^{-1}(t) u_a^\alpha(t) u_a^\beta(t) \;, \tag{7.7}$$

2) It is worthwhile to remind that the spacetime indices $\alpha, \beta, ...$, etc. runs from 0 to 3 and are raised and lowered by means of $\eta_{\alpha\beta}$. Spatial indices $i, j, k, ...$, etc. run from 1 to 3 and are raised and lowered by means of the Kronecker symbol δ_{ij}, so that, actually, the upper and lower case spatial indices are not distinguished. Repeated Greek and Latin indices denote summation from 0 to 3 and from 1 to 3 respectively.

where m_a is the constant (relativistic) rest mass of the a-th particle, $\boldsymbol{x}_a(t)$ are spatial coordinates of the a-th massive particle which depend on time t, $\boldsymbol{v}_a(t) = d\boldsymbol{x}_a(t)/dt$ is velocity of the a-th particle, $\gamma_a(t) = 1/\sqrt{1-\beta_a^2(t)}$ is the (time-dependent) Lorentz factor, $\boldsymbol{\beta}_a = \boldsymbol{v}_a/c$, $u_a^\alpha(t) = \{\gamma_a(t),\ \gamma_a(t)\boldsymbol{\beta}_a(t)\}$ is the four-velocity of the a-th particle normalized to -1, $\delta(\boldsymbol{x})$ is the usual 3-dimensional Dirac delta-function. In particular, one has for components of the tensor of energy-momentum

$$T_a^{00}(t) = \frac{m_a c^2}{\sqrt{1-v_a^2(t)}}\ , \qquad T_a^{0i}(t) = \frac{m_a c^2 \beta_a^i(t)}{\sqrt{1-\beta_a^2(t)}}\ , \qquad T_a^{ij}(t) = \frac{m_a c^2 \beta_a^i(t)\beta_a^j(t)}{\sqrt{1-\beta_a^2(t)}}\ . \tag{7.8}$$

Solution of equations (7.3)-(7.8) has the form of the *Liénard-Wiechert* potential [Jackson, 1975]. In order to see how it looks like, the tensor of energy-momentum is represented in a form where the whole time-dependence history of the moving particle is formally included in a one-dimensional delta-function

$$T^{\alpha\beta}(t,\boldsymbol{x}) = \int_{-\infty}^{+\infty} dt'\,\delta(t'-t)\,T^{\alpha\beta}(t',\boldsymbol{x})\ . \tag{7.9}$$

Here t' is an independent parameter along the worldlines of the particles which does not depend on time t. The solution of equation (7.3) can be found by making use of the Green function's technique [Barton, 1989]. One takes only the retarded Green function and abandons the time-advanced solution as the N-body system under consideration is isolated from possible external gravitational environment. It is equivalent to the assumption that there is no gravitational radiation impinging onto the system in the first post-Minkowskian approximation. It is interesting to emphasize that in higher post-Minkowskian approximations existence of the tail gravitational radiation effects [Blanchet and Damour, 1988; Blanchet and Schäfer, 1993; Schäfer, 1990] brings about a small fraction of incoming radiation as being backscattered on the static part of the curvature generated by the monopole component in multipole expansion of the metric tensor. However, although the backscattered radiation is incoming, it does not come in from past null infinity called sometimes "scri minus" in relativity [Misner *et al.*, 1973, page 917]. Therefore, it has nothing to do with the advanced Green function and represents a purely outgoing radiation at future null infinity ("scri plus") [Misner *et al.*, 1973, page 917]. One omits such backscatter terms in what follows, for they appear only in the higher orders of the post-Minkowskian approximation scheme.

Choosing the retarded Green function of the wave equation and integrating with respect to spatial coordinates by making use of the properties of the 3-dimensional delta-function, one gets the metric tensor perturbation in the form of an one-dimensional, retarded-time integral

$$h^{\alpha\beta}(t,\boldsymbol{x}) = \sum_{a=1}^{N}\int_{-\infty}^{+\infty} h_a^{\alpha\beta}(t',t,\boldsymbol{x})dt' \;, \tag{7.10}$$

$$h_a^{\alpha\beta}(t',t,\boldsymbol{x}) = \frac{4G}{c^4}\left[T_a^{\alpha\beta}(t') - \frac{1}{2}\eta^{\alpha\beta}T^{\lambda}_{a\lambda}(t')\right]\frac{\delta\left[t'-t+\frac{1}{c}r_a(t')\right]}{r_a(t')} \;, \tag{7.11}$$

where $\boldsymbol{r}_a(t') = \boldsymbol{x} - \boldsymbol{x}_a(t')$, and $r_a(t') = |\boldsymbol{r}_a(t')|$ is the usual Euclidean length of the vector. The integral (7.10) can be performed explicitly as described in, e.g., [Jackson, 1975, section 14]). The result is the retarded *Liénard-Wiechert* tensor potential

$$h^{\alpha\beta}(t,\boldsymbol{x}) = \frac{4G}{c^4}\sum_{a=1}^{N}\frac{T_a^{\alpha\beta}(s) - \frac{1}{2}\eta^{\alpha\beta}T^{\lambda}_{a\lambda}(s)}{r_a(s) - \boldsymbol{\beta}_a(s)\cdot\boldsymbol{r}_a(s)} \;, \tag{7.12}$$

where the retarded time $s = s(t,\boldsymbol{x})$ for the a-th body is a solution of the gravity null-cone equation

$$s = t - \frac{1}{c}|\boldsymbol{x} - \boldsymbol{x}_a(s)| \;, \tag{7.13}$$

describing propagation of gravity from the moving body a to the field point, $x^\alpha = (t,\boldsymbol{x})$, with the fundamental speed c. Here, it is assumed that the gravity field is measured at time t and at the point $\boldsymbol{x}$.

Equation (7.13) is a complicated non-linear equation, which can not be solved analytically for arbitrary worldlines of the massive bodies. Nevertheless, analytic solution is possible if the bodies move along straight lines with constant velocities

$$\boldsymbol{x}_a(t) = \boldsymbol{x}_a(t_a) + \boldsymbol{v}_a(t - t_a) \;, \tag{7.14}$$

where t_a is an arbitrary instant of time,and $\boldsymbol{v}_a$ is constant. Substituting equation (7.14) to the retarded time equation (7.13) brings about a quadratic algebraic equation for s

$$c^2(s-t)^2 = |\boldsymbol{x} - \boldsymbol{x}_a(t_a) + \boldsymbol{v}_a(s - t_a)|^2 \;, \tag{7.15}$$

which can be solved exactly. The solution is [Kopeikin, 2009, equation 25]

$$s = t - \frac{\boldsymbol{R}\cdot\boldsymbol{\beta}_a + \sqrt{R^2 - (\boldsymbol{R}\cdot\boldsymbol{\beta}_a)^2}}{c(1-\beta_a^2)} \;, \tag{7.16}$$

and $\boldsymbol{R} = \boldsymbol{x} - \boldsymbol{x}_a(t)$ with $\boldsymbol{x}_a(t)$ defined in equation (7.14).

One points out that it would be more appropriate to denote the retarded time for the a-th body as s_a, which would have reflected the dependence of the retarded time on the number of the body. However, it would make notations and the presentation of subsequent formulas more cumbersome. For this reason the notation s is used instead of s_a keeping in mind that if it is not stated otherwise, coordinates, velocity, and acceleration of the a-th body are taken at the corresponding retarded time s_a.

This remark is crucial, e.g., in the discussion regarding the definition of the center of mass of the N body system [3].

Equations (7.12) and (7.13) of the metric perturbation $h_{\alpha\beta}(t, \boldsymbol{x})$ are used for integration of the equations of light geodesics in the next section. It is worth emphasizing that the expression for the metric tensor (7.12) is Lorentz-covariant and is valid in any harmonic coordinate system admitting a smooth transition to the asymptotically flat spacetime at infinity and relating to each other by the Lorentz transformations of the theory of special relativity [Fock, 1964]. A treatment of post-linear corrections to the *Liénard-Wiechert* potentials (7.12) is given, for example, in a series of papers by Kip Thorne and collaborators [Crowley and Thorne, 1977; Kovacs and Thorne, 1977, 1978; Thorne and Kovacs, 1975].

It is also interesting to notice that if the condition of the weak metric tensor perturbation is imposed, that is $|h^{\alpha\beta}| \ll 1$, it will lead to a stronger upper limit on the velocity of the moving bodies than the simple $\beta_a < 1$ [Kovacs and Thorne, 1978; Mashhoon, 1992; Westpfahl, 1985]. This limitation follows from equation (7.12) for the *Liénard-Wiechert* potentials. For example, if velocity $\boldsymbol{v}_a$ is perpendicular to a radius-vector $\boldsymbol{r}_a$ (circular orbit of a coalescing binary system), the weak-field approximation imposed on the *Liénard-Wiechert* potentials demands

$$\frac{Gm_a}{c^2 r_a} \ll (1-\beta_a^2)^{1/2} \qquad \longrightarrow \qquad \beta_a \ll \sqrt{1-\left(\frac{Gm_a}{c^2 r_a}\right)^2}\,. \tag{7.17}$$

In another particular case, where velocity $\boldsymbol{v}_a$ of the body is almost parallel to $\boldsymbol{r}_a$ one gets a stronger restriction:

$$\frac{Gm_a}{c^2 r_a} \ll (1-\beta_a)^{3/2} \qquad \longrightarrow \qquad \beta_a \ll 1-\left(\frac{Gm_a}{c^2 r_a}\right)^{2/3}\,. \tag{7.18}$$

These limitations are always satisfied in the solar system, because of its weak gravitational field $(Gm_a)/(c^2 r_a) \ll 1$, but may be violated, for example, in coalescing binary systems, where gravitational field is not weak just before the coalescence.

3) For more detail see explanation after equation (7.209).

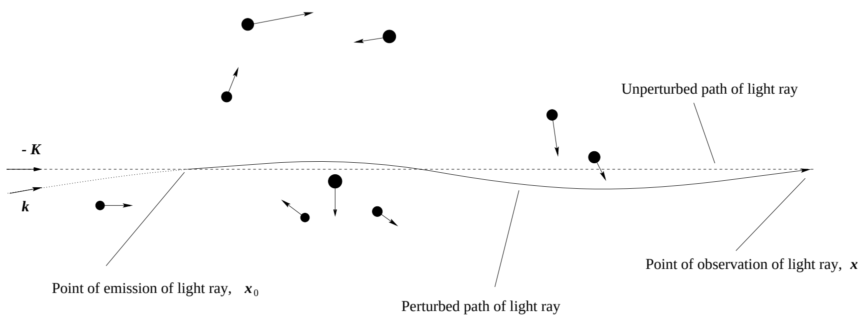

Figure 7.1 A light ray is emitted at the instant of time t_0 at the point $\boldsymbol{x}_0$ and arrives at the point of observation $\boldsymbol{x}$ at the instant of time t. Light-ray deflecting bodies move along accelerated worldlines during the time of propagation of the light ray; their velocities at some intermediate instant of time are shown by black arrows. In the absence of the light-ray deflecting bodies the light ray would propagate along an unperturbed path (the dashed line) which is a straight line passing through the points of emission, $\boldsymbol{x}_0$, and observation, $\boldsymbol{x}$. Unperturbed direction from observer to the source of light is determined by the unit vector $\boldsymbol{K} = -(\boldsymbol{x} - \boldsymbol{x}_0)/|\boldsymbol{x} - \boldsymbol{x}_0|$. In the presence of the light-ray-deflecting bodies the light ray propagates along the perturbed path (the solid line). The perturbed trajectory of the light ray is bent and twisted due to the effects of gravitoelectric (mass-induced) and gravitomagnetic (velocity-induced) fields of the bodies [Kopeikin and Fomalont, 2007]. The initial condition (7.19) for the equation of light propagation is determined by the unit vector $\boldsymbol{k}$ defined by the backward-in-time extension of the solution of the light-ray equation to past null infinity passing through the point of emission, $\boldsymbol{x}_0$, of light. Relationship between the unit vectors $\boldsymbol{k}$ and $\boldsymbol{K}$ includes relativistic bending of light and is given by equation (7.67).

7.3 Mathematical Technique for Integrating Equations of Propagation of Photons

One considers the motion of a light particle (photon) in the background gravitational field described by the metric (7.12). No back action of the photon on the gravitational field is assumed. Moreover, one will assume that there is no medium and photons propagate in vacuum. It allows to use equations of light geodesics directly applying the metric tensor (7.12) for calculation of the gravitational perturbation of the photon's trajectory. In real practice one can not ignore the presence of medium, and should take into account its effect on propagation of light. Astronomers have developed powerful methods for dealing with this problem and the reader is referred to a corresponding literature (see, for example, [Ginzburg, 1970; Yakovlev, 2002]) as discussion of this topic would take us too far beyond the scope of this book.

Let the motion of the photon be defined by fixing the initial-boundary conditions (see Figure 7.1)

$$\boldsymbol{x}(t_0) = \boldsymbol{x}_0 \,, \qquad \frac{d\boldsymbol{x}(-\infty)}{dt} = c\boldsymbol{k} \,, \tag{7.19}$$

where $\boldsymbol{k}^2 = 1$ and, henceforth, the spatial components of vectors are denoted by bold letters. These conditions define the coordinates $\boldsymbol{x}_0$ of the photon at the moment of emission of light, t_0, and its velocity at the infinite past and infinite distance from the origin of the spatial coordinates (that is, at the so-called, past null infinity [Misner *et al.*, 1973, page 917]).

The original equation of propagation of light rays in vacuum on a curved spacetime manifold is just a light-ray geodesic

$$\frac{d^2x^\alpha}{d\lambda^2}+\Gamma^\alpha_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}=0\,, \tag{7.20}$$

where λ is the affine parameter along the worldline of the light particle (photon), and the Christoffel symbols should be calculated in the linearized approximation as follows

$$\Gamma^\alpha_{\mu\nu}=\frac{1}{2}\eta^{\alpha\beta}\left(\partial_\nu h_{\beta\mu}+\partial_\mu h_{\beta\nu}-\partial_\beta h_{\mu\nu}\right)\,. \tag{7.21}$$

The affine parameter is related to coordinate time t by equation

$$c\frac{d^2t}{d\lambda^2}+\Gamma^0_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}=0\,, \tag{7.22}$$

which is just a time component of equation (7.20). After replacing the affine parameter λ with coordinate time t in equation (7.20), it can be reduced to equation for only spatial components of the photon

$$\frac{d^2x^i}{dt^2}+\Gamma^i_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}-\frac{1}{c}\Gamma^0_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}\frac{dx^i}{dt}=0\,. \tag{7.23}$$

Making use of equation (7.21) yields components of the Christoffel symbols

$$\Gamma^0_{00} = -\frac{1}{2c}\partial_t h_{00}\,, \tag{7.24}$$

$$\Gamma^0_{0k} = -\frac{1}{2}\partial_k h_{00}\,, \tag{7.25}$$

$$\Gamma^0_{kp} = -\frac{1}{2}\left(\partial_p h_{0k}+\partial_k h_{0p}-\frac{1}{c}\partial_t h_{kp}\right)\,, \tag{7.26}$$

$$\Gamma^i_{00} = -\frac{1}{2}\partial_i h_{00}+\frac{1}{c}\partial_t h_{0i}\,, \tag{7.27}$$

$$\Gamma^i_{0k} = -\frac{1}{2}\left(\partial_i h_{0k}-\partial_k h_{0l}-\frac{1}{c}\partial_t h_{ik}\right)\,, \tag{7.28}$$

$$\Gamma^i_{kp} = -\frac{1}{2}\left(\partial_i h_{kp}-\partial_p h_{ik}-\partial_k h_{ip}\right)\,, \tag{7.29}$$

where $\partial_t=\partial/\partial t$ and $\partial_i=\partial/\partial x^i$. Substituting them to the light ray equation (7.23) yields

$$\begin{aligned}\ddot{x}^i =\;& \frac{c^2}{2}\partial_i h_{00}-c\partial_t h_{0i}-c\left(\partial_k h_{0i}-\partial_i h_{0k}+\frac{1}{c}\partial_t h_{ik}\right)\dot{x}^k-\frac{1}{2}\partial_t h_{00}\dot{x}^i \\ -\;& \partial_k h_{00}\dot{x}^k\dot{x}^i-\left(\partial_p h_{ik}-\frac{1}{2}\partial_i h_{kp}\right)\dot{x}^k\dot{x}^p-\left(\frac{1}{c}\partial_p h_{0k}-\frac{1}{2c^2}\partial_t h_{kp}\right)\dot{x}^p\dot{x}^k\dot{x}^i\,,\end{aligned} \tag{7.30}$$

where $\dot{x}^i = dx^i/dt$, and $\ddot{x}^i = d^2x^i/dt^2$, and one has put all terms depending on the metric tensor perturbations $h_{\alpha\beta}$ to the right side so that they can be interpreted as a gravity force acting on the propagating photon by moving massive bodies.

The propagation equation (7.30) looks rather complicated and difficult for solution. However, it can be simplified and reduced to much simpler form that admits direct integration along the null cone with a method of successive approximations [Kopeikin and Schäfer, 1999; Kopeikin *et al.*, 1999]. In the first approximation the propagation of light ray is not affected by gravity force, and equation (7.30) is reduced to a rather trivial one

$$\ddot{x}^i = 0 . \tag{7.31}$$

Solution of this equation is the unperturbed trajectory of the light ray in a flat spacetime with the Minkowski metric $\eta_{\alpha\beta}$ that is a straight line

$$x^i(t) = x^i_N(t) = x^i_0 + ck^i \ (t - t_0) \ , \tag{7.32}$$

where t_0, x^i_0, and the unit vector $k^i = \boldsymbol{k}$ have been defined in equation (7.19). In this approximation, the coordinate speed of the photon is $\dot{x}^i = ck^i$ and is considered as a constant in the expression for the light-ray-perturbing force. In order to solve equation (7.30) in the second approximation one needs to understand better the properties of the unperturbed solution (7.32).

It is convenient to introduce a new independent parameter τ along the photon's trajectory according to the rule [Kopeikin *et al.*, 1999], [Kopeikin, 1997c]

$$\tau = \frac{1}{c}\boldsymbol{k}\cdot\boldsymbol{x}_N(t) = t - t_0 + \frac{1}{c}\boldsymbol{k}\cdot\boldsymbol{x}_0, \tag{7.33}$$

where here and everywhere else the dot between two spatial vectors denotes their Euclidean dot product. The time $t = t_0$ of the light signal's emission corresponds to the numerical value of the parameter $\tau_0 = c^{-1}\boldsymbol{k}\cdot\boldsymbol{x}_0$, and the numerical value of the parameter $\tau = 0$ corresponds to the time

$$t^* = t_0 - \frac{1}{c}\boldsymbol{k}\cdot\boldsymbol{x}_0 \ , \tag{7.34}$$

which is the time of the closest approach of the unperturbed trajectory of the photon to the origin of the coordinate system used for calculation of the metric tensor perturbations. One emphasizes that the numerical value of the moment t^* is constant for a chosen trajectory of light ray (that is, for another light ray it will have a different numerical value) and depends only on the spacetime coordinates of the point of emission of the photon and the point of its observation. Thus, one finds the relationships

$$\tau \equiv t - t^*, \qquad \tau_0 = t_0 - t^* \ , \tag{7.35}$$

which reveal that the new variable τ is negative from the point of emission up to the point of the closest approach $x^i(t^*) \equiv \xi^i$, and is positive otherwise [4]. The differential identity $dt = d\tau$ is valid and, for this reason, the integration along the light ray's path with respect to time t can be always replaced by the integration with respect to variable τ.

Making use of the parameter τ, the equation (7.32) of the unperturbed trajectory of the light ray can be represented as

$$x^i(\tau) = x^i_N(\tau) = ck^i\tau + \xi^i \,, \tag{7.36}$$

and the distance, $r(\tau) = |\boldsymbol{x}_N(t)|$, of the photon from the origin of the coordinate system obeys to the Pythagorean relationship

$$r(\tau) = \sqrt{c^2\tau^2 + d^2} \,. \tag{7.37}$$

The constant vector $\xi^i = \boldsymbol{\xi} = \boldsymbol{k} \times (\boldsymbol{x}_0 \times \boldsymbol{k}) = \boldsymbol{k} \times (\boldsymbol{x}_N(t) \times \boldsymbol{k})$ is called the impact parameter of the unperturbed trajectory of the light ray, $d = |\boldsymbol{\xi}|$ is the length of the impact parameter, and the symbol " $\times$ " between two vectors denotes the usual Euclidean cross product of two vectors. One notes that vector $\boldsymbol{\xi}$ is transverse to vector $\boldsymbol{k}$ and directed from the origin of the coordinate system towards the point of the closest approach of the light ray to the origin. One emphasizes that this vector plays just an auxiliary role in the discussion and, in general, has no essential physical meaning as it can be easily changed by the shift of the origin of the coordinates. At the given stage the freedom in choosing the origin of the coordinate system is not fixed by assuming, for example, that the origin coincides with the center of mass of the N-body system. Specific choices of the coordinate origin will be done later on.

Let us consider a set (bundle) of unperturbed light rays $x^i_N(\tau) = ck^i\tau + \xi^i$ with different values of vectors k^i and ξ^i. The unit vector field k^i, defined along an each line $x^i_N(\tau)$, describes the direction of a bundle of light rays and introduces a natural $(2+1)$ splitting of 3-dimensional space. The vector ξ^i defines a point of intersection of any of those rays with the plane being orthogonal to the bundle of the light rays. This vector does not depend on τ and can be given in the form of the following relationship

$$\xi^i = P^i{}_j x^j \,, \tag{7.38}$$

where

$$P^{ij} = \delta^{ij} - k^i k^j \,, \tag{7.39}$$

is the operator of projection onto the plane being orthogonal to vector k^i. The operator has only two algebraically independent components and satisfies the relationship

$$P^{ik}P_{kj} = P^i{}_j \,. \tag{7.40}$$

4) It is worth noting that this statement is true only if the origin of the coordinate system is between the source of light and observer. Under other circumstances the variable τ may be always either positive or negative.

Because of this property one can recast equation (7.38) into the form

$$\xi^i = P^i_j \xi^j \,, \tag{7.41}$$

which shows explicitly that the vector ξ^i is constrained to lie in the 2-dimensional plane. Thus, the operation of partial differentiation in this plane is

$$\frac{\partial \xi^i}{\partial \xi^j} = P^i_j \,. \tag{7.42}$$

It is worth noting that the projection operator can be used to raise and lower indices of any geometrical object lying in the plane orthogonal to vector k^i.

Implementing the two new parameters τ, $\boldsymbol{\xi}$ and introducing the four-dimensional light-ray vector $k^\alpha = (1, k^i)$ allows us to derive the principal differential identity of the bundle of the light geodesics. Specifically, for any smooth function $F(t, \boldsymbol{x})$ taken on the bundle of the light rays one has

$$\begin{aligned} &\left[\frac{\partial F(t,\boldsymbol{x})}{\partial x^i} + \frac{k_i}{c}\frac{\partial F(t,\boldsymbol{x})}{\partial t}\right]_{\boldsymbol{x}=c\boldsymbol{k}\,(t-t_0)+\boldsymbol{x}_0} = \\ &\frac{\partial F(\tau+t^*, c\boldsymbol{k}\tau+\boldsymbol{\xi})}{\partial \xi^i} + \frac{k_i}{c}\frac{\partial F(\tau+t^*, c\boldsymbol{k}\tau+\boldsymbol{\xi})}{\partial \tau} \,, \end{aligned} \tag{7.43}$$

where the equation for the light-ray trajectory is substituted to the left side of this equation *after* calculating the partial derivatives from $F(t, \boldsymbol{x})$ while in the right side we, first, substitute the light-ray trajectory to $F(t, \boldsymbol{x}) = F(\tau + t^*, c\boldsymbol{k}\tau + \boldsymbol{\xi})$ and, then, differentiate. Equation (7.43) demonstrates that one can always switch the order of mathematical operations between the taking the partial derivative and the substitution for the unperturbed light ray trajectory to function $F(t, \boldsymbol{x})$. This identity helps us to significantly simply the original equation (7.30 for light geodesics to continue its integration in the second approximation. One also notices two other, useful relationships for integrals from function $F(\tau, \boldsymbol{\xi})$, which decays to zero value at past null infinity sufficiently fast. Namely, one has

$$\int_{-\infty}^{\tau} \frac{dF(\sigma,\boldsymbol{\xi})}{d\sigma} d\sigma = F(\tau, \boldsymbol{\xi}) \,, \tag{7.44}$$

$$\int_{-\infty}^{\tau} \hat{\partial}_i F(\sigma,\boldsymbol{\xi}) d\sigma = \hat{\partial}_i \int_{-\infty}^{\tau} F(\sigma,\boldsymbol{\xi}) d\sigma \,, \tag{7.45}$$

where the first equation is almost self-evident, and the second one is fulfilled because the variables σ and $\boldsymbol{\xi}$ are independent for the bundle of the light rays.

Substituting the unperturbed velocity of the light ray, $\dot{x}^i = ck^i$, to the right side of equation (7.30) and applying equation (7.43) to the metric tensor perturbation $h_{\alpha\beta}$, one can re-write equation (7.30) of light geodesics as follows [Kopeikin and Schäfer, 1999; Kopeikin *et al.*, 1999]

$$\begin{aligned} \frac{d^2 x^i(\tau)}{d\tau^2} &= \frac{1}{2} c^2 k_\alpha k_\beta \hat{\partial}_i h^{\alpha\beta}(\tau, \boldsymbol{\xi}) \\ &- c\frac{d}{d\tau}\left[k_\alpha h^{\alpha i}(\tau,\boldsymbol{\xi}) + \frac{1}{2}k^i h^{00}(\tau,\boldsymbol{\xi}) - \frac{1}{2}k^i k_p k_q h^{pq}(\tau,\boldsymbol{\xi})\right] \,, \end{aligned} \tag{7.46}$$

where one has introduced a new notation for a partial derivative with respect to vector ξ^i of the impact parameter: $\hat{\partial}_i \equiv P_{ij}\partial/\partial\xi^j$, and all quantities on the right side of equation (7.46) are taken along the light-ray trajectory at the point corresponding to a numerical value of the running parameter τ while the parameter $\boldsymbol{\xi}$ is assumed as constant. In other words, equation (7.46) should be considered as an ordinary, second-order differential equation in time variable τ. More precisely, this kind of equation is known in the literature as "retarded-functional differential system" because of the dependence of the gravitational potentials on the retarded time argument s (see equation (7.12)). Such equations belong to the framework of "predictive relativistic mechanics" [Bel, 1970; Bel *et al.*, 1981; Bel and Fustero, 1976; Bel *et al.*, 1973; Currie, 1966; Damour, 1983; Hill, 1967].

The given form of equation (7.46) already shows that only the first term on the right hand side of it can contribute to the deflection of light if observer and source of light are at spatial infinity. Indeed, integration of the right side of equation (7.46) with respect to time from $-\infty$ to $+\infty$ brings all terms showing time derivative, to zero due to the asymptotic flatness of the metric tensor.

However, if the observer and the source of light are located at finite distances from the origin of the coordinate system, one needs to know how to perform the integrals from the metric perturbations (7.10) with respect to the parameter τ along the unperturbed trajectory of light ray. Let us denote those integrals as

$$B^{\alpha\beta}(\tau,\boldsymbol{\xi}) = c\int_{-\infty}^{\tau} h^{\alpha\beta}[\sigma,\boldsymbol{x}(\sigma)]d\sigma \,, \tag{7.47}$$

$$D^{\alpha\beta}(\tau,\boldsymbol{\xi}) = c\int_{-\infty}^{\tau} B^{\alpha\beta}(\sigma,\boldsymbol{\xi})d\sigma \,, \tag{7.48}$$

where the metric perturbation $h^{\alpha\beta}[\sigma,\boldsymbol{x}(\sigma)]$ is defined by the *Liénard-Wiechert* potential (7.10) and σ is a parameter of integration along the light ray having the same meaning as the parameter τ in equation (7.33). In order to calculate the integrals (7.47), (7.48) it is useful to change in the integrands the time argument, σ, to the new one, ζ, defined by the gravity null-cone equation (7.13) which, after substitution for $\boldsymbol{x}$ the unperturbed light-ray trajectory (7.36), reads as follows [5)]

$$\sigma + t^* = \zeta + \frac{1}{c}|\boldsymbol{\xi} + c\boldsymbol{k}\sigma - \boldsymbol{x}_a(\zeta)| \,. \tag{7.49}$$

Total differential of this equation yields a relationship between differentials of the time variables σ and ζ, and parameters t^*, ξ^i, k^i. More specifically,

$$d\zeta\,(r_a - \boldsymbol{\beta}_a\cdot\boldsymbol{r}_a) = d\sigma\,(r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a) + r_a dt^* - \frac{1}{c}\boldsymbol{r}_a\cdot d\boldsymbol{\xi} - \sigma\boldsymbol{r}_a\cdot d\boldsymbol{k} \,, \tag{7.50}$$

where the coordinates, $\boldsymbol{x}_a = \boldsymbol{x}_a(\zeta)$, and the velocity, $\boldsymbol{v}_a = \boldsymbol{v}_a(\zeta)$, of the a-th body are taken at the retarded time ζ, and coordinates of the photon, $\boldsymbol{x} = \boldsymbol{x}(\sigma)$, are taken at the

5) One again emphasizes that the new parameter ζ depends on the index of each body. For this reason it would be reasonable to denote it as ζ_a. One does not use this notation to avoid the appearance of a large number of sub-indices.

time $\sigma = \sigma(\zeta)$. From equation (7.50) one immediately obtains the partial derivatives with respect to the parameters

$$\frac{\partial\zeta}{\partial t^*} = \frac{r_a}{r_a - \boldsymbol{\beta}_a \cdot \boldsymbol{r}_a} , \qquad \frac{\partial\zeta}{\partial\xi^i} = -\frac{1}{c}\frac{P_{ij}r_a^j}{r_a - \boldsymbol{\beta}_a \cdot \boldsymbol{r}_a} , \qquad \frac{\partial\zeta}{\partial k^i} = -\frac{\sigma r_a^i}{r_a - \boldsymbol{\beta}_a \cdot \boldsymbol{r}_a} . \tag{7.51}$$

Furthermore, equation (7.50) yields a unique relationship between the time differentials of time σ on the worldline of the photon (light-ray cone) and the retarded time ζ along the gravity null-cone (7.13), which reads as follows

$$d\sigma = d\zeta \, \frac{r_a - \boldsymbol{\beta}_a \cdot \boldsymbol{r}_a}{r_a - \boldsymbol{k} \cdot \boldsymbol{r}_a} . \tag{7.52}$$

If parameter σ runs from $-\infty$ to $+\infty$, the new parameter ζ runs from $\zeta_{-\infty} = -\infty$ to $\zeta_{+\infty} = t^* + c^{-1}\boldsymbol{k} \cdot \boldsymbol{x}_a(\zeta_{+\infty})$ provided the motion of each body is restricted inside a bounded domain of space, like in the case of a binary system. In case the bodies move along straight lines with constant velocities, the parameter σ runs from $-\infty$ to $+\infty$, and the parameter ζ runs from $-\infty$ to $+\infty$ as well. In addition, one notices that when the numerical value of the parameter σ is equal to the time of observation τ, the numerical value of the parameter ζ equals to $s(\tau)$, which is found from the equation of the gravity null-cone cone (7.13) in which the point $\boldsymbol{x}$ denotes the spatial coordinates of observer.

After transforming time arguments the integrals (7.47), (7.48) take the form

$$B^{\alpha\beta}(s) = \sum_{a=1}^{N} B_a^{\alpha\beta}(s) , \tag{7.53}$$

$$B_a^{\alpha\beta}(s) = \frac{4G}{c^3}\int_{-\infty}^{s} \frac{T_a^{\alpha\beta}(\zeta) - \frac{1}{2}\eta^{\alpha\beta}T_{a\lambda}^{\lambda}(\zeta)}{r_a(\sigma,\zeta) - \boldsymbol{k} \cdot \boldsymbol{r}_a(\sigma,\zeta)} d\zeta , \tag{7.54}$$

$$D^{\alpha\beta}(s) = c\sum_{a=1}^{N}\int_{-\infty}^{\tau} B_a^{\alpha\beta}[\zeta(\sigma)]d\sigma , \tag{7.55}$$

where retarded time s in the upper limit of integration in equation (7.54) depends on the index of each body as it has already been mentioned in the previous text (i.e., $s = s_a$). Now one presents a remarkable, exact relationship

$$r_a(\sigma,\zeta) - \boldsymbol{k} \cdot \boldsymbol{r}_a(\sigma,\zeta) = ct^* + \boldsymbol{k} \cdot \boldsymbol{x}_a(\zeta) - c\zeta , \tag{7.56}$$

which can be proved by direct use of the gravity null-cone equation (7.13) and the expression (7.36) for the unperturbed trajectory of light ray. It is important to note that in the given relationship t^* is a constant time corresponding to the moment of the closest approach of the photon to the origin of coordinate system. The equation (7.56) shows that the integrand on the right side of equations (7.54) does not actually depend on time σ of the light ray at all, and the integration is performed only with respect to the retarded time variable ζ of the gravity null-cone.

Thus, as soon as the motion of the bodies $\boldsymbol{x}_a(t)$ is known as a function of time, the integral (7.54) can be calculated either analytically or numerically without solving the complicated gravity null-cone equation (7.13) to establish the relationship between the light-ray time and the gravity retarded time arguments. This procedure does not work for the integral (7.55) because the transformation to the new variable (7.52) does not eliminate explicit dependence on the argument of time τ from the integrand of this integral. Fortunately, as it is evident from the structure of equation (7.46) that one does not need to calculate this integral for finding the perturbed light-ray trajectory, but only its partial derivative with respect to ξ^i, which can be already calculated on the hypersurface of the null cone in the same way as the integral (7.54).

In order to find $\hat{\partial}_i D^{\alpha\beta}$ one uses equation (7.45) and note that the integrand of $B^{\alpha\beta}(s)$ does not depend on the variable ξ^i. This dependence manifests itself only indirectly through the upper limit of the integral (7.54), which depends on the retarded time $s = s(\tau, \boldsymbol{\xi})$ taken at the point of observation of the light ray. At this point the structure of the gravity null-cone equation has the following form

$$\tau + t^* = s + \frac{1}{c}|\boldsymbol{\xi} + c\boldsymbol{k}\tau - \boldsymbol{x}_a(s)| \,, \tag{7.57}$$

which assumes that the retarded time s is an implicit function of τ, $\boldsymbol{\xi}$, and t^*. For this reason, a straightforward differentiation of $B^{\alpha\beta}(s)$ with respect to the retarded time s and the implementation of formula (7.51) for the calculation of the derivative $\partial s/\partial\xi^i$ at the point of observation yields [6)]

$$\begin{aligned}\hat{\partial}_i B^{\alpha\beta}(s) &= \frac{\partial B^{\alpha\beta}(s)}{\partial s}\frac{\partial s}{\partial \xi^i} \\ &= -\frac{4G}{c^4}\sum_{a=1}^{N}\frac{T_a^{\alpha\beta}(s) - \frac{1}{2}\eta^{\alpha\beta}T^{\lambda}_{a\lambda}(s)}{r_a(s) - \boldsymbol{k}\cdot\boldsymbol{r}_a(s)}\;\frac{P^i{}_j\, r_a^j(s)}{r_a(s) - \boldsymbol{\beta}_a(s)\cdot\boldsymbol{r}_a(s)}\,.\end{aligned} \tag{7.58}$$

This result elucidates a remarkable fact that $\hat{\partial}_i B^{\alpha\beta}(s)$ is actually not an integral but the instantaneous function of time that can be calculated directly for arbitrary trajectory of the gravitating bodies. While calculating $\hat{\partial}_i D^{\alpha\beta}(s)$ one uses, first, formula (7.58) and, then, the replacement of variables (7.52). Proceeding in this way one arrives at the result

$$\begin{aligned}\hat{\partial}_i D^{\alpha\beta}(s) &= c\sum_{a=1}^{N}\int_{-\infty}^{\tau}\hat{\partial}_i B_a^{\alpha\beta}[\zeta(\sigma)]d\sigma \\ &= -\frac{4G}{c^4}\sum_{a=1}^{N}\int_{-\infty}^{s}\frac{T_a^{\alpha\beta}(\zeta) - \frac{1}{2}\eta^{\alpha\beta}T^{\lambda}_{a\lambda}(\zeta)}{[r_a(\sigma,\zeta) - \boldsymbol{k}\cdot\boldsymbol{r}_a(\sigma,\zeta)]^2}P^i{}_j\, r_a^j(\sigma,\zeta)d\zeta\,.\end{aligned} \tag{7.59}$$

It is worthwhile to emphasize that denominator of the integrand in right side of equation (7.59) can be represented as a function of the only variable ζ in accordance with

6) Calculation of the derivative $\hat{\partial}_i B^{\alpha\beta}(s_0)$ at the point of light emission is obtained from the formula (7.58) where all quantities involved are to be taken at the retarded time s_0.

equation (7.56), which is a consequence of the gravity null-cone equation (7.13). One also notices that

$$P^i{}_j \, r^j_a(\sigma,\zeta) = \xi^i - P^i{}_j x^j_a(\zeta) \,, \tag{7.60}$$

is function of the variable ζ only. Going back to the equation (7.56) one finds that the integrand of the integral (7.59) depends only on the retarded time argument ζ, so that it can be re-cast to the next form

$$\begin{aligned} \hat{\partial}_i D^{\alpha\beta}(s) &= -\frac{4G}{c^4}\sum_{a=1}^{N}\xi^i \int_{-\infty}^{s} \frac{T^{\alpha\beta}_a(\zeta) - \frac{1}{2}\eta^{\alpha\beta}T^{\lambda}_{a\lambda}(\zeta)}{[ct^* + \boldsymbol{k}\cdot\boldsymbol{x}_a(\zeta) - c\zeta]^2} d\zeta \\ &+ \frac{4G}{c^4}\sum_{a=1}^{N} P^i{}_j \int_{-\infty}^{s} \frac{T^{\alpha\beta}_a(\zeta) - \frac{1}{2}\eta^{\alpha\beta}T^{\lambda}_{a\lambda}(\zeta)}{[ct^* + \boldsymbol{k}\cdot\boldsymbol{x}_a(\zeta) - c\zeta]^2} \; x^j_a(\zeta)\, d\zeta \,, \end{aligned} \tag{7.61}$$

where the numerical value of the parameter s in the upper limit of the integral is calculated by solving the gravity null-cone equation (7.13). Hence, again, as it has been proven for $B^{\alpha\beta}(s)$, the integral (7.59) admits a direct calculation as soon as the motion of the gravitating bodies is prescribed.

7.4 Gravitational Perturbations of Photon's Trajectory

Gravitational perturbations of trajectory of the photon are found by straightforward integration of the equation of light-ray geodesics (7.46) with using definitions (7.47), (7.48) and the rules of integration (7.44), (7.45). After performing the calculation one fins,

$$\dot{x}^i(\tau) = ck^i + \dot{\Xi}^i(\tau)\,, \tag{7.62}$$

$$x^i(\tau) = x^i_N(\tau) + \Xi^i(\tau) - \Xi^i(\tau_0)\,, \tag{7.63}$$

where τ and τ_0 correspond, respectively, to the moment of observation and emission of the photon. Functions $\dot{\Xi}^i(\tau)$ and $\Xi^i(\tau)$ are relativistic perturbations given as follows

$$\dot{\Xi}^i(\tau) = \frac{1}{2}ck_\alpha k_\beta \hat{\partial}_i B^{\alpha\beta}(\tau) - k_\alpha h^{\alpha i}(\tau) - \frac{1}{2}k^i h^{00}(\tau) + \frac{1}{2}k^i k_p k_q h^{pq}(\tau)\,, \tag{7.64}$$

$$\Xi^i(\tau) = \frac{1}{2}k_\alpha k_\beta \hat{\partial}_i D^{\alpha\beta}(\tau) - k_\alpha B^{\alpha i}(\tau) - \frac{1}{2}k^i B^{00}(\tau) + \frac{1}{2}k^i k_p k_q B^{pq}(\tau)\,, \tag{7.65}$$

where functions $h^{\alpha\beta}(\tau)$, $B^{\alpha\beta}(\tau)$, $\hat{\partial}_i B^{\alpha\beta}(\tau)$, and $\hat{\partial}_i D^{\alpha\beta}(\tau)$ are defined by relationships (7.12), (7.54), (7.58), and (7.59) respectively.

Equations (7.64), (7.65) represent solution of the light-ray equation (7.46) with the initial-boundary condition (7.19). It means that the unit vector $\boldsymbol{k}$ defines a reference position of the source of light on the celestial sphere taken at the null past infinity. However, realistic sources of light are located at finite distances from observer, which assumes that the boundary value problem for equation (7.19) of the light-ray geodesics is to be solved. In this case the initial position, $\boldsymbol{x}_0 = \boldsymbol{x}(t_0)$, and the final position, $\boldsymbol{x} = \boldsymbol{x}(t)$, of a photon are given instead of its initial position $\boldsymbol{x}_0$ and the direction of light propagation $\boldsymbol{k}$ given at the past null infinity. All what is needed for the formulation of the boundary value problem is the relationship between the unit vector $\boldsymbol{k}$ and the unit vector

$$\boldsymbol{K} = -\frac{\boldsymbol{x} - \boldsymbol{x}_0}{|\boldsymbol{x} - \boldsymbol{x}_0|}\,, \tag{7.66}$$

which defines a geometric direction of the light propagation from observer to the source of light in flat spacetime. Formulas (7.63) and (7.65) yield

$$k^i = -K^i - \beth^i(\tau, \boldsymbol{\xi}) + \beth^i(\tau_0, \boldsymbol{\xi})\,, \tag{7.67}$$

where the relativistic corrections $\beth^i$ describing the difference between vectors $\boldsymbol{k}$ and K^i, are defined as follows

$$\beth^i(\tau, \boldsymbol{\xi}) = \frac{\frac{1}{2}k_\alpha k_\beta \hat{\partial}_i D^{\alpha\beta}(\tau) - k_\alpha P^i{}_j B^{\alpha j}(\tau)}{|\boldsymbol{x} - \boldsymbol{x}_0|}\,, \tag{7.68}$$

$$\beth^i(\tau_0, \boldsymbol{\xi}) = \frac{\frac{1}{2}k_\alpha k_\beta \hat{\partial}_i D^{\alpha\beta}(\tau_0) - k_\alpha P^i{}_j B^{\alpha j}(\tau_0)}{|\boldsymbol{x} - \boldsymbol{x}_0|}\,. \tag{7.69}$$

One emphasizes that the vectors $\beth^i(\tau, \xi) \equiv \beth$ and $\beth^i(\tau_0, \xi) \equiv \beth_0$ are orthogonal to the unit vector $\boldsymbol{k}$ and are taken at the points of observation and emission of the photon respectively. These vectors vanish if the distance $|\boldsymbol{x} - \boldsymbol{x}_0|$ between observer and the source of light goes to infinity. The relationships obtained in this section are used for the discussion of observable relativistic effects in the following sections, where the appropriate integrals will be calculated explicitly.

One emphasizes that the formalism under discussion admits to work with world-lines of arbitrary moving bodies without restricting them to straight lines as in the Lorentz-transformation technique of calculation of the gravitational light-ray perturbations proposed by Klioner [Klioner, 2003b]. More precisely, in the harmonic gauge (7.2) the equations of motion of the bodies result from the harmonic coordinate conditions (7.2). In the first post-Minkowskian approximation these conditions allow motion of the bodies only along straight lines with constant speeds. However, if in finding the metric tensor the non-linear terms in the Einstein equations are taken into account, the bodies may show accelerated motion without structurally changing the linearized form of the *Liénard-Wiechert* solution for the metric tensor, which is used for integration of equations of motion of a photon. Hence, the mathematical technique based on the Liénard-Wichert potentials supersedes the technique of the Lorentz transformations. Mathematical technique of the Synge's world function [Synge, 1964] proposed by French theory group [Teyssandier *et al.*, 2008] for calculating gravitational deflection and time delay of light in the program of space mission Gaia [Lindegren, 2009; Lindegren *et al.*, 1995; Lindegren and Perryman, 1996], is limited to the case of light propagating in the field of a static gravitational mass and requires further elaboration to include relativistic effects caused by motion of planets and the Sun [Malkin *et al.*, 2009].

The knowledge of trajectory of motion of photons in the gravitational field formed by a N-body system of arbitrary-moving massive bodies is necessary but not sufficient for the unambiguous physical interpretation of observational astrometric effects. In the most general case it also requires to know how observer and source of light move in the gravitational field of this system. This is because coordinate $\boldsymbol{x}$ of the observer and that $\boldsymbol{x}_0$ of the source of light enter the solution of the light-ray geodesics. Hence, as time passes and the observer receives photons from the source of light the observed direction $\boldsymbol{K}$ to the source of light and the magnitude and direction of relativistic perturbations change.

Let us stipulate that observer and the source of light are point-like test particles which move along timelike geodesic worldlines. Then, in the post-Minkowskian approximation, equation of motion of the particles, assuming no restriction on their velocities except for that $\dot{x} < c$, is exactly the same as equation (7.30). However,the most realistic situation is when observer and the source of light move with the speed much less than the speed of light: $\dot{x} \ll c$.

In the given coordinate system for velocities of the particles much smaller than the speed of light, equation (7.30) can be simplified to a more pragmatic form depending on the accuracy to which the velocity of the observer and/or the source of light must be taken into account. For example, in the Newtonian limit

$$\ddot{x}^i(t) = \frac{1}{2}c^2 h_{00,i} \ . \tag{7.70}$$

Regarding specific physical conditions either the post-Minkowski equation (7.30) or the Newtonian equation (7.70) should be integrated with respect to time to give the worldline of observer, $\boldsymbol{x}(t)$, and the source of light, $\boldsymbol{x}_0(t_0)$, as a function of time of observation, t, and of time of emission of light, t_0, respectively. This problem is not extensively treated in the present book as its solution has been developed with necessary accuracy by a number of previous authors. In particular, the post-Minkowskian approach for solving equations of motion of massive particles is thoroughly treated in [Bel *et al.*, 1981; Damour, 1983; Ibañez *et al.*, 1984], which also give a number of other, useful references. The post-Newtonian approach to the problem of motion is outlined in Chapters 4–6. In what follows, one assumes the motion of observer, $\boldsymbol{x}(t)$, and that of the source of light, $\boldsymbol{x}_0(t_0)$, are known with a required precision. In astrometry, these motions are connected through the solution of the light-ray geodesic equations (7.62)-(7.65).

It is also instructive to review propagation of photon through the gravitational field of N-body system on the Minkowski spacetime diagram as shown in Fig 7.2. A photon propagates from a source of light towards observer along light cone. Its motion is perturbed by the retarded gravitational field of the bodies expressed in terms of the indexLiénard-Wiechert potentials*Liénard-Wiechert* potentials. Also shown in Figure 7.2 are positions of the bodies (marked by the unfilled circles) taken on the space-like hypersurfaces (dashed lines) of fixed times t_i ($i = 0, 1, ..., 6$). As the photon approaches the system (events 0,1) it moves through a variable gravitational field of the two bodies. However, after crossing the system (events 5,6) the gravitational field at the photon's position gets "frozen" as it moves on the same null cone along which the gravitational field of the system propagates. Gravitational field is getting "frozen" during propagation of the photon inside the system (events 2,3,4). Spatial position $\boldsymbol{x}_a(s)$ of the gravitating bodies taken at any retarded instant of time $s = t - r_a/c$is very close to that when photon is crossing the system. The retarded, $\boldsymbol{x}_a(s)$, and instantaneous, $\boldsymbol{x}_a(t)$, spatial positions of the gravitating bodies are drastically different when the photon is very far outside the near zone of the system.

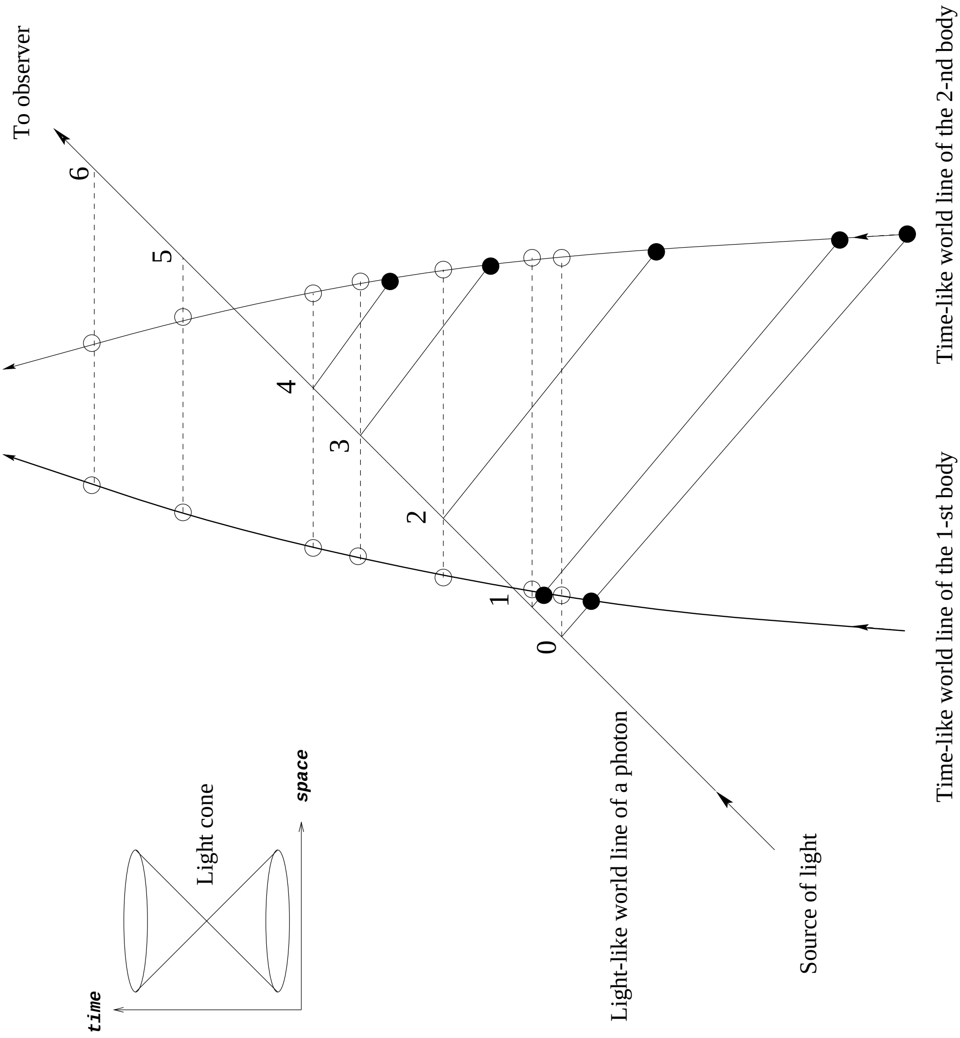

Figure 7.2 Minkowski diagram showing relationship between positions of a photon taken at times t_i (events $i = 0, 1, ..., 6$ on the photon's worldline) and positions of the light-ray-deflecting bodies (marked by the black circles) taken at the retarded times $s_i = t_i - r_a(s_i)/c$ corresponding to instants t_i $(i = 0, 1, ..., 6)$. For simplicity only gravitationally bounded two-body system (for example, a binary pulsar) is shown.

7.5 Observable Relativistic Effects

7.5.1 Gravitational time delay

Relativistic time delay in propagation of electromagnetic signals passing through the *static, spherically-symmetric* gravitational field of the Sun was predicted by Irwin Shapiro [Shapiro, 1964] who also made its first measurement by using radar observations of planets [Shapiro, 1966]. The basic idea involved estimating the unknown parameters of the planetary orbits from time-delay measurements made, when the target planet was far away from the direction to the Sun, and using the measurements made, when the path of the radar wave passes near the Sun to distinguish the effect of solar gravity on the delay. Detailed mathematical description of the Shapiro time delay can be found in any introductory textbook on relativity (see, for example, [Misner *et al.*, 1973; Weinberg, 1972]). Here, Shapiro's equation is generalized for the case of the propagation of light through *time-dependent* gravitational field formed by an ensemble of N arbitrary-moving massive bodies. The result, is valid not only when the light ray propagates outside the system of the bodies but also when light goes through the system.

The total time of propagation of an electromagnetic signal from the point $\boldsymbol{x}_0$ to the point $\boldsymbol{x}$ is derived from equations (7.63), (7.65). First, use equation (7.63) to express the difference $\boldsymbol{x}-\boldsymbol{x}_0$ between the point of emission and that of observation through the other terms of the equation. Then, square this difference using the properties of the Euclidean dot product. Finally, find the total coordinate time of propagation of light, $t-t_0$, extracting the square root from the product, and using the expansion with respect to the relativistic parameter $(Gm_a)/(c^2 r_a)$ which is assumed to be small. It results in

$$c(t-t_0) \quad = \quad |\boldsymbol{x}-\boldsymbol{x}_0| - \boldsymbol{k}\cdot\boldsymbol{\Xi}(\tau) + \boldsymbol{k}\cdot\boldsymbol{\Xi}(\tau_0)\,, \tag{7.71}$$

or

$$t-t_0 \quad = \quad \frac{1}{c}|\boldsymbol{x}-\boldsymbol{x}_0| + \Delta(t,t_0)\,, \tag{7.72}$$

where $|\boldsymbol{x}-\boldsymbol{x}_0|$ is the usual Euclidean distance between the points of emission, $\boldsymbol{x}_0$, and observation, $\boldsymbol{x}$, of the electromagnetic signal (photon), and $\Delta(t,t_0)$ is the generalization of the Shapiro time delay produced by the gravitational field of the moving bodies [Kopeikin and Schäfer, 1999]

$$\Delta(t,t_0) \quad = \quad \frac{1}{2c}k_\alpha k_\beta B^{\alpha\beta}(\tau) - \frac{1}{2c}k_\alpha k_\beta B^{\alpha\beta}(\tau_0) = \frac{2G}{c^3}\sum_{a=1}^{N} m_a B_a(s,s_0)\,. \tag{7.73}$$

Here, in the integral

$$B_a(s,s_0) \quad = \quad \int_{s_0}^{s}\frac{[1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a(\zeta)]^2}{\sqrt{1-\beta_a^2(\zeta)}}\,\frac{d\zeta}{t^*+c^{-1}\boldsymbol{k}\cdot\boldsymbol{x}_a(\zeta)-\zeta}\,, \tag{7.74}$$

the retarded time s is obtained by solving equation (7.13) for the time t of observation of the photon, and the retarded time s_0 is found by solving the same equation written down for the time t_0 of emission of the photon [7])

$$s_0 \quad = \quad t_0 - \frac{1}{c}|\boldsymbol{x}_0 - \boldsymbol{x}_a(s_0)| \,. \tag{7.75}$$

Relationships (7.72), (7.73) for the time delay have been derived with respect to the coordinate time t. Transformation from the coordinate time to the proper time $\mathcal{T}$ of observer moving with velocity $\boldsymbol{v}$, is made by integrating the infinitesimal increment of the proper time along the worldline $\boldsymbol{x}(t)$ of the observer as defined in equation (3.274) which takes the following form in the first post-Minkowskian approximation

$$\begin{aligned}\mathcal{T} = \int_{t_\mathrm{i}}^{t} & \Big\{1 - \beta^2(t) - h_{00}[t, \boldsymbol{x}(t)] - 2h_{0i}[t, \boldsymbol{x}(t)]\beta^i(t) \\ & - \quad h_{ij}[t, \boldsymbol{x}(t)]\beta^i(t)\beta^j(t)\Big\}^{1/2} dt \,,\end{aligned} \tag{7.76}$$

where $\boldsymbol{\beta} = \boldsymbol{v}/c$, t_i is the initial epoch of observation, and t is the current time of observation.

Calculation of the integral (7.74) is performed by means of using a new variable

$$y \quad = \quad t^* + \frac{1}{c}\boldsymbol{k}\cdot\boldsymbol{x}_a(\zeta) - \zeta \,, \tag{7.77}$$

$$dy \quad = \quad -[1 - \boldsymbol{k}\cdot\boldsymbol{\beta}_a(\zeta)]d\zeta \,, \tag{7.78}$$

so that the above integral (7.74) reads

$$B_a(s, s_0) \quad = \quad -\int_{s_0}^{s} \frac{1 - \boldsymbol{k}\cdot\boldsymbol{\beta}_a[\zeta(y)]}{\sqrt{1 - \beta_a^2[\zeta(y)]}}\, d\ln y; \,. \tag{7.79}$$

Integrating by parts and applying equation (7.56) to return from the time variable presentation of the gravity null cone to its coordinate presentation, results in

$$\begin{aligned}B_a(s, s_0) \quad = \quad & -\frac{1 - \boldsymbol{k}\cdot\boldsymbol{\beta}_a(s)}{\sqrt{1 - \beta_a^2(s)}} \ln[r_a(s) - \boldsymbol{k}\cdot\boldsymbol{r}_a(s)] \\ & + \frac{1 - \boldsymbol{k}\cdot\boldsymbol{\beta}_a(s_0)}{\sqrt{1 - \beta_a^2(s_0)}} \ln[r_a(s_0) - \boldsymbol{k}\cdot\boldsymbol{r}_a(s_0)] \\ & - \int_{s_0}^{s} \frac{\ln(r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a)}{(1 - \beta_a^2)^{3/2}} [\boldsymbol{k} - \boldsymbol{\beta}_a - \boldsymbol{\beta}_a \times (\boldsymbol{k}\times\boldsymbol{\beta}_a)]\cdot\dot{\boldsymbol{\beta}}_a \, d\zeta \,,\end{aligned} \tag{7.80}$$

where the notations are: $\boldsymbol{r}_a(s) = \boldsymbol{x} - \boldsymbol{x}_a(s)$, $\boldsymbol{r}_a(s_0) = \boldsymbol{x}_0 - \boldsymbol{x}_a(s_0)$, $r_a(s) = |\boldsymbol{r}_a(s)|$, $r_a(s_0) = |\boldsymbol{r}_a(s_0)|$, $\boldsymbol{x} = \boldsymbol{x}(t)$, $\boldsymbol{x}_0 = \boldsymbol{x}(t_0)$, $\boldsymbol{\beta}_a = \boldsymbol{\beta}_a(\zeta) = \boldsymbol{v}_a(\zeta)/c$, $\boldsymbol{v}_a(s) = d\boldsymbol{x}_a(s)/ds$, $\boldsymbol{v}_a(s_0) = d\boldsymbol{x}_a(s_0)/ds_0$, and the retarded times s and s_0 should be calculated from the gravity null-cone equations (7.13) and (7.75) respectively. The first and second terms

7) Again it would be better to denote the retarded time as s_{0a} emphasizing its dependence on the index a of the body under consideration. It was not used to avoid confusion of indices.

in the right side of this equation describe the generalized Lorentz-invariant form of the gravitational time delay for the case of the bodies moving with arbitrary velocities limited from above by the conditions (7.17), (7.18). The last term in the right side of equation (7.80) depends on the body's acceleration and may be comparable, in general case, to the first two terms in this equation. If motion of the bodies is bounded by gravitational interaction the acceleration-dependent terms are rather small and should be proportional to the relativistic corrections to the static Shapiro delay of the order of β_a^2. The acceleration-dependent term in equation (7.80) is identically zero if one can work in the approximation of the bodies moving along straight lines with constant velocities. Otherwise, one has to know the law of motion of the bodies for its precise calculation.

Neglecting all terms of order β_a^2 for the gravitational time delay one obtains its simplified expression

$$
\begin{aligned}
\Delta(t, t_0) \quad = \quad & -\frac{2G}{c^3} \sum_{a=1}^{N} m_a \Bigg\{ \ln\left(\frac{r_a - \boldsymbol{k} \cdot \boldsymbol{r}_a}{r_{0a} - \boldsymbol{k} \cdot \boldsymbol{r}_{0a}} \right) \\
& -(\boldsymbol{k} \cdot \boldsymbol{\beta}_a) \ln(r_a - \boldsymbol{k} \cdot \boldsymbol{r}_a) + (\boldsymbol{k} \cdot \boldsymbol{\beta}_{a0}) \ln(r_{0a} - \boldsymbol{k} \cdot \boldsymbol{r}_{0a}) \\
& + \int_{s_0}^{s} \ln\left[t^* + \boldsymbol{k} \cdot \boldsymbol{x}_a(\zeta) - \zeta\right] \left[\boldsymbol{k} \cdot \dot{\boldsymbol{\beta}}_a(\zeta)\right] d\zeta \Bigg\} ,
\end{aligned}
\tag{7.81}
$$

where $\boldsymbol{\beta}_a = \boldsymbol{\beta}_a(s)$, $\boldsymbol{\beta}_{a0} = \boldsymbol{\beta}_a(s_0)$, $\boldsymbol{r}_a = \boldsymbol{r}_a(s)$, $\boldsymbol{r}_{0a} = \boldsymbol{r}_a(s_0)$. The first term on the right side of the expression (7.81) looks identical with the static Shapiro time delay [Brumberg, 1991; Soffel, 1989; Will, 1993]. However, the reader should keep in mind that the argument of the logarithmic function depends on the worldlines of the moving bodies, which positions are not fixed in space but changes as time goes on. Moreover, the positions of the massive bodies are tightly connected to the position of observer, $\boldsymbol{x}$, at the time of observation t and that of the source of light, $\boldsymbol{x}_0$, at the time of emission t_0 of the observed light by equations (7.13), (7.75) of the gravity null cone, which are null characteristics of the linearized Einstein equations. This theoretical prediction of general relativity can be tested in precise astrometric experiments dealing with the gravitational time delay of light passing through the field of a moving, massive planet (Jupiter, Saturn) [Fomalont and Kopeikin, 2008, 2003; Kopeikin, 2001; Malkin *et al.*, 2009]. This kind of astrometric experiments allow us to measure whether the fundamental speed in general relativity has the same numerical value as the speed of light. Because the fundamental speed in general relativity defines the speed of propagation of gravity, the astrometric experiment measuring the gravitational time delay of light by a moving planet became known as "the speed of gravity" experiment [Kopeikin and Fomalont, 2002], which is discussed later in this section in more detail. Velocity-dependent corrections to the static Shapiro delay, which appear in the second line of equation (7.81) have been also obtained in the post-Newtonian approximate analysis with a different mathematical techniques under the assumption that gravitating bodies move uniformly along straight lines [Bertotti *et al.*, 2008; Klioner and Kopeikin, 1992; Wex, 1995; Wex and Kopeikin, 1999]. It should be emphasized once again that this assumption works well enough only if the light travel time does not exceed the characteristic Keplerian period of the gravitating system. Previous mathematical techniques were insufficient to prove that the assumption of uniform motion of bodies can be applied, e.g., for treatment of the Shapiro time delay in binary pulsars. This problem is discussed more deeply in the next sections of this book.

Expressions (7.80), (7.81) for gravitational time delay of light vastly extends previously known results, for they are applicable to the case of bodies moving along arbitrary worldlines whereas the calculations of all previous authors were severely restricted by the assumption that either the gravitating bodies are fixed in space or move uniformly with constant velocities. In addition, there was no reasonable theoretical understanding of how light interacts with gravitational field of a moving body. The rigorous theoretical derivation of formulas (7.80) and (7.81) has made a significant progress in clarifying this question and proved for the first time that in calculating the gravitational time delay of an electromagnetic wave the positions of the gravitating masses must be taken at the retarded times defined by the gravity null cone equations (7.13), (7.75) which correspond to the instants of emission and observation of the electromagnetic signal.

In the case of a uniformly moving bodies their velocity is constant and equation for the gravitational time delay can be represented in an explicitly Lorentz-invariant form

$$\Delta(t,t_0)=\frac{2G}{c^3}\sum_{a=1}^{N}m_a\left(k_\alpha u_a^\alpha\right)\ln\left[\frac{k_\alpha r_a^\alpha(s)}{k_\alpha r_a^\alpha(s_0)}\right], \tag{7.82}$$

where the four-velocity $u_a^\alpha = \gamma_a(1,\boldsymbol{\beta}_a)$, $\gamma_a = (1-\beta_a^2)^{-1/2}$, $k^\alpha = (1,\boldsymbol{k})$, $r_a^\alpha(s) = (r_a(s),\boldsymbol{r}_a(s))$, $r_a^\alpha(s_0) = (r_a(s_0),\boldsymbol{r}_a(s_0))$. One notices that the four-vector k^α is directed along the propagation of light ray, which is a null characteristic of the Maxwell equations, while the four-vectors $\boldsymbol{r}_a(s)$ and $\boldsymbol{r}_a(s_0)$ are directed along the null characteristics of the linearized Einstein equations. The explicit dependence of the argument of the logarithmic function on the dot product $k_\alpha r_a^\alpha$ of two null vectors k^α and r_a^α makes it evident that the gravitational time delay is sensitive to the null-cone effect because of the finite speed of propagation of gravity [Kopeikin, 2001; Kopeikin and Fomalont, 2006]. Equation (7.82) can be recast to another covariant form, which is very similar to that used in practical calculation of the time delay in Orbit Determination Program of NASA JPL [Moyer, 2003].

One starts from the following exact relationships

$$r_a-\boldsymbol{k}\cdot\boldsymbol{r}_a = \frac{|\boldsymbol{r}_{a0}-\boldsymbol{x}_a+\boldsymbol{x}_{a0}|^2-(r-r_a)^2}{2r}, \tag{7.83}$$

$$r_{a0}-\boldsymbol{k}\cdot\boldsymbol{r}_{a0} = -\frac{|\boldsymbol{r}_a+\boldsymbol{x}_a-\boldsymbol{x}_{a0}|^2-(r+r_{a0})^2}{2r}, \tag{7.84}$$

where $r=|\boldsymbol{r}|$, $\boldsymbol{r}=\boldsymbol{x}-\boldsymbol{x}_0$, so that

$$r^\alpha = rk^\alpha = (r,\boldsymbol{r}), \tag{7.85}$$

is a null vector in the flat spacetime connecting coordinates of the point of emission and reception of the electromagnetic wave: $\eta_{\alpha\beta}r^\alpha r^\beta = 0$. Because the gravitating body moves uniformly with constant speed $\boldsymbol{v}_a$, its retarded coordinate $\boldsymbol{x}(s)$ is not constant and can be expanded as follows

$$\boldsymbol{x}_a(s) = \boldsymbol{x}_{a0} + \boldsymbol{v}_a\,(s - s_0)\ , \tag{7.86}$$

where the time interval $s - s_0$ can be expressed in terms of the null-cone distances by making use of the retarded time equations (7.13), (7.75), and the unperturbed part of equation (7.71). One has,

$$s - s_0 \equiv (s - t) + (t - t_0) + (t_0 - s_0) = \frac{1}{c}\,(r + r_{a0} - r_a)\ . \tag{7.87}$$

Plugging equation (7.87) to (7.86), and substituting it into equations (7.83), (7.84) allow us to transform the ranging time delay logarithm to the following form

$$\ln\left[\frac{r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a}{r_{a0} - \boldsymbol{k}\cdot\boldsymbol{r}_{a0}}\right] = -\ln\left[\frac{r_a + r_{a0} + r - 2(\boldsymbol{r}_a\cdot\boldsymbol{\beta}_a) - \beta_a^2\,(r + r_{a0} - r_a)}{r_a + r_{a0} - r - 2(\boldsymbol{r}_{a0}\cdot\boldsymbol{\beta}_a) + \beta_a^2\,(r + r_{a0} - r_a)}\right]\ . \tag{7.88}$$

Let us now introduce definitions of the Lorentz-invariant distances

$$\rho_R = -u_{a\alpha} r_a^\alpha = \frac{r_a - \boldsymbol{\beta}_a\cdot\boldsymbol{r}_a}{\sqrt{1-\beta_a^2}}\ , \tag{7.89}$$

$$\rho_{0R} = -u_{a\alpha} r_{a0}^\alpha = \frac{r_{a0} - \boldsymbol{\beta}_a\cdot\boldsymbol{r}_{a0}}{\sqrt{1-\beta_a^2}}\ . \tag{7.90}$$

Tedious but straightforward calculations reveal that

$$r_a + r_{a0} + r - 2(\boldsymbol{r}_a\cdot\boldsymbol{\beta}_a) - \beta_a^2\,(r + r_{a0} - r_a) = \sqrt{1-\beta_a^2}\,(\rho_R + \rho_{0R} - r k_\alpha u_a^\alpha)\ , \tag{7.91}$$

$$r_a + r_{a0} - r - 2(\boldsymbol{r}_{a0}\cdot\boldsymbol{\beta}_a) + \beta_a^2\,(r + r_{a0} - r_a) = \sqrt{1-\beta_a^2}\,(\rho_R + \rho_{0R} + r k_\alpha u_a^\alpha)\ . \tag{7.92}$$

These equations taken along with equation (7.85) allows us to reduce the time delay logarithm in equation (7.88) to another Lorentz-invariant form

$$\ln\left(\frac{r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a}{r_{a0} - \boldsymbol{k}\cdot\boldsymbol{r}_{a0}}\right) = -\ln\left(\frac{\rho_R + \rho_{0R} - \rho_{12}}{\rho_R + \rho_{0R} + \rho_{12}}\right)\ , \tag{7.93}$$

where the ranging distance $\rho_{12} = r k_\alpha u_a^\alpha = u_{a\alpha} r^\alpha$ is invariant with respect to the Lorentz transformation. It represents transvection of the null vector r^α defined in equation (7.85) with four-velocity u_a^α of the gravitating body. The null vector r^α determines propagation of the electromagnetic signal. Distances ρ_{0R}, ρ_R are defined in equations (7.89), (7.90), and they also represent transvection of the null vectors r_{a0}^α, r_a^α with four-velocity u_a^α of the gravitating body. However, contrary to vector r^α, vectors r_{a0}^α, r_a^α describe the null characteristics of the gravitational field.

Accounting for equation (7.93) the Lorentz-invariant expression for the time delay assumes the following form

$$\Delta(t, t_0) = \frac{2G}{c^3}\sum_{a=1}^{N} m_a \frac{1 - \boldsymbol{k}\cdot\boldsymbol{\beta}_a}{\sqrt{1-\beta_a^2}} \ln\left(\frac{\rho_R + \rho_{0R} - \rho_{12}}{\rho_R + \rho_{0R} + \rho_{12}}\right)\ . \tag{7.94}$$

This equation is apparently Lorentz-invariant, valid for any value of the velocity $\boldsymbol{v}_a$ of the light-ray deflecting body. This equation has been derived in [Kopeikin, 2009] and represent Lorentz-invariant generalization of the gravitational time delay equation used in NASA Orbit Determination Program [Moyer, 2003].

7.5.2 Gravitational bending and deflection angle of light

A photon arriving at time t from a source of light to observer, located at the point of observation $\boldsymbol{x}$, has a coordinate direction defined in spacetime by a null four-vector $p^\alpha = (1, p^i)$, where

$$p^i = -\frac{1}{c}\dot{x}^i \quad = \quad -k^i - \frac{1}{c}\dot{\Xi}^i(\tau, \boldsymbol{\xi}) \,, \tag{7.95}$$

and the minus sign makes the space vector p^i directed from the observer to the source of light. However, the coordinate direction p^i to the source of light is not a directly observable quantity. A real observable direction towards the source of light is defined with respect to the local inertial frame of the observer by another null vector that is denoted $s^\alpha = (1, s^i)$. In this frame $s^i = -c^{-1}d\mathcal{X}^i/d\mathcal{T}$, where $\mathcal{T}$ is the observer's proper time and $\mathcal{X}^i$ are spatial coordinates of the local inertial frame. One will assume for simplicity that the observer is at rest with respect to the global harmonic coordinate system (t, x^i). If observer is moving with respect to the global coordinates with velocity v^i the additional Lorentz transformation must be applied. It is described by the matrix $L^\alpha{}_\beta$ with components given earlier in equation (2.82)

$$L^0{}_0 = \gamma \equiv \frac{1}{\sqrt{1-\beta^2}} \,, \qquad L^0{}_i = L^i{}_0 = -\beta\gamma n^i \,, \qquad L^i{}_j = \delta^{ij} + (\gamma - 1)n^i n^j \,, \tag{7.96}$$

where $\beta = v/c$, and $n^i = v^i/v$ is the unit vector in the direction of motion of the observer.

In case, when observer is at rest, the infinitesimal transformation from the global coordinates, (t, x^i), to the local inertial ones, $(\mathcal{T}, \mathcal{X}^i)$, is given by

$$c d\mathcal{T} = c\Lambda^0{}_0\, dt + \Lambda^0{}_j\, dx^j \quad , \qquad d\mathcal{X}^i = c\Lambda^i{}_0\, dt + \Lambda^i{}_j\, dx^j \,, \tag{7.97}$$

where the matrix of transformation $\Lambda^\alpha{}_\beta$ depends on the spacetime coordinates of the point of observation and is defined by the requirement of orthonormalization

$$g_{\alpha\beta} \quad = \quad \eta_{\mu\nu}\Lambda^\mu{}_\alpha\Lambda^\nu_\beta \,. \tag{7.98}$$

The matrix $\Lambda^\alpha{}_\beta$ represents components of the four basis vectors (tetrad) of the local inertial frame of the observer, expressed in terms of the global coordinates. In particular, the orthonormalization condition (7.98) pre-assumes that spatial angles and lengths at the point of observation are measured with the help of the Euclidean metric δ_{ij}. For this reason, as the four-vector s^α is null, one concludes that the Euclidean length $|\boldsymbol{s}|$ of vector s^i is equal to 1. Indeed, one has

$$\eta_{\alpha\beta}s^\alpha s^\beta \quad = \quad -1 + \boldsymbol{s}^2 = 0 \,. \tag{7.99}$$

Hence, $|\boldsymbol{s}| = 1$, and the unit vector $\boldsymbol{s}$ defines the astrometric position of the source of light on the unit celestial sphere attached to the point of observation. This is a, so-called, topocentric celestial sphere in the local coordinates $\mathcal{X}^\alpha = (\mathcal{T}, X$. The positions of celestial objects will differ from one observer to another, and may vary periodically at one site due to the rotation and orbital revolution of the Earth or if the observer is doing observations from an orbiting satellite.

In the linearized post-Minkowskian approximation the matrix of the tetrad transformation is as follows [Kopeikin *et al.*, 1999]

$$\begin{aligned} \Lambda^0{}_0 &= 1 - \frac{1}{2} h_{00}(t, \boldsymbol{x}) \,, \qquad (7.100)\\ \Lambda^0{}_i &= -h_{0i}(t, \boldsymbol{x}) \,, \\ \Lambda^i{}_0 &= 0 \,, \\ \Lambda^i{}_j &= \delta_{ij} + \frac{1}{2} h_{ij}(t, \boldsymbol{x}) \,. \end{aligned}$$

Using transformation (7.97) one obtains a relationship between the observable unit vector s^i and the coordinate direction p^i

$$s^i = \frac{\Lambda^i{}_j \, p^j - \Lambda^i{}_0}{\Lambda^0{}_0 - \Lambda^0{}_j \, p^j} \,. \qquad (7.101)$$

In the linearized approximation it takes the form

$$s^i = \left(1 + \frac{1}{2} h_{00} - h_{0j} p^j\right) p^i + \frac{1}{2} h_{ij} p^j \,. \qquad (7.102)$$

Remembering that $|\boldsymbol{s}| = 1$, one obtains for the Euclidean norm of vector p^i

$$|\boldsymbol{p}| = 1 - \frac{1}{2} h_{00} + h_{0j} p^j - \frac{1}{2} h_{ij} p^i p^j \,, \qquad (7.103)$$

which brings equation (7.102) to the form [8)]

$$s^i = m^i + \frac{1}{2} P^{ij} m^q h_{jq}(t, \boldsymbol{x}) \,, \qquad (7.104)$$

with the Euclidean unit vector $m^i = p^i / |\boldsymbol{p}|$. One notices that a set of the unit vectors m^i also defines a topocentric celestial sphere in the global coordinates $x^\alpha = (t, \boldsymbol{x})$. Equation (7.104) makes it evident that gravitational field at the point of observation makes the components of a topocentric vector directed towards one and the same source of light different in the local and global coordinates. This is a specific feature of general relativistic astrometry, which is not present in classic astrometry where gravitational field does not invoke any difference between the coordinates because space and time are absolute.

Let now denote by α^i a dimensionless vector describing the angle of a total deflection of the light ray caused by the gravitational field from the point of emission of

8) Note that one is allowed to use the substitution $\delta_{ij} - p_i p_j = \delta_{ij} - k_i k_j = P_{ij}$ in the relativistic terms of any formula of the present chapter.

light to the point of observation, measured at the point of observation and calculated with respect to vector k^i of the light ray given by the boundary condition imposed on the light ray at the past null infinity. It is defined according to relationship [Kopeikin *et al.*, 1999]

$$c\alpha^i(\tau,\boldsymbol{\xi}) \quad = \quad -P^i{}_j\,\dot{\Xi}^j(\tau,\boldsymbol{\xi}) = k^i\left[\boldsymbol{k}\cdot\dot{\boldsymbol{\Xi}}(\tau,\boldsymbol{\xi})\right] - \dot{\Xi}^i(\tau,\boldsymbol{\xi})\ . \tag{7.105}$$

Now one notices that definition (7.95) yields for the norm of the vector p^i

$$|\boldsymbol{p}| = 1 + \frac{1}{c}\boldsymbol{k}\cdot\dot{\boldsymbol{\Xi}}\ . \tag{7.106}$$

Comparing it with equation (7.103) allows us to calculate

$$\frac{1}{c}\boldsymbol{k}\cdot\dot{\boldsymbol{\Xi}} = -\frac{1}{2}h_{00} + h_{0j}k^j - \frac{1}{2}h_{ij}k^ik^j\ . \tag{7.107}$$

As a consequence of definitions (7.95) and equations (7.105), (7.106) one can establish a relationship between the unit vectors k^i and m^i that

$$m^i \quad = \quad -k^i + \alpha^i(\tau,\boldsymbol{\xi})\ . \tag{7.108}$$

The boundary value of the light vector k^i relates to the geometric direction K^i of the light ray by equation (7.67). Hence, taking into account this equation and expressions (7.101), (7.103), (7.105), one obtains the observed astrometric direction to the source of light

$$s^i(\tau,\boldsymbol{\xi}) \quad = \quad K^i + \alpha^i(\tau,\boldsymbol{\xi}) + \beth^i(\tau,\boldsymbol{\xi}) - \beth^i(\tau_0,\boldsymbol{\xi}) + \varkappa^i(\tau,\boldsymbol{\xi})\ , \tag{7.109}$$

where all terms following vector K^i have various relativistic origin due to the presence of the gravitational field, relativistic corrections $\beth^i$ are defined by the equation (7.68) and the local frame relativistic correction

$$\varkappa^i(\tau,\boldsymbol{\xi}) \quad = \quad -\frac{1}{2}P^{ij}k^q h_{jq}(t,\boldsymbol{x}) \tag{7.110}$$

describes the light deflection caused by the deformation of space at the point of observations due to the presence of the gravitational field.

If two sources of light are observed along the directions s^i_1 and s^i_2, correspondingly, the measured angle ψ between them is defined in the local inertial frame as follows

$$\cos\psi \quad = \quad \boldsymbol{s}_1\cdot\boldsymbol{s}_2\ , \tag{7.111}$$

where the dot denotes the usual Euclidean scalar product. It is worth emphasizing that the observed direction to the source of light (7.109) includes the relativistic deflection of the light ray which depends not only on quantities taken at the point of observation but also on those $\beth^i(\tau_0,\boldsymbol{\xi})$ taken at the point of emission of light. Usually this term is rather small and can be neglected. However, it becomes important in the problem of propagation of light in the field of gravitational waves [Kopeikin and Korobkov, 2005; Kopeikin *et al.*, 2006, 1999] or for a proper treatment of high-precision astrometric observations of objects being within the boundary of the solar system (for example, tracking motion of interplanetary spacecrafts with VLBI [Ichikawa *et al.*, 2004; Lanyi *et al.*, 2007]).

Without going into specific technical details of the observational procedure one gives an explicit expression for the angle $\alpha^i(\tau)$ that follows from its definition (7.105) and expression (7.64) for the relativistic perturbation of the coordinate velocity of the light ray

$$\alpha^i(\tau) = -\frac{1}{2}k_\alpha k_\beta \hat{\partial}_i B^{\alpha\beta}(\tau) + k_\alpha\, P^i{}_j\, h^{\alpha j}(\tau)\,. \tag{7.112}$$

Relationships (7.12), (7.58) taken along with the definition of the tensor of energy-momentum (7.6)-(7.8) allow us to recast the previous expression into the following explicit form

$$\begin{aligned}\alpha^i(\tau) &= \frac{2G}{c^2}\sum_{a=1}^{N}\frac{m_a}{\sqrt{1-\beta_a^2}}\frac{(1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a)^2}{r_a-\boldsymbol{k}\cdot\boldsymbol{r}_a}\frac{P^i{}_j\, r_a^j}{r_a-\boldsymbol{\beta}_a\cdot\boldsymbol{r}_a} \\ &- \frac{4G}{c^2}\sum_{a=1}^{N}\frac{m_a}{\sqrt{1-\beta_a^2}}\frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a}{r_a-\boldsymbol{\beta}_a\cdot\boldsymbol{r}_a}P^i{}_j\,\beta_a^j\,,\end{aligned} \tag{7.113}$$

where coordinates and velocity of the massive a-th body have to be taken at the retarded time s, which relates to the time of observation t by the gravity null-cone equation (7.13). Equation (7.113) gives Lorentz-invariant expression for the gravitational deflection angle α^i, which consists of two terms. The first term in the right side of equation (7.113) reduces, in case of $\boldsymbol{\beta}_a = 0$, to the textbook's expression for the deflection angle of light in gravitational field of a static, spherically-symmetric mass predicted by Einstein and, first, measured by Eddington in case of starlight grazing limb of the Sun [Weinberg, 1972, section 8.5]. Relativistic corrections of the order of β_a to the value of the static-case deflection angle are small (less than one microarcsecond [Brumberg *et al.*, 1990]) for the solar system bodies and can be currently neglected. However, relativistic corrections associated with the retardation in the positions of the light-ray deflecting bodies due to the finite speed of gravity are important, can be measured [Fomalont and Kopeikin, 2003], and should be taken seriously in relativity-testing astrometric experiments [Kopeikin and Fomalont, 2007; Kopeikin and Makarov, 2007].

Neglecting all terms of the order of β_a one obtains a simplified form of the previous expression

$$\alpha^i(\tau) = \frac{2G}{c^2}\sum_{a=1}^{N}\frac{m_a}{r_a}\frac{P^i{}_j\, r_a^j}{(r_a-\boldsymbol{k}\cdot\boldsymbol{r}_a)}\,, \tag{7.114}$$

which may be compared to the analogous expression for the deflection angle obtained previously by many other authors in the framework of the post-Newtonian approximation (see, for example, [Brumberg, 1991], and references therein). One notes that all previous authors fixed the moment of time, at which coordinates $\boldsymbol{x}_a$ of the gravitating bodies were to be calculated rather arbitrary, without having a rigorous justification for their choice. Our approach based on the retarded *Liénard-Wiechert* potentials, gives a unique answer to this question and makes it obvious that the coordinates $\boldsymbol{x}_a$ of the bodies must be taken at the moment of retarded time s relating to the time of observation t by the gravity null-cone equation (7.13).

The next step in finding the explicit expression for the observed coordinate direction s^i is the computation of the quantity $\beth^i(\tau)$ given in (7.68). One has from formulas (7.53), (7.59) the following result for the numerator of $\beth^i(\tau)$

$$\frac{1}{2}k_\alpha k_\beta \hat{\partial}_i D^{\alpha\beta}(\tau) - k_\alpha P^i{}_j B^{\alpha j}(\tau) = -\frac{2G}{c^2}\sum_{a=1}^{N} m_a \left[\xi^i C_a(s) - P^i{}_j D_a^j(s)\right] + \frac{4G}{c^2}\sum_{a=1}^{N} m_a\, P^i{}_j E_a^j(s)\,, \tag{7.115}$$

where the integrals $C_a(s)$, $D_a^j(s)$ and $E_a^j(s)$ read as follows

$$C_a(s) = c\int_{-\infty}^{s}\left[\frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a(\zeta)}{ct^* + \boldsymbol{k}\cdot\boldsymbol{x}_a(\zeta) - c\zeta}\right]^2 \frac{d\zeta}{\sqrt{1-\beta_a^2(\zeta)}}\,, \tag{7.116}$$

$$D_a^j(s) = c\int_{-\infty}^{s}\left[\frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a(\zeta)}{ct^* + \boldsymbol{k}\cdot\boldsymbol{x}_a(\zeta) - c\zeta}\right]^2 \frac{x_a^j(\zeta)}{\sqrt{1-\beta_a^2(\zeta)}}\,d\zeta\,, \tag{7.117}$$

$$E_a^j(s) = \int_{-\infty}^{s}\frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a(\zeta)}{ct^* + \boldsymbol{k}\cdot\boldsymbol{x}_a(\zeta) - c\zeta}\,\frac{\beta_a^j(\zeta)}{\sqrt{1-\beta_a^2(\zeta)}}\,d\zeta\,. \tag{7.118}$$

Making use of the new variable y introduced in (7.77) and integrating by parts yields

$$C_a(s) = \frac{1}{\sqrt{1-\beta_a^2}}\,\frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a}{r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a} + \int_{-\infty}^{s}\frac{[\boldsymbol{k}-\boldsymbol{\beta}_a-\boldsymbol{\beta}_a\times(\boldsymbol{k}\times\boldsymbol{\beta}_a)]\cdot\dot{\boldsymbol{\beta}}_a}{r_a-\boldsymbol{k}\cdot\boldsymbol{r}_a}\,\frac{d\zeta}{(1-\beta_a^2)^{3/2}}\,, \tag{7.119}$$

$$D_a^j(s) = \frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a}{\sqrt{1-\beta_a^2}}\,\frac{x_a^j}{r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a} + \int_{-\infty}^{s}\frac{[\boldsymbol{k}-\boldsymbol{\beta}_a-\boldsymbol{\beta}_a\times(\boldsymbol{k}\times\boldsymbol{\beta}_a)]\cdot\dot{\boldsymbol{\beta}}_a}{r_a-\boldsymbol{k}\cdot\boldsymbol{r}_a}\,\frac{x_a^j\,d\zeta}{(1-\beta_a^2)^{3/2}} - E_a^j(s)\,, \tag{7.120}$$

$$E_a^j(s) = -\frac{\beta_a^j}{\sqrt{1-\beta_a^2}}\ln(r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a) + \int_{-\infty}^{s}\ln(r_a-\boldsymbol{k}\cdot\boldsymbol{r}_a)\,\Pi^j_k\,\dot{\beta}_a^k\,\frac{d\zeta}{(1-\beta_a^2)^{1/2}}\,, \tag{7.121}$$

where $\Pi^j_k(\zeta) = \delta^j_k + u^j(\zeta)u_k(\zeta)$ is the spatial part of the operator of projection onto the plane being perpendicular to the worldline of the a-th body, and the bodies' coordinates and velocities in all terms, being outside the signs of the integrals, are taken at the moment of the retarded time $s = t - r_a/c$. Equations (7.119)-(7.121) will be used in section 7.6.2 for the discussion of the gravitational lens equation with taking into account velocity-dependent relativistic corrections of the body deflecting light rays.

Finally, the quantity $\varkappa^i(\tau)$ can be explicitly given by the following expression

$$\varkappa^i(\tau) = -\frac{2G}{c^2}\sum_{a=1}^{N}\frac{m_a}{\sqrt{1-\beta_a^2}}\,\frac{(\boldsymbol{k}\cdot\boldsymbol{\beta}_a)}{r_a-\boldsymbol{\beta}_a\cdot\boldsymbol{r}_a}\,(P^i{}_j\,\beta_a^j)\,, \tag{7.122}$$

where coordinates and velocities of the bodies must be taken at the retarded time s according to equation (7.13). Notice that the angle $\varkappa^i$ is a very small quantity being proportional to the product $(Gm_a/c^2 r_a)\beta_a^2$ and can be neglected in all practical cases of the astrometric observations inside the solar system.

7.5.3
Gravitational shift of electromagnetic-wave frequency

Exact calculation of gravitational shift of frequency between the emitted and observed electromagnetic wave is crucial for adequate interpretation of high-precision measurements of radial velocities [Lindegren and Dravins, 2003], the inhomogeneity induced anisotropy of cosmic microwave background (CMB) radiation also known as the Sachs-Wolfe effect [Linder, 1997; White and Hu, 1997], Doppler tracking gravitational experiments in the solar system [Cowsik, 2000; Estabrook and Wahlquist, 1975; Iess *et al.*, 1999; Mashhoon and Grishchuk, 1980] and other spectroscopic astronomical investigations. In the last decade, for instance, radial velocity measuring technique has reached unprecedented precision 1 m/s [Cochran, 1996; Lovis *et al.*, 2005; Valenti *et al.*, 1995]. In the future, much better precision can be hoped for, on instruments designed for the “extremely large telescopes" [Pasquini *et al.*, 2006] when measurement of the post-Newtonian relativistic effects in optical binary and/or multiple star systems will be possible [Kopeikin and Ozernoy, 1999; Zucker and Alexander, 2007].

Let a source of light move with respect to the coordinate system $x^\alpha = (t, x^i)$ with velocity $\boldsymbol{v}_0 = \boldsymbol{v}_0(t_0) = d\boldsymbol{x}_0(t_0)/dt_0$, which is a function of the time of emission t_0, and emit electromagnetic wave (photon) with frequency $\nu_0 = 1/(\delta\mathcal{T}_0)$, where $\delta\mathcal{T}_0$ is the period of the wave expressed in the proper time $\mathcal{T}_0$ of the source of light. One denotes by $\nu = 1/(\delta\mathcal{T})$ the observed frequency of the electromagnetic wave measured at the point of observation by an observer moving with velocity $\boldsymbol{v} = \boldsymbol{v}(t) = d\boldsymbol{x}/dt$ with respect to the coordinate system $x^\alpha = (t, x^i)$. One stipulates that the time periods $\delta\mathcal{T}_0$ and $\delta\mathcal{T}$ are so small (high-frequency approximation) that one can treat them as infinitesimal differentials. Therefore, the observed gravitational shift of frequency $1 + z = \nu/\nu_0$ can be defined through the consecutive differentiation of the proper time of the source of light, $\mathcal{T}_0$, with respect to the proper time of the observer, $\mathcal{T}$, [Brumberg, 1972; Synge, 1964]. In other words,

$$1 + z \quad = \quad \frac{d\mathcal{T}_0}{d\mathcal{T}} = \frac{d\mathcal{T}_0}{dt_0}\frac{dt_0}{dt}\frac{dt}{d\mathcal{T}} \,, \tag{7.123}$$

where the derivative dt_0/dt is calculated by making use of the time delay equation (7.72), the derivative

$$\frac{d\mathcal{T}_0}{dt_0} = \left[1 - \beta_0^2(t_0) - h_{00}(t_0, \boldsymbol{x}_0) - 2h_{0i}(t_0, \boldsymbol{x}_0)\,\beta_0^i(t_0) - h_{ij}(t_0, \boldsymbol{x}_0)\,\beta_0^i(t_0)\,\beta_0^j(t_0)\right]^{1/2} \,, \tag{7.124}$$

is taken at the point of emission of light, and the derivative

$$\frac{dt}{d\mathcal{T}} = \left[1 - \beta^2(t) - h_{00}(t,\boldsymbol{x}) - 2h_{0i}(t,\boldsymbol{x})\,\beta^i(t) - h_{ij}(t,\boldsymbol{x})\,\beta^i(t)\,\beta^j(t)\right]^{-1/2} , \tag{7.125}$$

is calculated at the point of observation, $\boldsymbol{\nu} = \boldsymbol{v}/c$ and $\boldsymbol{\nu}_0 = \boldsymbol{v}_0/c$.

Synge calls relationship (7.123) the *Doppler effect in terms of frequency* [Synge, 1964, page 122]. It is fully consistent with a definition of the *Doppler shift in terms of energy* [Synge, 1964, page 231], when one compares the energy of photon at the point of emission against that at the point of observation of light. The *Doppler shift in terms of energy* is given by equation

$$1 + z \quad = \quad \frac{\nu}{\nu_0} = \frac{u^\alpha \mathcal{K}_\alpha}{u_0^\alpha \mathcal{K}_{0\,\alpha}} , \tag{7.126}$$

where u_0^α, u^α are four-velocities of the source of light and observer, and $\mathcal{K}_{0\,\alpha}$, $\mathcal{K}_\alpha$ are four-momenta of the photon at the points of emission and observation respectively. It is quite easy to see that equation (7.126) is equivalent to its counterpart (7.123). Indeed, taking into account that $u^\alpha = dx^\alpha/d\mathcal{T}$ and $\mathcal{K}_\alpha = \partial\varphi/\partial x^\alpha$, where φ is the phase of the electromagnetic wave [Misner *et al.*, 1973, equation (22.26a)], one obtains $u^\alpha \mathcal{K}_\alpha = d\varphi/d\mathcal{T}$. Thus,

$$1 + z \quad = \quad \frac{d\varphi}{d\varphi_0}\frac{d\mathcal{T}_0}{d\mathcal{T}} . \tag{7.127}$$

However, the phase φ of electromagnetic wave remains constant along the corresponding light ray trajectory [Misner *et al.*, 1973, Box 22.3] because of the null character of electromagnetic field. Indeed,

$$\frac{d\varphi}{d\lambda} = \frac{\partial\varphi}{\partial x^\alpha}\frac{dx^\alpha}{d\lambda} = \mathcal{K}_\alpha \mathcal{K}^\alpha = 0 , \tag{7.128}$$

where λ is the affine parameter along the light ray. For this reason, $d\varphi/d\varphi_0 = 1$, and equation (7.127) coincides with equation (7.123).

The time derivative dt_0/dt along the light-ray trajectory is calculated from equation (7.72) where one has to take into account that the function $B_a(s, s_0)$ depends on times t_0 and t not only through the retarded times $s_0 = s_0(t_0, \boldsymbol{x}_0(t_0))$ and $s = s(t, \boldsymbol{x}(t))$ standing in the upper and lower limits of the integral (7.74) but also through the time of the closest approach, t^*, and the unit vector $\boldsymbol{k}$ that must be considered as time-dependent parameters. Indeed, parameters t^* and $\boldsymbol{k}$ are defined by equations (7.34) and (7.67) which demonstrate clearly that they depend on time in case of moving source of light and/or observer. Hence, the derivative along the light ray must be written down as follows

$$\frac{dt_0}{dt} = \frac{1 + \boldsymbol{K}\cdot\boldsymbol{\beta} - \frac{2G}{c^3}\sum\limits_{a=1}^{N} m_a \left[\frac{\partial s}{\partial t}\frac{\partial}{\partial s} + \frac{\partial s_0}{\partial t}\frac{\partial}{\partial s_0} + \frac{\partial t^*}{\partial t}\frac{\partial}{\partial t^*} + \frac{\partial k^i}{\partial t}\frac{\partial}{\partial k^i}\right] B_a(s, s_0, t^*, \boldsymbol{k})}{1 + \boldsymbol{K}\cdot\boldsymbol{\beta}_0 + \frac{2G}{c^3}\sum\limits_{a=1}^{N} m_a \left[\frac{\partial s}{\partial t_0}\frac{\partial}{\partial s} + \frac{\partial s_0}{\partial t_0}\frac{\partial}{\partial s_0} + \frac{\partial t^*}{\partial t_0}\frac{\partial}{\partial t^*} + \frac{\partial k^i}{\partial t_0}\frac{\partial}{\partial k^i}\right] B_a(s, s_0, t^*, \boldsymbol{k})} ,$$

(7.129)

where the unit vector $\boldsymbol{K}$ is defined in (7.66) and one shows the explicit dependence of function B_a on all parameters which implicitly depend on time. Notice that taking times τ and τ_0 as primary quantities instead of t and t_0 brings in the retarded times s and s_0 dependence on the time of the closest approach t^*. Thus, one has to chose these functions either as $s = s(t, t_0)$, $s_0 = s_0(t_0)$ or $s = s(\tau, \tau_0, t^*)$, $s_0 = s_0(\tau_0, t^*)$. The latter choice introduces partial derivatives of s and s_0 with respect to t^* and modifies formula (7.129). One prefers to operate with functions $s = s(t, t_0)$, $s_0 = s_0(t_0)$ and equation (7.129) because the time of the closest approach t^* is an auxiliary quantity that does not enter any final result. Hence, it is more preferable to avoid dealing with it in the intermediate calculations any time as it is possible.

Partial time derivatives of vector $\boldsymbol{k}$ are calculated by using the approximation $\boldsymbol{k} = -\boldsymbol{K}$ and formula (7.66), where the coordinates of the source of light, $\boldsymbol{x}_0(t_0)$, and of the observer, $\boldsymbol{x}(t)$, are functions of the corresponding time. Straightforward differentiation yields

$$\frac{1}{c}\frac{\partial k^i}{\partial t} = \frac{(\boldsymbol{k}\times(\boldsymbol{\beta}\times\boldsymbol{k}))^i}{R}\,, \qquad \frac{1}{c}\frac{\partial k^i}{\partial t_0} = -\frac{(\boldsymbol{k}\times(\boldsymbol{\beta}_0\times\boldsymbol{k}))^i}{R}\,, \tag{7.130}$$

where $R = |\boldsymbol{x} - \boldsymbol{x}_0|$ is the coordinate distance between the observer and the source of light.

Partial derivative of the retarded times s with respect to t and that of s_0 with respect to t_0 are calculated from the gravity null-cone equations (7.13) and (7.75) respectively, where one has to take into account that the spatial position of the point of observation is not taken arbitrary but always connected to the point of emission of light by the unperturbed trajectory of light, $\boldsymbol{x}(t) = \boldsymbol{x}_0(t_0) + c\boldsymbol{k}(t - t_0)$. More explicitly, one has to use for the calculations of the partial derivatives the following relationships

$$s = t - \frac{1}{c}|\boldsymbol{x}_0(t_0) + c\boldsymbol{k}(t, t_0)\,(t - t_0) - \boldsymbol{x}_a(s)|\,, \tag{7.131}$$

$$s_0 = t_0 - \frac{1}{c}|\boldsymbol{x}_0(t_0) - \boldsymbol{x}_a(s_0)|\,, \tag{7.132}$$

where the unit light vector $\boldsymbol{k}$ must be considered as a two-point function of times t, t_0 with the partial time derivatives taken from equation (7.130). Physical meaning of relationship (7.131) is explained in Figure 7.3 demonstrating that in general relativity gravity and light propagate on one and the same bi-characteristic hypersurface of the null cone [Frolov, 1979]. A photon is emitted at time t_0 at the point $\boldsymbol{x}_0$ and arrives to observer located at point $\boldsymbol{x}$ at time t. Light propagates along a null characteristic of the Maxwell equations connecting the point of emission and the point of observation. As light propagates, the light-ray deflecting body a moves along its own worldline. Gravitational interaction of the body with the photon is not instantaneous because its obey the retarded *Liénard-Wiechert* solution of the Einstein equations. The null characteristic of the linearized Einstein equations connect the body's position $\boldsymbol{x}_a(s)$ taken at the retarded instant of time $s = t - r_a/c$. In fact, equation (7.131) and Figure

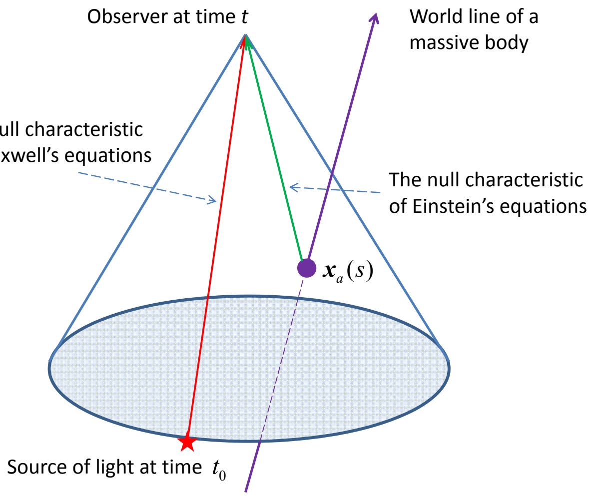

Figure 7.3 Null-cone is a hypersurface made of characteristics of the Maxwell and Einstein equations. Gravitational field of a moving massive body affects propagation of light on the null-cone hypersurface so that the light that has been received by observer at time t, experiences the pull of gravity of the massive body located at the retarded position $\boldsymbol{x}_a(s)$ on the null-cone.

7.3 say that the observed photon is affected by the gravitational field of the moving body from its retarded position. Calculation of infinitesimal variation of equations (7.131), (7.132) immediately gives for the partial derivatives

$$\frac{\partial s}{\partial t} = \frac{r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a}{r_a - \boldsymbol{\beta}_a\cdot\boldsymbol{r}_a} - \frac{(\boldsymbol{k}\times\boldsymbol{\beta})\cdot(\boldsymbol{k}\times\boldsymbol{r}_a)}{r_a - \boldsymbol{\beta}_a\cdot\boldsymbol{r}_a}\,, \tag{7.133}$$

$$\frac{\partial s}{\partial t_0} = \frac{1 - \boldsymbol{k}\cdot\boldsymbol{\beta}_0}{r_a - \boldsymbol{\beta}_a\cdot\boldsymbol{r}_a}(\boldsymbol{k}\cdot\boldsymbol{r}_a)\,, \tag{7.134}$$

$$\frac{\partial s_0}{\partial t_0} = \frac{r_{0a} - \boldsymbol{\beta}_0\cdot\boldsymbol{r}_{0a}}{r_{0a} - \boldsymbol{\beta}_{a0}\cdot\boldsymbol{r}_{0a}}\,, \tag{7.135}$$

$$\frac{\partial s_0}{\partial t} = 0\,. \tag{7.136}$$

Partial time derivatives of the parameter t^* are calculated from its original definition, $t^* = t_0 - c^{-1}\boldsymbol{k}\cdot\boldsymbol{x}_0(t_0)$, and are given by

$$\frac{\partial t^*}{\partial t_0} = 1 - \boldsymbol{k}\cdot\boldsymbol{\beta}_0 + \frac{\boldsymbol{\beta}_0\cdot\boldsymbol{\xi}}{R}\,, \qquad \frac{\partial t^*}{\partial t} = -\frac{\boldsymbol{\beta}\cdot\boldsymbol{\xi}}{R}\,, \tag{7.137}$$

where $\boldsymbol{\xi}$ is vector of the impact parameter of the light ray, and terms of the order of ξ/R in both formulas originate from the time derivatives of vector $\boldsymbol{k}$

Partial derivatives of the function $B_a(s, s_0, t^*, \boldsymbol{k})$ defined by the integral (7.74) read as follows

$$\frac{1}{c}\frac{\partial B_a}{\partial s} = \frac{1}{\sqrt{1-\beta_a^2}}\frac{(1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a)^2}{r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a}\,, \tag{7.138}$$

$$\frac{1}{c}\frac{\partial B_a}{\partial s_0} = -\frac{1}{\sqrt{1-v_{a0}^2}}\frac{(1-\boldsymbol{k}\cdot\boldsymbol{\beta}_{a0})^2}{r_{0a} - \boldsymbol{k}\cdot\boldsymbol{r}_{0a}}\,, \tag{7.139}$$

$$\frac{1}{c}\frac{\partial B_a}{\partial t^*} = C_a(s_0) - C_a(s)\,, \tag{7.140}$$

$$\frac{\partial B_a}{\partial k^i} = D_a^i(s_0) - D_a^i(s) + 2\left[E_a^i(s_0) - E_a^i(s)\right]\,, \tag{7.141}$$

where the partial derivative $\partial B_a/\partial t^*$ is expressed in terms of relationships (7.116), (7.119), while calculation of the partial derivative $\partial B_a/\partial k^i$ yields the results in terms of functions (7.117), (7.118) and (7.120), (7.121) respectively. One notices that the integral functions C_a, D_a, E_a defined by equations (7.119)-(7.121) may be not calculable analytically in general case of arbitrary worldlines of the massive bodies. However, if accelerations of gravitating bodies are small enough, velocity of each body can be considered as constant, thus, making derivatives (7.140), (7.141) approximated by simpler expressions

$$\frac{1}{c}\frac{\partial B_a}{\partial t^*} = -\frac{1}{\sqrt{1-\beta_a^2}}\frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a}{r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a} + \frac{1}{\sqrt{1-\beta_{a0}^2}}\frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_{a0}}{r_{0a} - \boldsymbol{k}\cdot\boldsymbol{r}_{0a}}\,, \tag{7.142}$$

$$\begin{aligned}\frac{\partial B_a}{\partial k^i} = &-\frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a}{\sqrt{1-\beta_a^2}}\frac{x_a^j(s)}{r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a} + \frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_{a0}}{\sqrt{1-\beta_{a0}^2}}\frac{x_a^j(s_0)}{r_{0a} - \boldsymbol{k}\cdot\boldsymbol{r}_{0a}} \\ &+\frac{2\beta_a^j}{\sqrt{1-\beta_a^2}}\ln(r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a) - \frac{2\beta_{a0}^j}{\sqrt{1-\beta_{a0}^2}}\ln(r_{0a} - \boldsymbol{k}\cdot\boldsymbol{r}_{0a})\,.\end{aligned} \tag{7.143}$$

Residual terms to these equations can be always calculated from the integrals in (7.119)-(7.121) whenever it is necessary. They are unimportant for interpretation of current astrometric measurements.

Careful inspection of previous equations given in this section, reveals that

$$1 + \boldsymbol{K}\cdot\boldsymbol{\beta} - \frac{2G}{c^3}\sum_{a=1}^{N} m_a\left[\frac{\partial t^*}{\partial t}\frac{\partial B_a}{\partial t^*} + \frac{\partial k^i}{\partial t}\frac{\partial B_a}{\partial k^i}\right] = 1 - \boldsymbol{k}\cdot\boldsymbol{\beta}\,, \tag{7.144}$$

$$\begin{aligned}1 + \boldsymbol{K}\cdot\boldsymbol{\beta}_0 + \frac{2G}{c^3}\sum_{a=1}^{N} m_a\left[\frac{\partial t^*}{\partial t_0}\frac{\partial B_a}{\partial t^*} + \frac{\partial k^i}{\partial t_0}\frac{\partial B_a}{\partial k^i}\right] &= (1 - \boldsymbol{k}\cdot\boldsymbol{\beta}_0)\\ &\times\left\{1 + \frac{2G}{c^3}\sum_{a=1}^{N} m_a\left[C_a(s_0) - C_a(s)\right]\right\}\,,\end{aligned} \tag{7.145}$$

where one has used equation (7.140) and equations (7.67)-(7.69) connecting the unit vectors $\boldsymbol{K}$ and $\boldsymbol{k}$. Hence, equation (7.129) can be recast in

$$\frac{dt_0}{dt} = \frac{1 - \boldsymbol{k}\cdot\boldsymbol{\beta} - \frac{2G}{c^3}\sum_{a=1}^{N} m_a\frac{\partial s}{\partial t}\frac{\partial B_a}{\partial s}}{(1 - \boldsymbol{k}\cdot\boldsymbol{\beta}_0)\left\{1 + \frac{2G}{c^3}\sum_{a=1}^{N} m_a\left[C_a(s_0) - C_a(s)\right]\right\} + \frac{2G}{c^3}\sum_{a=1}^{N} m_a\left[\frac{\partial s}{\partial t_0}\frac{\partial B_a}{\partial s} + \frac{\partial s_0}{\partial t_0}\frac{\partial B_a}{\partial s_0}\right]}\,, \tag{7.146}$$

where the partial derivatives of the function B_a are calculated from equations (7.138), (7.139).

7.6
Applications to Relativistic Astrophysics and Astrometry

7.6.1
Gravitational time delay in binary pulsars

7.6.1.1 *Pulsars - Rotating Radio Beacons*

Radio astronomical measurements of pulsars allow the study of a variety of questions in fundamental physics. A pulsar is a neutron star which emits a beam of electromagnetic radiation towards observer on the Earth. The radiation is not continuous but arrives in the form of pulses due to a misalignment of the neutron star's rotation axis and its magnetic axis. In other words, pulsars pulse because the radiation generated within the magnetic field sweeps in and out of the line of sight of the observer with a regular period of the pulsar's revolution around its own axis. The observed periods of pulsars range from 1.4 milliseconds to 8.5 seconds [Lorimer and Kramer, 2004].

Binary pulsar consists of a pulsar and its companion star both being in orbit around their mutual center of gravity (barycenter). The pulsar's companion can be either an ordinary star of a main sequence, or a white dwarf, or a neutron star, or even a black hole. When companion of the pulsar is a neutron star which is also observed as a pulsar, the binary system is called a double pulsar [Kramer and Wex, 2009].

Timing of binary pulsars is one of the most important methods of testing general relativity in the strong gravitational field regime [Damour, 1984; Damour and Esposito-Farèse, 1992, 1998a; Damour and Schäfer, 1991; Damour and Taylor, 1992; Kramer *et al.*, 2006; Lorimer and Kramer, 2004; Stairs, 2005; van Straten *et al.*, 2001; Zaglauer, 1992]. Relativistic effects in binary pulsars are parameterized by a set of the post-Keplerian (PK) parameters introduced by Damour and Deruelle [1986]. Some of the parameters characterize the orbital motion of the stars (see section 6.4.3.2 of this book for more detail) and the others propagation of radio signals through the gravitational field of the pulsar and its companion. The basic PK parameters quantify different relativistic effects and can be analyzed using a theory-independent procedure in which masses of the two stars are the only dynamic unknowns [Damour and Taylor, 1992; Edwards *et al.*, 2006; Hobbs *et al.*, 2006]. Each of the basic PK parameters depends on the masses of two orbiting stars in a different functional way. Consequently, if three or more PK parameters can be measured, the number of equations connecting the masses with observations becomes overdetermined that can be used to test the consistency of the gravitational theory. Besides masses, other PK parameters depend on the angles characterizing orientation of pulsar's rotational axis and their measurement can tell us how strong the pulsar's spin couples with the orbital momentum of the binary system (geodetic precession) [Kramer, 1998; Stairs *et al.*, 2004; Weisberg and Taylor, 2002].

Especially important for this test are the binary pulsars with relativistic orbits being visible nearly edge-on. In such systems pulsar astronomers can determine masses of orbiting stars by measuring two propagation-related PK parameters - the *range* and *shape* of the gravitational time delay in propagation of radio pulses from pulsar to observer - independently of other relativistic effects. Perhaps, the most famous examples of the nearly edge-on binary pulsars are PSR B1855+09 and PSR B1534+12 [Lorimer and Kramer, 2004]. The sine of inclination angle, i, of the orbit of PSR B1855+09 to the line of sight makes up a value of about 0.9992 and the range PK parameter reaches 1.27 μs (microseconds) [Kaspi *et al.*, 1994]. The corresponding quantities for PSR B1534+12 are $\sin i = 0.982$ and 6.7 μs [Stairs *et al.*, 1998]. Few years ago a double pulsar PSR J0737-3039 which is a member of a highly relativistic double-neutron-star binary with an orbital period of 2.4 hours, was discovered [Burgay *et al.*, 2003; Possenti *et al.*, 2004]. For this system, the shape PK parameter $\sin i = 0.99974$ and the range is 6.21 μs are giving an unprecedented opportunity to test the gravitational time delay in the strong gravity-field regime [Kramer and Wex, 2009].

All binary pulsars emit gravitational waves which take the orbital energy and angular momentum of the binary pulsar away from the system. The process of generation and emission of the gravitational waves leads to the appearance in equations of motion of the binary system the gravitational radiation-reaction force that causes a secular decrease in the binary's orbital period [Damour, 1983; Grishchuk and Kopeikin, 1986; Schäfer, 1985]. This general-relativistic prediction was confirmed in binary pulsar PSR B1913+16 with the precision of about 0.3% by Taylor and collaborators [Taylor, 1994]. The orbital period decay of the double pulsar PSR J0737-3039 agrees with general relativity at the level of 1.4% [Kramer *et al.*, 2006] but can be further improved down to 0.01 % uncertainty level [Deller *et al.*, 2009]. New achievements in technological development, building and continuous upgrading the largest radio telescopes extend the potential of experimentalists to measure with a higher precision relativistic effects associated with propagation of a radio pulse from pulsar through the time-dependent gravitational field of binary pulsars towards observer on the Earth. The temporal variability of the gravitational field during propagation of a pulsar's pulse produces additional effects in pulsar timing observations which reveal themselves in the form of small corrections to the *static* Shapiro time delay making its interpretation a more challenging problem. This is not a goal of the present book to discuss all details of the binary pulsar timing observations. The discussion will be focused on relativistic effects in propagation of radio pulses and present the exact Lorentz covariant theory of the gravitational time delay in binary pulsars which includes, besides of the well known *static* Shapiro delay, all relativistic corrections for the velocities of pulsar and its companion. The general formula will be restricted to terms which are linear with respect to the velocities. The reason is that the terms being quadratic with respect to velocities are proportional to gravitational energy of the system due to the virial theorem. Hence, the proper treatment of the quadratic with respect to velocity terms can be achieved only within the second post-Minkowskian approximation, which is much more complicated and is not considered in the present book. Gravitational time delay in propagation of radio pulses from the pulsar to observer has been computed by Blandford & Teukolsky [Blandford and Teukolsky, 1976] and reproduced later in many other works under assumption that gravitational field of the binary pulsar is static and weak everywhere. This assump-

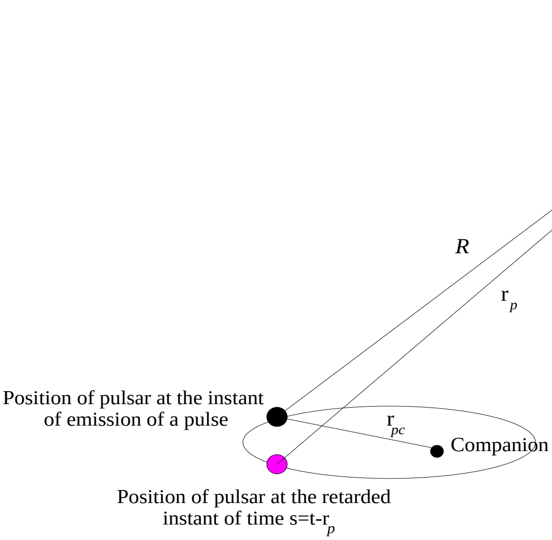

Figure 7.4 Propagation of radio pulse from a binary pulsar to observer. The pulsar emits radio signal at time t_0 which reaches observer at time t. Gravitational time delay in propagation of the pulse depends on the retarded positions of the pulsar and its companion taken at the retarded instants of time $s_0 = t_0 - r_{0a}/c$ and $s = t - r_a/c$ corresponding to times t_0 and t.

tion significantly simplifies calculations but does not take into account higher-order effects caused by time-dependent part of the gravitational field of the binary system. Indeed, as the electromagnetic pulse moves toward observer the time-dependent part of the gravitational field of the binary system (gravitational wave) propagates to the observer as well. As gravity and light wave move with the same fundamental speed c in general relativity, this could potentially lead to a secular accumulation of the impact of the gravitational wave on the radio pulse leading to subsequent amplification of the amplitude (range) of the gravitational time delay and to the change in its detected profile (shape). Some improvement in calculation of the gravitational time delay was achieved in papers by Klioner [1991a]; Wex [1995] where the propagation of a radio wave through gravitational field of a uniformly moving masses was considered. Integration of the pulse's propagation equation in the field of the *Liénard-Wiechert* gravitational potentials allows us to remove any restrictions on the worldlines of the pulsar and its companion. This retarded-time approach to solving the problem of calculation of gravitational time delay in binary pulsars was proposed by Kopeikin and Schäfer [1999] and is discussed in the rest of this section.

7.6.1.2 *The approximation scheme*

Figure 7.4 illustrates geometry of a pulsar's radio pulse propagation and the mutual positions of the binary pulsar and observer. Let us stipulate that the origin of the coordinate system is at the barycenter of the binary pulsar. Radio pulses are emitted rather close to the surface of the pulsar and the coordinates of the point of emission,

$\boldsymbol{x}_0$, can be approximated by equation

$$\boldsymbol{x}_0 = \boldsymbol{x}_p(t_0) + \boldsymbol{X}(t_0) , \tag{7.147}$$

where $\boldsymbol{x}_p$ are the barycentric coordinates of the pulsar's center-of -mass, and $\boldsymbol{X}$ is the barycentric radius-vector of the point of emission both taken at the moment of emission of the radio pulse, t_0. At the moment of emission the spatial orientation of the pulsar's radio beam is almost the same with respect to observer at the Earth otherwise the pulse will not reach the observer. Hence, one is allowed to assume with a good approximation that vector $\boldsymbol{X}$ is constant at every instant of time when an emission of a radio pulse takes place. More precisely, the coordinates of the point of emission are constant in the pulsar proper reference frame, which construction is similar to the geocentric frame in the solar system (see section 4.5 of this book). Relativistic post-Newtonian transformation from the proper reference frame of the pulsar to the binary pulsar barycentric coordinate system is also similar to the transformation between the geocentric and barycentric coordinates of the solar system given in section 5.1 of the present book. The transformation reveals that if the pulsar moves along elliptic orbit the barycentric vector $\boldsymbol{X}$ actually depends on time because of the relativistic contractions induced by the orbital velocity and gravitational field of the companion. However, this periodic relativistic perturbation of the vector is of the order of $(|\boldsymbol{X}|/c)(v_p^2/c^2)$ where v_p is a characteristic velocity of the pulsar with respect to the barycenter of the binary system. For a typical distance $|\boldsymbol{X}| \simeq 50 \div 100$ km and the orbital velocity $v_p/c \simeq 10^{-3}$ [Lorimer and Kramer, 2004] this is too small for being measurable. Another reason for temporal variations of the barycentric vector $\boldsymbol{X}$ arises due to the effects of aberration of the pulsar's beam [Smarr and Blandford, 1976], the orbital pulsar parallax [Kopeikin, 1995], and the gravitational bending delay [Doroshenko and Kopeikin, 1995]. These effects can be directly measured but they are small and can be neglected in calculation of the gravitational time delay.

In what follows, the calculation of the gravitational time delay is performed in the approximation which is linear with respect to velocities of the pulsar and its companion. Hence, formula (7.81) plays a key role where one has to evaluate the integral depending on the acceleration of the massive bodies. However, before performing the integral in this formula it is instructive to derive a relationship between the retarded times s and s_0 given by expressions (7.13) and (7.75) respectively. Subtracting equation (7.75) from (7.13) and taking into account relationship (7.72), one obtains

$$s - s_0 = \frac{1}{c}\left(R - r_a + r_{0a}\right) + \Delta(t, t_0) , \tag{7.148}$$

where $R = |\boldsymbol{R}|$, $\boldsymbol{R} = \boldsymbol{x} - \boldsymbol{x}_0$, $r_a = |\boldsymbol{x} - \boldsymbol{x}_a(s)|$, and $r_{0a} = |\boldsymbol{x}_0 - \boldsymbol{x}_a(s_0)|$, $\boldsymbol{x}$ is the coordinates of observer, and the index a numerates the pulsar and its companion[9]. Notice that the point of observation, $\boldsymbol{x}$, is separated from the binary system by a very large distance typically equal to several hundred parsec [Lorimer and Kramer, 2004], and it can be approximated as $|\boldsymbol{x}| \simeq R$. On the other hand, the size of the binary system can not exceed the distance denoted as r_{0a}. Thus, the Taylor expansion of r_a with respect to

9) It may be worth to remind again that the retarded time s is calculated for each body separately.

the small parameter r_{0a}/R is admissible. The first two terms of the expansion are

$$r_a = |\boldsymbol{R} + \boldsymbol{x}_0 - \boldsymbol{x}_a(s)| = R - \boldsymbol{K}\cdot[\boldsymbol{x}_0 - \boldsymbol{x}_a(s)]\,, \tag{7.149}$$

where the unit vector $\boldsymbol{K} = -\boldsymbol{R}/R$ as defined in (7.66), and the residual terms are negligibly small. Using the approximation $\boldsymbol{K} = -\boldsymbol{k} + O(G)$, where $O(G)$ indicates the gravitational perturbation from equation (7.67), formula (7.148) is reduced to the form

$$s - s_0 = \frac{1}{c}\left\{(r_{0a} - \boldsymbol{k}\cdot\boldsymbol{r}_{0a}) + \boldsymbol{k}\cdot[\boldsymbol{x}_a(s) - \boldsymbol{x}_a(s_0)]\right\} + O(G)\,, \tag{7.150}$$

which explicitly shows that the difference between the retarded times s and s_0 is of the order of time interval being required for both gravity and light to cross the binary system (light-crossing time). It is this interval which is characteristic in the problem of propagation of radio pulses from the binary system to observer on the Earth. Therefore, the retarded time s taken along the light ray trajectory changes only a little (few seconds) during the entire process of propagation of light from pulsar to observer while the coordinate time t changes enormously (hundred of light years). This remarkable fact is important for understanding how the time-dependent gravity field of the binary system affects propagation of electromagnetic signals from distant astronomical systems to observer on the Earth.

In addition to the expression (7.150) , one can show that time differences, $s_0 - t_0$, and, $s - t_0$, are also of the same order of magnitude as $s - s_0$. Indeed, assuming that the velocities of pulsar and its companion are small compared to the fundamental speed c, one gets from (7.75) and (7.150) for these intervals

$$s_0 - t_0 = -\frac{1}{c}|\boldsymbol{x}_0 - \boldsymbol{x}_a(s_0)| = -\frac{1}{c}\left(\rho_{0a} - \boldsymbol{\rho}_{0a}\cdot\boldsymbol{\beta}_a\right) + O\left(\beta_a^2\right) + O(G)\,, \tag{7.151}$$

$$s - t_0 = -\frac{1}{c}(\boldsymbol{k}\cdot\boldsymbol{\rho}_{0a})(1 - \boldsymbol{k}\cdot\boldsymbol{\beta}_a) + O\left(\beta_a^2\right) + O(G)\,, \tag{7.152}$$

where $\boldsymbol{\rho}_{0a} = \boldsymbol{x}_0 - \boldsymbol{x}_a(t_0)$, $\rho_{0a} = |\boldsymbol{\rho}_{0a}|$, and $\boldsymbol{\beta}_a \equiv \boldsymbol{v}_a(t_0)/c$. The relationships (7.151), (7.152) prove that the time intervals under discussion are really small and comparable with the time taken by light to cross the binary system (few seconds or less). It means that the coordinates and their time derivatives of the pulsar and its companion can be expanded in the integrands of all integrals taken over the null cone in a rapidly converging Taylor series around the time of emission, t_0, of a radio signal in powers of $s - t_0$ and/or $s_0 - t_0$. Figure 7.2 explains the relationships between positions of photon on the light-ray trajectory and the retarded positions of pulsar (index p) and its companion (index c).

In what follows, the efforts are focused on the derivation of the linear with respect to velocity of moving bodies, corrections to the static-field Shapiro delay. Calculations are done by using expression (7.81) for the time delay which is already explicitly contains some terms being proportional to the ratio v_a/c. Equation (7.81) also contains time-dependent integral from the orbital acceleration, $\dot{\boldsymbol{\beta}}_a = \dot{\boldsymbol{v}}_a/c$, of the bodies. This is not a table integral even in the simplest case of a circular orbit. Therefore, one should use an approximation method. The integral is taken from the retarder time s_0 to the retarded time s and, according to equations (7.148), (7.152) the time difference $s - s_0$ is about the same as $s - t_0$, and is about the time taken by light to cross the binary system that is much smaller than the orbital period P_b. For this reason, one is allowed to expand the coordinates and acceleration of the massive body in the Taylor series

$$\boldsymbol{x}_a(\zeta) = \boldsymbol{x}_a(t_0) + \boldsymbol{v}_a(t_0)(\zeta - t_0) + ... \tag{7.153}$$

$$\dot{\boldsymbol{\beta}}(\zeta) = \dot{\boldsymbol{\beta}}(t_0) + \ddot{\boldsymbol{\beta}}(t_0)(\zeta - t_0) + ... \,, \tag{7.154}$$

where the ellipsis denote terms of higher order. In fact, the second term in the right side of equations (7.153), (7.154) is smaller than the first one by a factor of $\beta_a = v_a/c$, which is neglected. Hence, in order to perform the integration in equation (7.81) it is sufficient to take into account only the first terms in expansions (7.153), (7.154). Then, the integral under consideration is approximated as follows

$$\int_{s_0}^{s} \ln(r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a)(\boldsymbol{k}\cdot\dot{\boldsymbol{v}}_a)\, d\zeta = \boldsymbol{k}\cdot\dot{\boldsymbol{v}}_a(t_0) \int_{s_0}^{s} \ln\left[t^* + \boldsymbol{k}\cdot\boldsymbol{x}_a(t_0) - \zeta\right] d\zeta \,, \tag{7.155}$$

and is reduced to the table form. Accounting for (7.150)-(7.152), the result of integration yields

$$\begin{aligned}\int_{s_0}^{s} \ln\left[t^* + \boldsymbol{k}\cdot\boldsymbol{x}_a(t_0) - \zeta\right] d\zeta = & \frac{1}{c}(r_{0a} - \boldsymbol{k}\cdot\boldsymbol{r}_{0a})\left[\ln(r_{0a} - \boldsymbol{k}\cdot\boldsymbol{r}_{0a}) - 1\right] \\ & -\frac{1}{c}(r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a)\ln(r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a) + O(\beta_a)\,,\end{aligned} \tag{7.156}$$

where r_a and r_{0a} have the same meaning as in equation (7.148). The result (7.156) is multiplied by the radial acceleration of the gravitating body according to equation (7.155). Simple numerical estimate of the integral (7.155) tells us that it can reach in a binary system the maximal magnitude of the order of $(Gm_a/c^3)(x/P_b)\beta_a \ln(1 - \sin i)$, where x is the projected semi-major axis of the binary system expressed in light seconds, P_b is its orbital period, and i is the angle of inclination of the orbital plane of the binary system to the line of sight. For a typical edge-on binary pulsar like PSR B1534+12 the terms under discussion are about $10^{-4} \div 10^{-5}$ μs which is too small to be measured. For this reason, all terms depending on the acceleration of the pulsar and its companion will be omitted from the following considerations.

Let us expand coordinates of the a-th body taken at the retarded time s in Taylor series in the neighborhood of time s_0

$$\boldsymbol{x}_a(s) = \boldsymbol{x}_a(s_0) + \boldsymbol{v}_a(s_0)(s - s_0) + O[(s - s_0)^2] \tag{7.157}$$

Accounting for equation (7.150) and approximating $\boldsymbol{r}_{0a} = \boldsymbol{\rho}_{0a}$ in all velocity-dependent terms, allows us to recast previous equation to

$$\boldsymbol{x}_a(s) = \boldsymbol{x}_a(s_0) + \boldsymbol{\beta}_a\,(\rho_{0a} - \boldsymbol{k}\cdot\boldsymbol{\rho}_{0a}) + O(\beta_a^2)\,. \tag{7.158}$$

Making use of this expansion, one can prove that a very large distance, r_a, between the binary system and observer, relates to the small one, r_{0a}, by important relationship

$$r_a^2 - (\boldsymbol{k}\cdot\boldsymbol{r}_a)^2 = r_{0a}^2 - (\boldsymbol{k}\cdot\boldsymbol{r}_{0a})^2 - 2(\rho_{0a} - \boldsymbol{k}\cdot\boldsymbol{\rho}_{0a})\left[\boldsymbol{\beta}_a\cdot\boldsymbol{\rho}_{0a} - (\boldsymbol{k}\cdot\boldsymbol{\rho}_{0a})(\boldsymbol{k}\cdot\boldsymbol{\beta}_a)\right] + O(\beta_a^2)\,.$$

(7.159)

Moreover,

$$r_{0a} + \boldsymbol{k}\cdot\boldsymbol{r}_{0a} = \rho_{0a} + \boldsymbol{k}\cdot\boldsymbol{\rho}_{0a} + \boldsymbol{\beta}_a\cdot\boldsymbol{\rho}_{0a} + (\boldsymbol{k}\cdot\boldsymbol{\beta}_a)\rho_{0a} + O(\beta_a^2)\,. \tag{7.160}$$

As a consequence of simple algebra, one obtains

$$\frac{r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a}{r_{0a} - \boldsymbol{k}\cdot\boldsymbol{r}_{0a}} = \frac{r_a^2 - (\boldsymbol{k}\cdot\boldsymbol{r}_a)^2}{r_{0a}^2 - (\boldsymbol{k}\cdot\boldsymbol{r}_{0a})^2}\,\frac{r_{0a} + \boldsymbol{k}\cdot\boldsymbol{r}_{0a}}{r_a + \boldsymbol{k}\cdot\boldsymbol{r}_a}\,, \tag{7.161}$$

and after making use of equations (7.159), (7.160) it yields the following result

$$\frac{r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a}{r_{0a} - \boldsymbol{k}\cdot\boldsymbol{r}_{0a}} = \frac{1 + \boldsymbol{k}\cdot\boldsymbol{\beta}_a}{r_a + \boldsymbol{k}\cdot\boldsymbol{r}_a}\left[\rho_{0a} + \boldsymbol{k}\cdot\boldsymbol{\rho}_{0a} - \boldsymbol{\beta}_a\cdot\boldsymbol{\rho}_{0a} + (\boldsymbol{k}\cdot\boldsymbol{\rho}_{0a})(\boldsymbol{k}\cdot\boldsymbol{\beta}_a)\right] + O\left(\beta_a^2\right)\,. \tag{7.162}$$

It is straightforward to prove that

$$r_a + \boldsymbol{k}\cdot\boldsymbol{r}_a = 2\,(R + \boldsymbol{k}\cdot\boldsymbol{r}_{0a}) + O\left(\frac{r_{0a}^2}{R}\right)\,, \tag{7.163}$$

where $R = |\boldsymbol{R}|$ is the distance from the point of emission to the point of observation. This distance is expanded as follows

$$\boldsymbol{R} \;=\; \boldsymbol{\mathcal{R}} + \boldsymbol{x}_E + \boldsymbol{w} - \boldsymbol{x}_p - \boldsymbol{X}\,, \tag{7.164}$$

where $\boldsymbol{\mathcal{R}}$ is the distance between the barycenter of the binary pulsar and that of the solar system, $\boldsymbol{x}_E$ is the distance from the barycenter of the solar system to the center of mass of the Earth (geocenter), $\boldsymbol{w}$ is the geocentric position of the radio telescope, $\boldsymbol{x}_p$ are coordinates of the center of mass of the pulsar with respect to the barycenter of the binary system, and $\boldsymbol{X}$ are coordinates of the point of emission of the radio pulse with respect to the pulsar proper reference frame. The distance $\boldsymbol{\mathcal{R}}$ is very large and gradually changing because of the proper (tangential) motion of the binary system in the sky or its radial motion with respect to the solar system. Proper motion of any star is small and can be neglected in calculation of the gravitational time delay of light. Hence, the distance $\boldsymbol{\mathcal{R}}$ can be taken as a constant in the gravitational time delay. All other distances in formula (7.164) are of order of either diurnal, or annual, or pulsar's orbital parallax [Kopeikin, 1995] with respect to the very large distance $\boldsymbol{\mathcal{R}}$. Such a relationship between the distances is more than sufficient to approximate

$$\ln(r_a + \boldsymbol{k}\cdot\boldsymbol{r}_a) \;=\; \ln(2\mathcal{R}) + O\left(\frac{r_{0a}}{R}\right) \propto \text{const.}\,, \tag{7.165}$$

where $\mathcal{R} = |\boldsymbol{\mathcal{R}}|$. Constant terms and polynomials of time are not directly observable in pulsar timing because they are absorbed in the parameters of the rotational phase of the pulsar [Lorimer and Kramer, 2004]. For this reason, the term $\ln(r_a + \boldsymbol{k}\cdot\boldsymbol{r}_a)$ will be omitted from the final expression for the gravitational time delay in the binary pulsar.

Accounting for all approximations having been developed in this section, equation (7.81) is simplified and reduced to the form

$$\Delta(t,t_0) = -\frac{2G}{c^3}\sum_{a=1}^{2} m_a\Big\{(1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a)\ln\left[\rho_{0a}+\boldsymbol{k}\cdot\boldsymbol{\rho}_{0a}-\boldsymbol{\beta}_a\cdot\boldsymbol{\rho}_{0a}+(\boldsymbol{k}\cdot\boldsymbol{\rho}_{0a})(\boldsymbol{k}\cdot\boldsymbol{\beta}_a)\right]+\boldsymbol{k}\cdot\boldsymbol{\beta}_a\Big\}$$
$$+O\left(\frac{Gm_a}{c^3}\frac{v_a^2}{c^2}\right)+O\left(\frac{Gm_a}{c^3}\frac{v_a}{c}\frac{x}{P_b}\right)+O\left(\frac{Gm_a}{c^3}\frac{x}{\mathcal{R}}\right). \tag{7.166}$$

This post-Newtonian formula was derived by Kopeikin and Schäfer [1999] and it completes the analytic derivation of the velocity-dependent corrections to the static Shapiro time delay in binary systems. One has included to this equation estimates of the residual terms, which help to vizualize the relativistic effects omitted from consideration.

7.6.1.3
Post-Newtonian versus post-Minkowski calculations of time delay in binary systems

Until recently the post-Newtonian approximations were a basic tool used by researchers in order to calculate relativistic effects in the light time delay, the light deflection angle, electromagnetic frequency change, etc. It works very well in case of static gravitational field because in this case gravity interacts with photon instantaneously. However, if the gravitational field is not static one has to take into account that gravity propagates with finite speed and its interaction with photon takes place on the hypersurface of the null cone. It looks like the problem gets enormously complicated from the mathematical point of view. This motivated development of other approximated techniques. In particular, in order to take into account gravitomagnetic effects of moving bodies on light propagation, the bodies were assumed to be moving along straight lines with constant velocities [Klioner, 1991a, 2003b; Klioner and Kopeikin, 1992; Wex, 1995; Wex and Kopeikin, 1999]. It is obvious that this assumption is valid for binary systems only for a short interval of time and can not be used for calculation of propagation of radio pulses from pulsar to observer if one wants to account for more subtle relativistic effects like the effects caused by gravitational waves emitted by the binary system. The Lorentz-invariant post-Minkowskian approach, adopted in this book, is superior with respect to the post-Newtonian approximations as it takes into account all possible effects of time-dependent gravitational field on the light propagation. It also helps to understand the limits of the post-Newtonian calculations of the light propagation made by previous authors and to recognize that it gives correct results if only the linear velocity-dependent terms in formula (7.166) for the gravitational time delay in binary systems is taken into account .

Let us recall that the post-Newtonian theory operates with the instantaneous values of the gravitational potentials (the metric tensor) calculated in the near zone of the gravitating system, that is in the region of space around the binary system with a characteristic size being equal to a length of gravitational waves emitted by the system. Coordinates $\boldsymbol{x}_a(t)$ of gravitating bodies are arguments of the metric tensor depending in the post-Newtonian scheme on the current value of the coordinate time t. Thus, the metric tensor taken on the light-ray trajectory is a function of time $g_{\alpha\beta}[\boldsymbol{x}, \boldsymbol{x}_a(t)]$, where $\boldsymbol{x} \equiv \boldsymbol{x}(t)$ is coordinate of the photon. If the body's coordinates are expanded around the time of emission, t_0, of a radio pulse one gets for the components of the metric tensor a Taylor expansion, which reads as follows

$$g_{\alpha\beta}[\boldsymbol{x}, \boldsymbol{x}_a(t)] = g_{\alpha\beta}(\boldsymbol{x}, \boldsymbol{x}_a) + \frac{\partial g_{\alpha\beta}(\boldsymbol{x}, \boldsymbol{x}_a)}{\partial x^i_a} v^i_a(t - t_0) + \frac{\partial^2 g_{\alpha\beta}(\boldsymbol{x}, \boldsymbol{x}_a)}{\partial x^i_a \partial x^j_a} v^i_a v^j_a (t - t_0)^2 + \dots , \tag{7.167}$$

where $\boldsymbol{x}_{a0} \equiv \boldsymbol{x}_a(t_0)$, $\boldsymbol{v}_a \equiv \boldsymbol{v}_a(t_0)$, and the ellipsis denotes cubic and higher-order terms with respect to time difference $t - t_0$. In the post-Newtonian scheme, this Taylor expansion is used in the equations of light propagation to find the photon's trajectory. It is easy to observe that the expansion is divergent if the time interval $t - t_0$ exceeds the orbital period P_b of the binary system. This is the reason why the assumption of the uniform and rectilinear motion of the bodies in the binary system does not work if the time of integration of the equations of light propagation is longer than the orbital period. In particular, it explains why the post-Newtonian scheme can not be used alone for calculation of the effect of gravitational waves on propagation of light rays.

On the other hand, the post-Minkowski scheme gives components of the metric tensor in terms of the *Liénard-Wiechert* potentials that are functions of the retarded time $s = s(t, \boldsymbol{x})$ given in (7.13). On the light-ray trajectory the metric tensor is now a function $g_{\alpha\beta}[\boldsymbol{x}(t), \boldsymbol{x}_a(s)]$. One has shown that in terms of the retarded time argument the characteristic time for the interaction of gravity with a radio pulse while it moves from the pulsar to observer corresponds to the interval of time being required for light to cross the system. During this time the gravitational potentials can not change their numerical values too much because of the slow motion of the gravitating bodies. Hence, if coordinates of the bodies are expanded around t_0 one gets for the metric tensor expressed in terms of the *Liénard-Wiechert* potentials the following Taylor expansion

$$g_{\alpha\beta}[\boldsymbol{x}, \boldsymbol{x}_a(s)] = g_{\alpha\beta}(\boldsymbol{x}, \boldsymbol{x}_a) + \frac{\partial g_{\alpha\beta}(\boldsymbol{x}, \boldsymbol{x}_a)}{\partial x^i_a} v^i_a(s - t_0) + \frac{\partial^2 g_{\alpha\beta}(\boldsymbol{x}, \boldsymbol{x}_a)}{\partial x^i_a \partial x^j_a} v^i_a v^j_a (s - t_0)^2 + \dots , \tag{7.168}$$

where $\boldsymbol{x} \equiv \boldsymbol{x}(t)$, $\boldsymbol{x}_a \equiv \boldsymbol{x}_a(t_0)$, $\boldsymbol{v}_a \equiv \boldsymbol{v}_a(t_0)$, and the ellipsis denotes cubic and higher-order terms with respect to time difference $s - t_0$. This type of the Taylor expansion always converges because the time difference $s - t_0$ is much smaller than the orbital period of the binary system (see equation (7.152)).

It can be easily seen that the first terms in the Taylor expansions (7.167) and (7.168) coincide exactly. Furthermore, the retarded time s is a linear function of time t if quadratic with respect to velocity terms, are omitted from the retarded time equation (7.16), which implies that solutions of the equations of light propagation depending only on the first and second terms in the Taylor expansions (7.167) and (7.168) should be formally identical independently on what kind of approximation scheme is used for finding the metric tensor. Differences in the solution of the light-ray equation will arise only if one employs the terms which are quadratic with respect to velocities in the Taylor expansions (7.167) and (7.168).

Thus, the post-Newtonian approximation works fairly well for finding the part of the solution of the equations of light geodesics, which depends linearly on velocities of the bodies. However, it can not be used for calculation of perturbations of the light-ray trajectory caused by the accelerations of massive bodies. The post-Newtonian scheme can be applied without restriction only if the length of the light ray trajectory is small compared with the size of the gravitating system. This situation is realized in the observations of the solar system objects which will be analyzed later in more detail. One also notices that calculations of the light-ray propagation based on the post-Newtonian approximations are more cumbersome than those based on the post-Minkowskian approximation scheme which properly accounts for all retardation effects in the light propagation caused by motion of bodies by means of the *Liénard-Wiechert* potentials [Kopeikin and Fomalont, 2006].

7.6.1.4 *Time Delay in the Parameterized Post-Keplerian Formalism*

The parameterized post-Keplerian (PPK) formalism was introduced by Damour and Deruelle [1985, 1986] and further improved by Damour and Taylor [1992]. PPK formalism utilizes the orbital solution for a binary system in the form of the Damour-Deruelle parametrization that has been discussed in section 6.4.3.2 of this book. A standard timing model in PPK formalism consists of the five Keplerian parameters $(P_b, T_0, e_0, \omega_0, x_0)$, and 8 post-Keplerian parameters $(k, \gamma, \dot{P}_b, r, s; \delta_\theta, \dot{e}, \dot{x})$. There is a set of supplementary parameters caused by aberration of light, bending of light, Lorentz transformation between various reference frames, geodetic precession of pulsar's spin, and others [Damour and Taylor, 1992; Doroshenko and Kopeikin, 1995; Edwards *et al.*, 2006; Kopeikin, 1994; Kramer and Wex, 2009]. For any particular theory of gravitation, the values of a post-Keplerian parameter are a function of the measured Keplerian parameters and the two masses in the system (m_1, m_2). Thus, measuring any of the post-Keplerian parameters defines an allowed line in the $m_1 - m_2$ plane. First and second post-Keplerian parameters, k and γ [10], provides a measurement of m_1 and m_2. Each further parameter tests the theory of gravity by requiring that all the lines in the $m_1 - m_2$ plane meet at a point, allowing up to 6 different tests from the 8 PPK parameters. Currently, the original binary pulsar (PSR B1913+16) has yielded good measurements of three PK parameters $(k, \gamma, \dot{P}_b)$. This yields one test of relativistic gravity, and shows that general relativistic prediction of the orbital period decay due to emission of gravitational waves is correct. Two more strong-field tests of gravity have come from the more recently discovered pulsar PSR B1534+12. Measurements of k, γ, r and s have yielded two more confirmations of general relativity. These results have ruled out large classes of other gravitational theories which, without these binary pulsar systems, would be extremely difficult

10) Historically, this parameter has been denoted by the same letter as the parameter of the parameterized post-Newtonian formalism measuring the deflection of light [Will, 1993]. It may be a little confusing so that the reader should stay on alert to avoid the confusion.

to test. Other, more impressive tests of general relativity are expected from timing of the double pulsar PSR J0737-3039 which admits four independent confirmations [Kramer and Wex, 2009].

Measurement of the "range" and "shape" PPK parameters r and s is particularly important for testing general relativity in the binary pulsars being visible nearly edge-wise [Doroshenko and Kopeikin, 1995; van Straten *et al.*, 2001]. Taking into account the particular nature of the double pulsar PSR J0737-3039, which orbital plane coincides with observer's line of sight, it looks reasonable to analyze the influence of the orbital motion of pulsar and its companion on the r and s PPK parameters. Such analysis will be provided in this section by making use of expression (7.166). A binary pulsar consists of two bodies - the pulsar (subindex "p") and its companion (subindex "c"). The emission of a radio pulse takes place very near to the surface of the pulsar and, according to equation (7.147) and the related discussion, one can approximate $\boldsymbol{X} = X\boldsymbol{k}$, where X is the distance from the center of mass of the pulsar to the pulse-emitting point. In this approximation one gets $\boldsymbol{\rho}_{0p} = X\boldsymbol{k}$, and, as a consequence,

$$\ln\left[\rho_{0p} + \boldsymbol{k}\cdot\boldsymbol{\rho}_{0p} - \boldsymbol{\beta}_p\cdot\boldsymbol{\rho}_{0p} + (\boldsymbol{k}\cdot\boldsymbol{\rho}_{0p})(\boldsymbol{k}\cdot\boldsymbol{\beta}_p)\right] \quad = \quad \ln(2X)\ . \tag{7.169}$$

The distance X is practically constant. Possible slow drift of the emission point in vertical or horizontal directions can be absorbed by fitting to the pulsar's rotational phase. Hence, one can not observe this effect. Accounting for equation (7.169), formula (7.166) for the gravitational time delay can be displayed in the form

$$\begin{aligned}\Delta(t,t_0) \quad = \quad & -\frac{2Gm_c}{c^3}\Big\{(1-\boldsymbol{k}\cdot\boldsymbol{\beta}_c)\ln\left[\rho_{0c} + \boldsymbol{k}\cdot\boldsymbol{\rho}_{0c} - \boldsymbol{\beta}_c\cdot\boldsymbol{\rho}_{0c} + (\boldsymbol{k}\cdot\boldsymbol{\rho}_{0c})(\boldsymbol{k}\cdot\boldsymbol{\beta}_c)\right] + \boldsymbol{k}\cdot\boldsymbol{\beta}_c\Big\} \\ & -\frac{2Gm_p}{c^3}\left[(1-\boldsymbol{k}\cdot\boldsymbol{\beta}_p)\ln(2X) + \boldsymbol{k}\cdot\boldsymbol{\beta}_p\right]\ ,\end{aligned} \tag{7.170}$$

where one has omitted the residual terms because they are negligibly small. Kopeikin [1994] have proved that any constant term multiplied by the dot product $\boldsymbol{k}\cdot\boldsymbol{\beta}_p$ and/or $\boldsymbol{k}\cdot\boldsymbol{\beta}_c$ is absorbed into the epoch of the first pulsar's passage through the periastron. Thus, the conclusion is that all terms relating to the pulsar along with the very last term in the curl brackets in formula (7.170) are not directly observable. For this reason, they are omitted in the following text since the logarithmic contribution caused by the pulsar's companion is dominant. According to formula (7.147) one has

$$\boldsymbol{\rho}_{0c} = \boldsymbol{r} + X\boldsymbol{k} \qquad , \qquad \rho_{0c} = r + \frac{X}{r}\boldsymbol{k}\cdot\boldsymbol{r} + \ldots\ , \tag{7.171}$$

where $\boldsymbol{r} = \boldsymbol{x}_p(t_0) - \boldsymbol{x}_c(t_0)$ is a vector of relative position of the pulsar with respect to its companion, $r = |\boldsymbol{r}|$, and ellipsis denote residual terms of higher order. Taking into account these equations and omitting the unobservable terms allows us to reduce the gravitational time delay to the form

$$\Delta(t,t_0) \quad = \quad -\frac{2Gm_c}{c^3}(1-\boldsymbol{k}\cdot\boldsymbol{\beta}_c)\ln\left[\left((1+\frac{X}{r}\right)(r+\boldsymbol{k}\cdot\boldsymbol{r}) - \boldsymbol{\beta}_c\cdot\boldsymbol{r} + (\boldsymbol{k}\cdot\boldsymbol{r})(\boldsymbol{k}\cdot\boldsymbol{\beta}_c)\right] \tag{7.172}$$

The ratio $X/r \simeq 10^{-3} \div 10^{-4}$ is usually small for most of binary pulsars and can be neglected. Then, the time delay (7.172) can be decomposed into three terms

$$\Delta(t,t_0) = -\frac{2Gm_c}{c^3}\ln(r+\boldsymbol{k}\cdot\boldsymbol{r}) + \frac{2Gm_c}{c^3}(\boldsymbol{k}\cdot\boldsymbol{\beta}_c)\ln(r+\boldsymbol{k}\cdot\boldsymbol{r}) + \frac{2Gm_c}{c^3}\frac{\boldsymbol{\beta}_c\cdot\boldsymbol{r}-(\boldsymbol{k}\cdot\boldsymbol{r})(\boldsymbol{k}\cdot\boldsymbol{\beta}_c)}{r+\boldsymbol{k}\cdot\boldsymbol{r}}\,. \tag{7.173}$$

The first term on the right hand side of equation (7.173) is the standard expression for the Shapiro time delay in binary pulsars [Damour and Deruelle, 1986]. The second and the third terms on the right hand side of equation (7.173) were discovered by Wex [1995] under the assumption of uniform and rectilinear motion of pulsar and its companion in the expression for the post-Newtonian metric tensor of the binary system. One understands now that this assumption was equivalent to taking into account primary terms of retardation effects in propagation of gravitational field from pulsar's companion to the radio pulse [Kopeikin, 2003a]. This approximation works fairly well only for terms being linear with respect to velocities of the stars. Currently, this approximation is fully sufficient for experimental purposes.

In what follows only the case of the elliptic motion of the pulsar with respect to its companion is of importance. We apply it to elaborate on equation (7.172) where, in the case of the nearly edgewise orbits, the magnitude of $r + \boldsymbol{k}\cdot\boldsymbol{r}$ term can be pretty small near the event of the superior conjunction of the pulsar and its companion, thus, leading to significant amplification of the gravitational delay of pulsar's signal. The size and the shape of an elliptic orbit of the pulsar with respect to its companion are characterized by the semi-major axis a_R and the eccentricity e $(0 \le e < 1)$. The orientation in space of the plane of the pulsar's orbit is defined with respect to the plane of the sky by the inclination angle i and the longitude of the ascending node Ω. For orientation of the pulsar's position in the orbital plane one uses the argument of the pericenter ω. More precisely, the orientation of the orbit is defined by three unit vectors $(\boldsymbol{l},\boldsymbol{m},\boldsymbol{n})$ having coordinates defined in equations (1.109)–(1.111) as follows

$$\begin{aligned}\boldsymbol{l} &= (\cos\Omega, \sin\Omega, 0)\,,\\ \boldsymbol{m} &= (-\cos i\sin\Omega, \cos i\cos\Omega, \sin i)\,,\\ \boldsymbol{n} &= (\sin i\sin\Omega, -\sin i\cos\Omega, \cos i)\,.\end{aligned} \tag{7.174}$$

In this coordinate system one has the unit vector $\boldsymbol{k}$ defining direction of the propagation of light to be equal to $\boldsymbol{k} = -\boldsymbol{K} = (0,0,-1)$. One also neglects the proper motion of the pulsar in the sky which brings about the small secular change in coordinates of the vector $\boldsymbol{k}$ but the effect is negligible in practical observations. The error of the approximation is about $(Gm_c/c^3)(\mu T_{span})/(1-\sin i)$, where μ is the proper motion of the pulsar and T_{span} is the total time span of observation. This error is much smaller than 100 nanoseconds - the current precision of pulsar timing.

The coordinates of the pulsar in the orbital plane are given by the radius-vector $\boldsymbol{r}$ and the angle of the true anomaly f. In terms of $\boldsymbol{r}$ and f one has, according to equation (1.143)

$$\boldsymbol{r} = r(\boldsymbol{P}\cos f + \boldsymbol{Q}\sin f)\,, \tag{7.175}$$

where the unit vectors $\boldsymbol{P}$, $\boldsymbol{Q}$ are defined by

$$\boldsymbol{P} = \boldsymbol{l}\cos\omega + \boldsymbol{m}\sin\omega \quad , \quad \boldsymbol{Q} = -\boldsymbol{l}\sin\omega + \boldsymbol{m}\cos\omega\ . \tag{7.176}$$

The coordinate velocity of the pulsar's companion is given by

$$\boldsymbol{\beta}_c = -\frac{m_p}{M}\frac{\dot{\boldsymbol{r}}}{c}\ , \tag{7.177}$$

$$\dot{\boldsymbol{r}} = \left(\frac{GM}{p}\right)^{1/2}\left[-\boldsymbol{P}\sin f + \boldsymbol{Q}(\cos f + e)\right]\ , \tag{7.178}$$

where $M = m_p + m_c$; $p = a_R(1-e^2)^{1/2}$ is the focal parameter of the elliptic orbit, and m_p and m_c are the masses of the pulsar and its companion. Accounting for relationships

$$r = a_R(1 - e\cos u)\ , \qquad r\cos f = a_R(\cos u - e)\ , \qquad r\sin f = a_R(1-e^2)^{1/2}\sin u\ , \tag{7.179}$$

where u is the eccentric anomaly relating to the time of emission, $t_0 \equiv T$, and the moment of the first passage of the pulsar through the periastron, T_0, by the Kepler transcendental equation

$$u - e\sin u = n_b(T - T_0)\ , \tag{7.180}$$

one obtains

$$\boldsymbol{k}\cdot\boldsymbol{r} = -a_R\sin i\left[(\cos u - e)\sin\omega + (1-e^2)^{1/2}\cos\omega\sin u\right]\ , \tag{7.181}$$

$$\boldsymbol{r}\cdot\boldsymbol{\beta}_c = -c^{-1}a_c\, a_R\, n_b\, e\sin u\ , \tag{7.182}$$

$$\boldsymbol{k}\cdot\boldsymbol{\beta}_c = c^{-1}a_c\, n_b\,(1-e^2)^{-1/2}\sin i\left[e\cos\omega + \frac{(\cos u - e)\cos\omega - (1-e^2)^{1/2}\sin\omega\sin u}{1 - e\cos u}\right]\ . \tag{7.183}$$

Here $a_c = a_R m_p/M$, and $n_b = (GM/a_R^3)^{1/2}$ is the orbital frequency related to the orbital period P_b by the equation $n_b = 2\pi/P_b$.

Ignoring all constant factors, the set of equations given in this section allows us to casts the gravitational time delay (7.172) in the form [Kopeikin, 2003a]

$$\Delta(t, t_0) = -2r\ln\left\{1 - e\cos(u+\varepsilon) - s\left[\sin\omega(\cos u - e) + \sqrt{1-e^2}\cos\omega\sin u\right]\right\}, \tag{7.184}$$

where e is the orbital eccentricity of the binary system, ω is the argument of the periastron, and u is the eccentric anomaly relating to the time of emission, $t_0 = T$, and the instant of the first passage of the pulsar through the periastron, T_0, by the Kepler transcendental equation (7.180). The other parameters entering equation (7.184) are

$$\varepsilon = \frac{2\pi}{\sin i}\frac{x}{P_b}\frac{m_p}{m_c}\ , \tag{7.185}$$

$$r = \frac{Gm_c}{c^3}\left[1 - \frac{\varepsilon\sin i}{\sqrt{1-e^2}}F(u)\right], \tag{7.186}$$

$$s = \sin i\left[1 + \frac{\varepsilon\sin i}{\sqrt{1-e^2}}F(u)\right], \tag{7.187}$$

where m_p and m_c are masses of the pulsar and its companion respectively, $x = a \sin i/c$ is the orbital semimajor axis of pulsar's orbit projected on the line of sight and measured in seconds, and function

$$F(u) = e\cos\omega + \frac{(\cos u - e)\cos\omega - \sqrt{1-e^2}\sin\omega\sin u}{1 - e\cos u}\,. \tag{7.188}$$

In the case of a nearly circular orbit, when $e \simeq 0$, equation (7.184) is simplified

$$\Delta(t, t_0) = -\frac{2Gm_c}{c^3}(1 - \varepsilon\sin i\cos\phi)\ln\left(1 - \sin i\sin\phi - \frac{1}{2}\varepsilon\sin^2 i\sin 2\phi\right)\,, \tag{7.189}$$

where $\phi = u+\omega$ is the orbital phase. One notices that the argument of the logarithmic function is modulated by the term having a double orbital frequency.

The magnitude of the velocity-dependent terms in the argument of the logarithm in equation (7.184) is of the order $10^{-3} \div 10^{-4}$ as compared with the main term. In equations (7.185)-(7.187) parameter ε is the new (constant) relativistic parameter and r and s are the range and shape parameters of the gravitational time delay. It is worth noting that both of these parameters are no longer constant but depend of the orbital phase. This distorts the amplitude and shape of the logarithmic curve especially near the time of the superior conjunction of pulsar with its companion - the effect which may be measurable in future timing observations when better precision and time resolution will be achieved [Hobbs *et al.*, 2009]. Measurement of this distortion allows to make a judgment about the strength of gravitomagnetic orbital effects in general theory of relativity [Kopeikin and Fomalont, 2007]. It is worth emphasizing that the magnitude of the new timing effects described in this section can be large enough only in the binary systems where the mass and orbital velocity of pulsar's companion are reasonably large. The binary can be visible edge-on but if pulsar's companion is a low-mass star the detection of the gravitomagnetic effect in the time delay is questionable. It is worth noting that existence of the, so-called, bending time delay [Doroshenko and Kopeikin, 1995] affects the times of arrival of radio pulsars in a similar way as predicted by formula (7.184) making observation of the relativistic effect caused by the velocity-dependent terms in the gravitational time delay, a more difficult problem [Kramer *et al.*, 2006].

7.6.2
Moving gravitational lenses

Theoretical study of astrophysical phenomena caused by a moving gravitational lens certainly deserves a special attention [Kopeikin and Schäfer, 1999]. Though the effects produced by the motion of the lens may be difficult to measure, they can give us additional valuable information on the lensing parameters [Bonvin, 2008; Frittelli, 2003a,b; Sereno, 2008]. In particular, a lensing object moving across the line of sight should cause a red-shift difference between the multiple images of a background object like a quasar lensed by a galaxy, and a brightness anisotropy in the microwave background radiation [Birkinshaw, 1989]. Moreover, velocity-dependent terms entering equation of gravitational lens along with the proper motion of the deflector

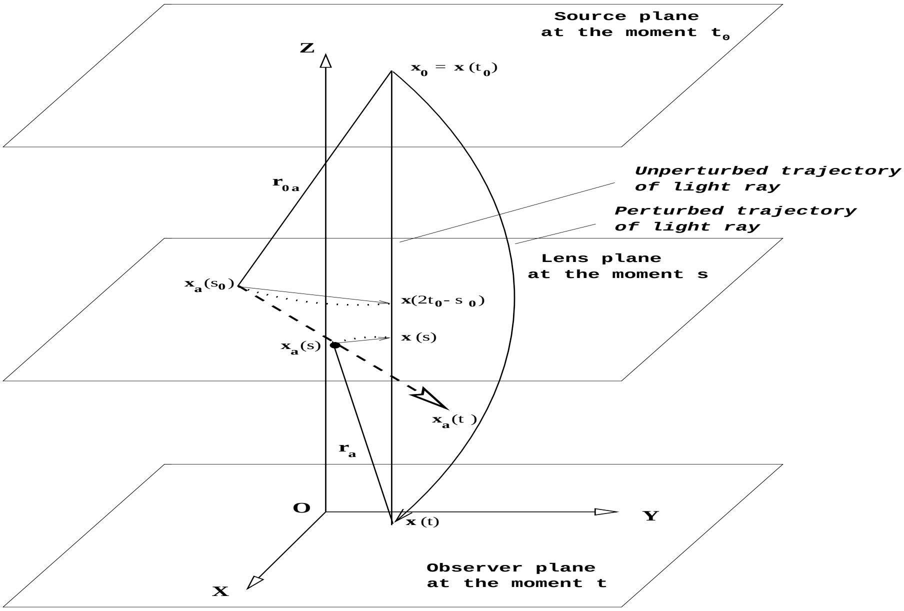

Figure 7.5 Relative configuration of observer, source of light and a moving gravitational lens deflecting a light ray which is emitted at the moment t_0 at the point $\boldsymbol{x}_0$, and is received at the moment t at the point $\boldsymbol{x}$, where the observer is located. The lens moves along a straight line with constant velocity starting from the retarded position $\boldsymbol{x}_a(s_0)$ through that $\boldsymbol{x}_a(s)$ and riches the point $\boldsymbol{x}_a(t)$ at the moment of observation. The process occupies the time of propagation of the photon from the point of its emission up to the point of its observation.

can distort the shape and the amplitude of magnification curve observed in the microlensing event. Slowly moving gravitational lenses are conventional astrophysical objects [Schneider *et al.*, 1992] but the effects caused by their motion are small and have not yet been detected. However, a hypothetical cosmic string, for example, moves with the speed of light and may produce a noticeable observable effect if it has sufficiently large mass per unit length. It is hopeful that gradually increasing precision of spectral and photometric astronomical observations will make it possible to measure all these and other possible effects in a foreseeable future [Zakharov and Sazhin, 1998].

7.6.2.1 *Gravitational Lens Equation*

This section derives the equation of a moving gravitational lens for the case of constant velocity $\boldsymbol{v}_a$ of the a-th light-ray-deflecting mass but without any other restrictions on its magnitude besides $v_a < c$. This assumption simplifies calculations of all required integrals allowing to bring them to a manageable form. Geometry of the lens is shown in Figure 7.5.

In what follows, it is convenient to introduce two vectors $\boldsymbol{\varsigma}_a = \boldsymbol{x}(s) - \boldsymbol{x}_a(s)$ and $\boldsymbol{\varsigma}_{0a} = \boldsymbol{x}(2t_0 - s_0) - \boldsymbol{x}_a(s_0)$. One will also suppose that the length of vector $\boldsymbol{\varsigma}_a$ is small compared to any of the distances: $R = |\boldsymbol{x} - \boldsymbol{x}_0|$, $r_a = |\boldsymbol{x} - \boldsymbol{x}_a(s)|$, or $r_{0a} = |\boldsymbol{x}_0 - \boldsymbol{x}_a(s_0)|$. It is not difficult to prove by straightforward calculations with taking into account the gravity null-cone equations (7.13), (7.75), that

$$\boldsymbol{\varsigma}_a = \boldsymbol{r}_a - \boldsymbol{k} r_a\,, \qquad \boldsymbol{\varsigma}_{0a} = \boldsymbol{r}_{0a} + \boldsymbol{k} r_{0a}\,, \tag{7.190}$$

where, as in the other parts of the present book, one has $\boldsymbol{r}_a = \boldsymbol{x} - \boldsymbol{x}_a(s)$ and $\boldsymbol{r}_{0a} = \boldsymbol{x}_0 - \boldsymbol{x}_a(s_0)$, and the unit vector $\boldsymbol{k}$ is directed along the unperturbed path of propagation of the light ray. From these equalities it follows that approximately

$$\boldsymbol{k}\cdot\boldsymbol{\varsigma}_a = -\frac{d_a^2}{2r_a}\,, \qquad \boldsymbol{k}\cdot\boldsymbol{\varsigma}_{0a} = \frac{d_{0a}^2}{2r_{0a}}\,, \tag{7.191}$$

and

$$r_{0a} - \boldsymbol{k}\cdot\boldsymbol{r}_{0a} = 2r_{0a} - \frac{d_{0a}^2}{2r_{0a}}\,, \tag{7.192}$$

where distances $d_a = |\boldsymbol{\varsigma}_a|$ and $d_{0a} = |\boldsymbol{\varsigma}_{0a}|$ are the Euclidean lengths of the corresponding vectors, which are small due to the initial assumption. One can see that making use of the relationships (7.190) yields

$$r_a - \boldsymbol{\beta}_a\cdot\boldsymbol{r}_a = r_a(1 - \boldsymbol{k}\cdot\boldsymbol{\beta}_a) - \boldsymbol{\varsigma}_a\cdot\boldsymbol{\beta}_a = r_a(1 - \boldsymbol{k}\cdot\boldsymbol{\beta}_a) + O(\beta_a d_a)\,, \tag{7.193}$$

where $\boldsymbol{\beta}_a \equiv \boldsymbol{v}_a/c$, and the residual terms can be neglected because of their smallness compared to the leading one.

It is worth noting that vector $\boldsymbol{\varsigma}_a$ is approximately equal to the impact parameter of the light ray trajectory measured with respect to the position of the deflector taken at the retarded time $s = t - r_a/c$. Indeed, let us introduce vectors $\xi^i = P^i{}_j x^j$ and $\xi^i_a = P^i{}_j x^j_a(s)$ which are lying in the plane that is orthogonal to the unperturbed trajectory of the light ray, that is $\boldsymbol{\xi}\cdot\boldsymbol{k} = 0$. Then, from definitions (7.190), (7.191) one immediately derives an exact relationship

$$\boldsymbol{\xi} - \boldsymbol{\xi}_a = \boldsymbol{\varsigma}_a + \boldsymbol{k}\,\frac{d_a^2}{2r_a}\,, \tag{7.194}$$

from which it follows that

$$\boldsymbol{\varsigma}_a = \boldsymbol{\xi} - \boldsymbol{\xi}_a - \boldsymbol{k}\,\frac{d_a^2}{2r_a}\,. \tag{7.195}$$

Similar relationships can be derived for $\boldsymbol{\varsigma}_{0a}$ in the same way. It is worthwhile to note that

$$P^i_j r^j_a = P^i_j \varsigma^j_a = \xi^i - \xi^i_a\,, \tag{7.196}$$

and

$$r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a = \frac{d_a^2}{2r_a} = \frac{|\boldsymbol{\xi} - \boldsymbol{\xi}_a|^2}{2r_a} + \frac{d_a^4}{8r_a^3} + \dots\,. \tag{7.197}$$

Let us denote the total angle of light deflection caused by the a-th body as (compare with the exact equation (7.113))

$$\alpha_a^i(\tau) = \frac{4Gm_a}{c^2}\frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a}{\sqrt{1-\beta_a^2}}\frac{\xi^i-\xi_a^i}{|\boldsymbol{\xi}-\boldsymbol{\xi}_a|^2}\,. \tag{7.198}$$

Thus, from the formulas (7.119)-(7.121) one obtains [11] for vectors α^i, $\daleth^i$, $\varkappa^i$ introduced earlier in equations (7.113), (7.68) and (7.122)

$$\alpha^i(\tau) = \sum_{a=1}^{N}\alpha_a^i(\tau) + O\left(\frac{Gm_a}{c^2r_a}\beta_a\right) + O\left(\frac{Gm_a}{c^2r_a}\frac{d_a}{r_a}\right)\,, \tag{7.199}$$

$$\begin{aligned}\daleth^i(\tau) &= -\frac{1}{R}\sum_{a=1}^{N}r_a\alpha_a^i(\tau) - \frac{2G}{c^2R}\sum_{a=1}^{N}\frac{m_a\beta_{aT}^i}{\sqrt{1-\beta_a^2}}\ln\left(\frac{|\boldsymbol{\xi}-\boldsymbol{\xi}_a|^2}{2r_a}\right)\\ &+ O\left(\frac{Gm_a}{c^2r_a}\beta_a\right) + O\left(\frac{Gm_a}{c^2r_a}\frac{d_a}{r_a}\right)\,,\end{aligned} \tag{7.200}$$

$$\daleth^i(\tau_0) = -\frac{2G}{c^2R}\sum_{a=1}^{N}\frac{m_a\beta_{aT}^i}{\sqrt{1-v_a^2}}\ln(2r_{0a}) + O\left(\frac{Gm_a}{c^2r_a}\beta_a\right) + O\left(\frac{Gm_a}{c^2r_a}\frac{d_a}{r_a}\right)\,, \tag{7.201}$$

$$\varkappa^i(\tau) = O\left(\frac{Gm_a}{c^2r_a}\beta_a^2\right)\,, \tag{7.202}$$

where by definition $\beta_{aT}^i \equiv v_{aT}^i/c$ and the transverse velocity $v_{aT}^i = P^i{}_j v_a^j$ is a projection of the velocity of the a-th body onto the plane being orthogonal to the unperturbed light trajectory.

Let us assume that the lens consists of a gravitationally bounded system of N bodies with mass m_a of each body. In the most simple case one considers only one body. The direction of the perturbed light trajectory at the point of observation is determined by the unit vector $\boldsymbol{s}$ according to equation (7.109). One usee that definition to draw a straight line originating from the point of observation and directed along the vector $\boldsymbol{s}$ up to the point of its intersection with the lens plane as shown in Figure 7.6. The line is parameterized by a running parameter λ and its equation is given by

$$x^i(\lambda) = x^i(t) + cs^i(\lambda - t)\,, \tag{7.203}$$

where t is the the moment of observation, and $x^i(t)$ are the spatial coordinates of the point of observation. Geometrically, the coordinates of the point $x^i(\lambda)$ at the instant of time λ^* when the line (7.203) intersects the lens plane are determined by as a sum of three vectors

$$x^i(\lambda^*) = X^i(\lambda^*) + \eta^i - \xi_L^i\,, \tag{7.204}$$

where $\eta^i = P^i_j\, x^i(\lambda^*)$ is the perturbed value of the impact parameter ξ^i caused by the influence of the combined gravitational fields of the (micro) lenses m_a, $X^i(\lambda^*) =$

11) Notice that the integrals in equations (7.119)-(7.121) are identically zero because of the assumption that the velocities of the gravitating bodies are constant.

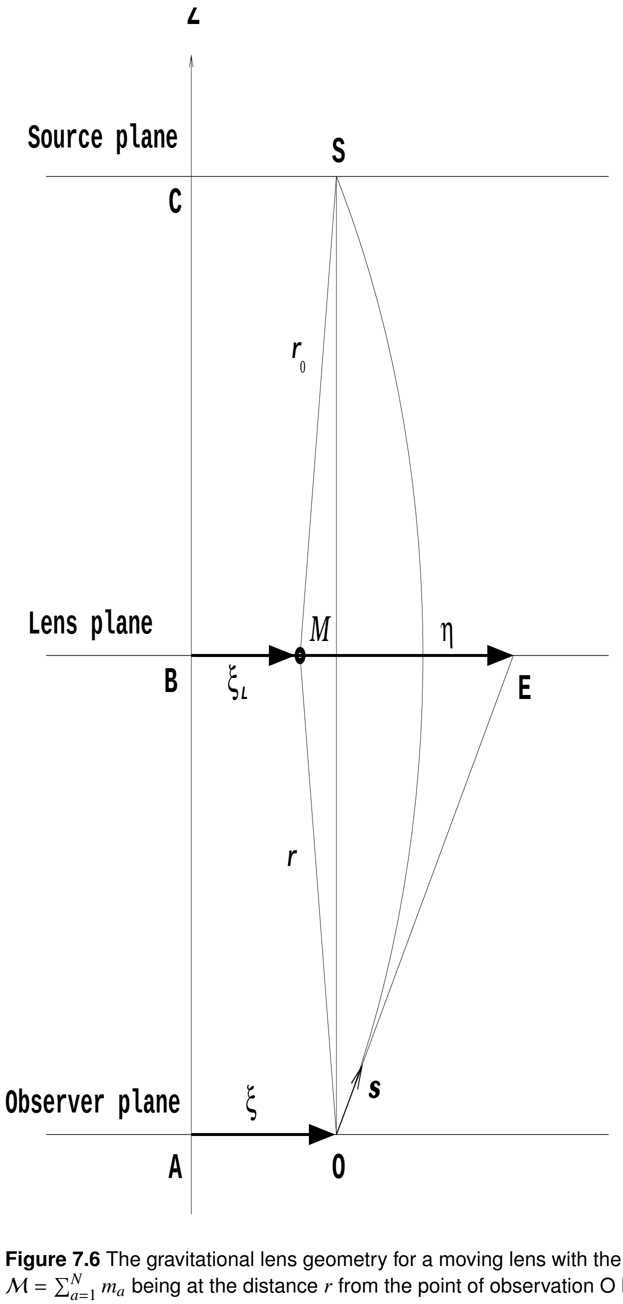

Figure 7.6 The gravitational lens geometry for a moving lens with the overall mass $\mathcal{M} = \sum_{a=1}^{N} m_a$ being at the distance r from the point of observation O having coordinates $x^i(t)$. A source of light S with coordinates $x_0(t_0)$ is at the distance R from the point O. Vector $\boldsymbol{\xi}$ is the impact parameter of the unperturbed path of photon in the observer plane with respect to the origin of coordinates. Vector $\boldsymbol{\xi}_L$ denotes position of the center of mass of the lensing objects in the lens plane. Vector $\boldsymbol{\eta} = \mathbf{BE}$ is the observed position of the image of a background source of light S shifted in the lens plane from its true position by the gravitational field of the lens to the point E. Coordinates of the center of mass of the lens are $X^i(\lambda^*) = \mathcal{M}^{-1} \sum_{a=1}^{N} m_a x_a^i(\lambda^*)$, and coordinates of the point E are $x^i(\lambda^*) = x^i(t) + s^i(\lambda^* - t)$.

$\mathcal{M}^{-1}\sum_{a=1}^{N} m_a x_a^i(\lambda^*)$ are coordinates of the center of mass of the lens at the moment λ^*. When the line (7.203) intersects the lens plane the numerical value of λ up to corrections of order $O(d/r)$ is equal to that of the retarded time s defined by equation like (7.13) in which r_a is replaced by r - the distance from observer to the lens. It means that at the lens plane $\lambda^* - t \simeq -r/c$. Taking the value of the parameter $\lambda = \lambda^*$, and applying the operator of projection P_{ij} to the equation (7.203), one obtains

$$\eta^i \quad = \quad \xi^i - \left[\alpha^i(\tau) + \beth^i(\tau) - \beth^i(\tau_0) + \varkappa^i(\tau)\right] r\,. \tag{7.205}$$

Finally, making use of the relationships (7.199)-(7.201) and expanding distances r_a, r_{0a} around the values r, r_0 respectively (see Figure 7.6 for explanation of the meaning of these distances), the equation of gravitational lens in vectorial notations reads as follows

$$\boldsymbol{\eta} \quad = \quad \boldsymbol{\xi} - \frac{r\, r_0}{R}\,\boldsymbol{\alpha}(\boldsymbol{\xi}) + \frac{r}{R}\,\boldsymbol{\kappa}(\boldsymbol{\xi})\,, \tag{7.206}$$

where

$$\boldsymbol{\alpha}(\boldsymbol{\xi}) \quad = \quad \frac{4G}{c^2}\sum_{a=1}^{N} m_a \frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a}{\sqrt{1-\beta_a^2}}\,\frac{\xi^i - \xi_a^i}{|\boldsymbol{\xi}-\boldsymbol{\xi}_a|^2}\,, \tag{7.207}$$

$$\boldsymbol{\kappa}(\boldsymbol{\xi}) \quad = \quad \frac{2G}{c^2}\sum_{a=1}^{N} \frac{m_a\beta_{aT}^i}{\sqrt{1-\beta_a^2}}\ln\left(\frac{|\boldsymbol{\xi}-\boldsymbol{\xi}_a|^2}{2r_a\, r_{0a}}\right). \tag{7.208}$$

It is not difficult to realize that the third term on the right hand side of the equation (7.206) is $(d_a/r_0)\beta_a$ times smaller than the second one. For this reason, one is allowed to neglect it and to represent the equation of gravitational lensing in its conventional form [Schneider *et al.*, 1992; Zakharov and Sazhin, 1998],

$$\boldsymbol{\eta} \quad = \quad \boldsymbol{\xi} - \frac{rr_0}{R}\,\boldsymbol{\alpha}(\boldsymbol{\xi})\,, \tag{7.209}$$

where the angle $\boldsymbol{\alpha}(\boldsymbol{\xi})$ is given by (7.207). It is worthwhile emphasizing that the assumption of constant velocities of particles $\boldsymbol{v}_a$ makes equation (7.209) valid under condition that the accelerations of the bodies have negligible contribution to the gravitational lens equation.

It is useful to compare the expression for the angle of deflection α^i given in equation (7.207) with that derived by Kopeikin *et al.* [1999]. In that paper one has considered different aspects of astrometric and timing effects of gravitational waves from localized sources. The gravitational field of the source was described in terms of static monopole, spin dipole, and time-dependent quadrupole moments. Time delay and the angle of light deflection α^i in case of gravitational lensing were obtained in the following form [Kopeikin *et al.*, 1999]

$$t - t_0 = \frac{1}{c}|\boldsymbol{x}-\boldsymbol{x}_0| - \frac{4}{c}\psi + \frac{2G\mathcal{M}}{c^3}\ln(4rr_0/d^2)\,, \qquad \alpha_i = 4\,\hat{\partial}_i\psi\,, \tag{7.210}$$

where d is a constant impact parameter, the partial ('projective') derivative reads $\hat{\partial}_i \equiv P^j{}_i\partial/\partial\xi^j$, and r and r_0 are constant distances from the lens to observer and the source of light respectively. The quantity ψ is the, so-called, gravitational lens potential [Schneider *et al.*, 1992; Zakharov and Sazhin, 1998] having the form [Kopeikin *et al.*, 1999]

$$\psi = \frac{G}{c^2}\left[\mathcal{M}+\frac{1}{c}\varepsilon_{jpq}k^p\mathcal{S}^q\hat{\partial}_j+\frac{1}{2}\,\mathcal{I}^{pq}(t^*)\,\hat{\partial}_{pq}\right]\ln|\boldsymbol{\xi}|\;, \tag{7.211}$$

and ε_{jpq} is the fully antisymmetric Levi-Civita symbol. The expression (7.211) includes the explicit dependence on the static mass $\mathcal{M}$, spin $\mathcal{S}^i$, and time-dependent quadrupole moment $\mathcal{I}^{ij}$ of the deflector taken at the moment t^* of the closest approach of the light ray to the origin of the coordinate system which was chosen at the center of mass of the deflector so that the dipole moment $\mathcal{I}^i$ of the system equals to zero identically. Equation (7.210) generalizes our result obtained independently in [Kopeikin, 1997c] to the case of the gravitational lens potential which is a function of time. In case of an isolated astronomical system of N bodies, the multipole moments are defined in the Newtonian approximation as follows

$$\mathcal{M}=\sum_{a=1}^{N}m_a\;,\qquad \mathcal{I}^i=\sum_{a=1}^{N}m_a x_a^i\;,\qquad \mathcal{S}^i=\sum_{a=1}^{N}m_a(\boldsymbol{x}_a\times\boldsymbol{v}_a)^i\;,\qquad \mathcal{I}_{ij}=\sum_{a=1}^{N}m_a\left(x_a^i x_a^j-\frac{1}{3}\boldsymbol{x}_a^2\,\delta^{ij}\right)\;, \tag{7.212}$$

where the symbol '×' denotes the usual Euclidean cross product between twovectors and, what is more important, coordinates and velocities of all bodies are taken at one and the same instant of time. The rest of this section stipulates that velocity of the light-ray-deflecting bodies are small and the origin of the coordinate frame is chosen at the barycenter of the gravitational lens. It means that

$$\mathcal{I}^i(t)=\sum_{a=1}^{N}m_a x_a^i(t)=0\;,\qquad \text{and}\qquad \dot{\mathcal{I}}^i(t)=\sum_{a=1}^{N}m_a v_a^i(t)=0\;. \tag{7.213}$$

Now it is worthwhile to note that coordinates of gravitating bodies in equation (7.207) are taken at different instants of the retarded time defined for each body a by the equation (7.13). In the case of gravitational lensing all the retarded times are close to the instant of the closest approach t^* and one is allowed to use the Taylor expansion of the quantity

$$\sum_{a=1}^{N}m_a x_a^i(s) = \sum_{a=1}^{N}m_a x_a^i(t^*)+\sum_{a=1}^{N}m_a v_a^i(t^*)(s-t^*)+O(s-t^*)^2\;. \tag{7.214}$$

Remembering that the retarded time s is defined by equation (7.13) and the moment of the closest approach is given by the relationship

$$t^* = t-\frac{1}{c}\boldsymbol{k}\cdot\boldsymbol{x}=t-\frac{1}{c}\boldsymbol{k}\cdot\boldsymbol{r}_a-\frac{1}{c}\boldsymbol{k}\cdot\boldsymbol{x}_a(s)\;, \tag{7.215}$$

one obtains after accounting for equation (7.197),

$$c\,(s-t^*) = \boldsymbol{k}\cdot\boldsymbol{x}_a(s)-\frac{d_a^2}{2r_a}\simeq\boldsymbol{k}\cdot\boldsymbol{x}_a(t^*)+O\left(\frac{d_a^2}{r_a}\right)+O\left(\beta_a x_a\right)\;. \tag{7.216}$$

Finally, one concludes that

$$\sum_{a=1}^{N} m_a x_a^i(s) \quad = \quad \sum_{a=1}^{N} m_a \beta_a^i(t^*)[\boldsymbol{k}\cdot\boldsymbol{x}_a(t^*)] + \ldots\,, \tag{7.217}$$

where ellipses denote terms of higher order of magnitude, and where the equation (7.213) has been used.

Let us assume that the impact parameter ξ^i is always larger than the distance ξ_a^i. Then, making use of the Taylor expansion of the right hand side of equation (7.207) with respect to ξ_a^i and $\beta_a = v_a/c$ one can prove that the deflection angle α^i is represented in the form

$$\alpha_i \quad = \quad 4\hat{\partial}_i\Psi\,, \tag{7.218}$$

where the potential Ψ is given as follows

$$\Psi \quad = \quad \frac{G}{c^2}\Bigg\{\sum_{a=1}^{N} m_a - \boldsymbol{k}\cdot\sum_{a=1}^{N} m_a\boldsymbol{\beta}_a(s) - \sum_{a=1}^{N} m_a x_a^j(s)\hat{\partial}_j + \tag{7.219}$$
$$\boldsymbol{k}\cdot\sum_{a=1}^{N} m_a\boldsymbol{\beta}_a(s)\, x_a^j(s)\,\hat{\partial}_j + \frac{1}{2}\sum_{a=1}^{N} m_a x_a^p(s)\, x_a^q(s)\,\hat{\partial}_{pq}\Bigg\}\ln|\boldsymbol{\xi}| + \ldots\,,$$

and ellipsis denote residual terms of the higher order of magnitude. Expanding all terms depending on the retarded time s in this formula with respect to the time t^*, noting that the second 'projective' derivative $\hat{\partial}_{pq}$ is traceless, and taking into account the relationship (7.217), the center-of-mass conditions (7.213), the definitions of multipole moments (7.212), and the vector equality

$$x_a^j(\boldsymbol{k}\cdot\boldsymbol{v}_a) - v_a^j(\boldsymbol{k}\cdot\boldsymbol{x}_a) \quad = \quad (\boldsymbol{k}\times(\boldsymbol{x}_a\times\boldsymbol{v}_a))^j\,, \tag{7.220}$$

one finds out that the gravitational lens potential is given by $\Psi = \psi$. Hence, under adopted assumptions the gravitational lens formalism based on the null-cone technique, gives the same result for the angle of deflection of light as it is shown in formulas (7.210), (7.211).

Notice that if one had supposed that the dipole moment of the lens $\mathcal{I}^i$ is not equal to zero, then the expression for the gravitational lens potential ψ would assume the form

$$\psi = \frac{G}{c^2}\left[\mathcal{M} - \frac{1}{c}\boldsymbol{k}\cdot\dot{\boldsymbol{\mathcal{I}}}(t^*) - \mathcal{I}^i(t^*)\,\hat{\partial}_i + \frac{1}{c}\,(\boldsymbol{k}\times\boldsymbol{\mathcal{S}})^i\,\hat{\partial}_i + \frac{1}{2}\,\mathcal{I}^{ij}(t^*)\,\hat{\partial}_{ij}\right]\ln|\boldsymbol{\xi}|\,, \tag{7.221}$$

where the impact parameter $\boldsymbol{\xi}$ is the distance from the origin of the coordinate system to the point of the closest approach of the light ray to the lens. The scrutiny examination of the multipole structure of the shape of the curves of constant value of the potential ψ in cosmological gravitational lenses [Bartelmann, 1998; Kaiser and Squires, 1993] may reveal the presence of dark matter in the lens and identify the position of its center of mass, velocity and density distribution which can be compared with the analogous characteristics of luminous matter in the lens. In case of the transparent gravitational lens the overall expression for the gravitational lens potential ψ

given in terms of the expansion with respect to the transverse-traceless (TT) internal and external multipole moments is given in the work [Kopeikin, 1997c]. Discussion of observational effects produced by the spin of the lens is given in other articles [Ciufolini *et al.*, 2003; Dymnikova, 1986; Kopeikin and Mashhoon, 2002].

7.6.2.2 ***Gravitational Shift of Frequency by Moving Bodies***

One stipulates that velocity $\boldsymbol{v}_a$ of each body comprising the lens is almost constant so that one can neglect the bodies' accelerations. One also assumes that the coordinate velocity of the source of light is $\boldsymbol{v}_0 = d\boldsymbol{x}_0/dt_0$, and that of observer is $\boldsymbol{v} = d\boldsymbol{x}/dt$. Calculation of the gravitational shift of frequency by an ensemble of moving bodies is performed by making use of a general equation (7.123). Let us further assume that both the source of light and observer are located very far away from the moving masses so that one can neglect the impact of the gravitational field on their motion. Under these circumstances, derivative of the proper time of the source of light, $\mathcal{T}_0$, with respect to the coordinate time t_0 at the point of emission of light is

$$\frac{d\mathcal{T}_0}{dt_0} = \sqrt{1-\beta_0^2}\,, \tag{7.222}$$

where $\beta_0 = v_0/c$. Derivative of the proper time of the observer, $\mathcal{T}$, with respect to the coordinate time t at the point of observation

$$\frac{dt}{d\mathcal{T}} = \frac{1}{\sqrt{1-\beta^2}}\,, \tag{7.223}$$

where $\beta = v/c$. Accounting for equation (7.136) in equation (7.129) one obtains that a derivative of the time of emission with respect to the time of observation

$$\frac{dt_0}{dt} = \frac{1+\boldsymbol{K}\cdot\boldsymbol{\beta} - \frac{2G}{c^2}\sum_{a=1}^{N} m_a\left[\frac{\partial s}{\partial t}\frac{\partial}{\partial s} + \frac{\partial t^*}{\partial t}\frac{\partial}{\partial t^*} + \frac{\partial k^i}{\partial t}\frac{\partial}{\partial k^i}\right]B_a(s,s_0,t^*,\boldsymbol{k})}{1+\boldsymbol{K}\cdot\boldsymbol{\beta}_0 + \frac{2G}{c^2}\sum_{a=1}^{N} m_a\left[\frac{\partial s}{\partial t_0}\frac{\partial}{\partial s} + \frac{\partial s_0}{\partial t_0}\frac{\partial}{\partial s_0} + \frac{\partial t^*}{\partial t_0}\frac{\partial}{\partial t^*} + \frac{\partial k^i}{\partial t_0}\frac{\partial}{\partial k^i}\right]B_a(s,s_0,t^*,\boldsymbol{k})}\,. \tag{7.224}$$

After taking the partial derivatives with the help of relationships (7.133)-(7.143), using expansions (7.192), (7.193), (7.197), neglecting small terms of the order of d_a/r_a, Gm_a/c^2r_a, Gm_a/c^2r_{0a}, and reducing similar terms, one gets

$$1+z = \left(\frac{1-\beta_0^2}{1-\beta^2}\right)^{1/2} \frac{1+(\boldsymbol{K}+\beth-\beth_0)\cdot\boldsymbol{\beta} + \frac{4G}{c^2}\sum_{a=1}^{N} m_a \frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a}{\sqrt{1-\beta_a^2}}\frac{(\boldsymbol{k}\times\boldsymbol{\beta})\cdot(\boldsymbol{k}\times\boldsymbol{r}_a)}{|\boldsymbol{\xi}-\boldsymbol{\xi}_a|^2}}{1+(\boldsymbol{K}+\beth-\beth_0)\cdot\boldsymbol{\beta}_0 + \frac{4G}{c^2}\sum_{a=1}^{N} m_a \frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_0}{\sqrt{1-\beta_a^2}}\frac{(\boldsymbol{k}\times\boldsymbol{\beta}_a)\cdot(\boldsymbol{k}\times\boldsymbol{r}_a)}{|\boldsymbol{\xi}-\boldsymbol{\xi}_a|^2}}\,, \tag{7.225}$$

where the relativistic corrections $\beth = \beth(\tau,\boldsymbol{\xi})$, $\beth_0 = \beth(\tau_0,\boldsymbol{\xi})$ are given by means of expressions (7.68), (7.69), (7.115) - (7.118). Making use of relationship (7.67) between the unit vectors $\boldsymbol{K}$ and $\boldsymbol{k}$, the previous formula can be displayed as follows

$$1+z=\left(\frac{1-\beta_0^2}{1-\beta^2}\right)^{1/2}\frac{1+\boldsymbol{k}\cdot\boldsymbol{\beta}+\dfrac{4G}{c^2}\displaystyle\sum_{a=1}^{N}m_a\frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a}{\sqrt{1-\beta_a^2}}\frac{(\boldsymbol{k}\times\boldsymbol{\beta})\cdot(\boldsymbol{k}\times\boldsymbol{r}_a)}{|\boldsymbol{\xi}-\boldsymbol{\xi}_a|^2}}{1+\boldsymbol{k}\cdot\boldsymbol{\beta}_0+\dfrac{4G}{c^2}\displaystyle\sum_{a=1}^{N}m_a\frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_0}{\sqrt{1-\beta_a^2}}\frac{(\boldsymbol{k}\times\boldsymbol{\beta}_a)\cdot(\boldsymbol{k}\times\boldsymbol{r}_a)}{|\boldsymbol{\xi}-\boldsymbol{\xi}_a|^2}}\,. \tag{7.226}$$

This formula corresponds to the case when both the source of light and observer are located at so large distance from the ensemble of the moving bodies that the difference between directions of vectors $\boldsymbol{K}$ and $\boldsymbol{k}$ is negligible.

Formula (7.226) for the gravitational shift of frequency is gauge invariant with respect to small coordinate transformations in the first post-Minkowskian approximation which leave the coordinates harmonic and asymptotically Minkowski. Moreover, formula (7.226) is invariant with respect to the Lorentz transformations and can be applied for arbitrary large velocities of observer, source of light, and that of a moving body. In case of the slow motion of the source of light, observer, and the body equation (7.226) can be further simplified by expanding it with respect to the powers of small quantities v_0/c, v/c, and v_a/c. Neglecting terms of the order of v^4/c^4, v_0^4/c^4, $(Gm_a/c^2d_a)(v^2/c^2)$, etc., yields for the frequency shift

$$\begin{aligned}\frac{\delta\nu}{\nu_0} &= \boldsymbol{k}\cdot(\boldsymbol{\beta}_0-\boldsymbol{\beta})\left[1+\boldsymbol{k}\cdot\boldsymbol{\beta}_0+(\boldsymbol{k}\cdot\boldsymbol{\beta}_0)^2-\frac{\beta_0^2}{2}+\frac{\beta^2}{2}\right]-\frac{\beta_0^2}{2}+\frac{\beta^2}{2} \\ &\quad+4\sum_{a=1}^{N}\frac{Gm_a}{c^2}\frac{[\boldsymbol{k}\times(\boldsymbol{\beta}-\boldsymbol{\beta}_a)]\cdot(\boldsymbol{k}\times\boldsymbol{r}_a)}{|\boldsymbol{\xi}-\boldsymbol{\xi}_a|^2}\,,\end{aligned} \tag{7.227}$$

where $\delta\nu=\nu-\nu_0$. The terms in the right side of this formula that depend only on the velocities of the source of light and observer are the part of the special relativistic Doppler shift of frequency caused by motion of the observer and the source of light. The last term in the right side of equation (7.227) describes gravitational shift of frequency caused by the relative motion of the body a with respect to observer. One emphasizes that in the linear with respect to $\boldsymbol{v}_a/c$ approximation, gravitational shift of frequency depends only on the transverse component of the relative motion of the massive body and observer. Dependence of the gravitational shift of frequency on the longitudinal motion of the body (its radial velocity) appears if and only if one takes into account quadratic and higher order powers in $\boldsymbol{v}_a/c$. Equation (7.227) shows that the gravitational shift of frequency appears only if there is a relative transverse velocity of the massive body with respect to observer which makes the impact parameter of the light ray with respect to the body, $\boldsymbol{\xi}-\boldsymbol{\xi}_a$, dependent on time.

Let us drop for a while the special relativistic terms out of equation (7.227) and discuss the case of an ensemble of moving massive bodies comprising a non-stationary gravitational lens. One will assume that the projective distance $\boldsymbol{\xi}_a$ of each body from the center of mass of the lens is small compared with the impact parameter $\boldsymbol{\xi}$ of the light ray. By expanding the last term in the expression (7.227) with respect to the powers of the small quantity, ξ/ξ_a, the gravitational shift of frequency can be written down in the following form

$$\left(\frac{\delta\nu}{\nu_0}\right)_{gr} = \frac{4}{c}\frac{\partial\psi}{\partial s} + \boldsymbol{\beta}\cdot\boldsymbol{\alpha}(\boldsymbol{\xi}) , \tag{7.228}$$

where the deflection angle $\boldsymbol{\alpha} = 4\hat{\partial}_i\psi$ is displayed in (7.210), and

$$\psi = \frac{G}{c^2}\sum_{a=1}^{N} m_a \ln|\boldsymbol{\xi} - \boldsymbol{\xi}_a| , \tag{7.229}$$

with $\boldsymbol{\xi}_a = \boldsymbol{\xi}_a(s)$ and $s = t - r_a/c$ being the retarded time (7.13). It is straightforward to prove that function ψ is exactly the gravitational lens potential given previously in equation (7.221), and that the retarded time $s = t^*$ in a very good approximation. It is also useful to notice that formula (7.228) can be derived as a direct consequence of equation (7.210) for the time delay of light in gravitational lensing [Kopeikin and Schäfer, 1999].

Simple relationship (7.228) can be compared with the result of previous calculations performed by Birkinshaw and Gull [1983] (see especially their equation 9). One has checked that the derivation of the corresponding formula for the gravitational shift of frequency given by these authors on the ground of a pure phenomenological approach and cited later on in paper [Birkinshaw, 1989], is consistent, at least, in the first order approximation with respect to the velocity of gravitational lens. Gurvits and Mitrofanov [1986] had derived similar formula but it is printed with a typo in the overall numerical factor.

Formula (7.228) shows that a photon passing the front side of the lens with respect to the direction of motion of the lens will see a weaker gravitational potential on its way into the lens than on the way out, and, hence, receives a net red shift. Similarly a photon on the other side will receive a net blue shift [Lewis and Challinor, 2006]. The moving lens, therefore, induces a dipole-like temperature anisotropy in the photons crossing the lens path. For a $10^{15}\ M_\odot$ mass cluster moving transverse to the line of sight at the speed 600 km/s the amplitude of the signal is 5×10^{-7}, corresponding to a temperature signal of 1 μK. Equation (7.228) also points out to a remarkable fact that the annual motion of observer with respect to the solar system barycenter should produce periodic changes in the observed spectra of images of the background sources (quasars) and in the measured anisotropy of CMBR. This is because the relative velocity $\boldsymbol{v} - \boldsymbol{v}_a$ of observer with respect to the lens experiences annual periodic variation due to observer's orbital motion. This effect produces small periodic variations in the temperature of the CMBR radiation in the sky in the directions being close to cosmological gravitational lenses. It will be technically challenging to observe these temporal variations of the temperature because of their smallness. For the above-given example of the cluster, they may reach value of $\delta T/T \simeq 10^{-8}$ - too small to be detectable by a space mission Planck [Heinämäki, 2006; Lawrence, 2009] but it may be considered for detection in future trans-Planckian missions. Numerical simulations of the CMBR anisotropy by a moving gravitational lens carried out in the paper [Aghanim *et al.*, 1998] on the premise of formula (7.228) under assumption $\boldsymbol{v} = 0$, confirm significance of the effect for future space experiments being designed for detection of the small scale temperature fluctuations of the CMBR.

Now, let us consider the case when the distance r between the lens and observer and that r_0 between the lens and the source of light are not large so that one can not neglect the difference between the unit vectors $\boldsymbol{K}$ and $\boldsymbol{k}$ as one did in derivation of equation (7.227). Vector $\boldsymbol{K}$ relates to $\boldsymbol{k}$ by transformation equation (7.67). Let us substitute this transformation to the right side of equation (7.227), thus, obtaining

$$\begin{aligned}\frac{\delta\nu}{\nu_0} &= \boldsymbol{K}\cdot(\boldsymbol{\beta}-\boldsymbol{\beta}_0)\left[1-\boldsymbol{K}\cdot\boldsymbol{\beta}_0+(\boldsymbol{K}\cdot\boldsymbol{\beta}_0)^2-\frac{\beta_0^2}{2}+\frac{\beta^2}{2}\right]-\frac{\beta_0^2}{2}+\frac{\beta^2}{2} \\ &\quad +4\sum_{a=1}^{N}\frac{Gm_a}{c^2}\frac{[\boldsymbol{K}\times(\boldsymbol{\beta}-\boldsymbol{\beta}_a)]\cdot(\boldsymbol{K}\times\boldsymbol{r}_a)}{|\boldsymbol{\xi}-\boldsymbol{\xi}_a|^2}+(\beth-\beth_0)\cdot(\boldsymbol{\beta}-\boldsymbol{\beta}_0)\,.\end{aligned} \tag{7.230}$$

where the gravitational deflection angle $\beth=-r/R\,\boldsymbol{\alpha}$, and the angle $\beth(\tau_0)$ is negligibly small, as follows from equations (7.200), (7.201)). Making use of these expressions for $\beth$ and $\beth_0$ in equation (7.230), one obtains for the observable gravitational shift of frequency by a non-stationary gravitational lens (special-relativistic terms are not shown)

$$\left(\frac{\delta\nu}{\nu_0}\right)_{gr} = \frac{4}{c}\frac{\partial\psi}{\partial t^*}+\frac{r_0}{R}(\boldsymbol{\beta}\cdot\boldsymbol{\alpha})+\frac{r}{R}(\boldsymbol{\beta}_0\cdot\boldsymbol{\alpha})\ , \tag{7.231}$$

where r and r_0 are distances from observer to the lens and from the lens to the source of light respectively, $R=|\boldsymbol{x}-\boldsymbol{x}_0|\simeq r+r_0$, and the retarded time s was approximated by the time of the light-ray closest approach t^*. It is easy to see that in the limit $r_0\to+\infty$, $r=const.$, the equation (7.231) goes back to equation (7.228).

The last two terms in the right side of equation (7.231) have been derived by Bertotti & Giampieri [Bertotti and Giampieri, 1992] who used a different mathematical technique assuming that the lens is static. This assumption did not allow them to derive the first term in the right side of equation (7.231) discovered by Birkinshaw & Gull [Birkinshaw and Gull, 1983] who, in their own turn, neglected the important contributions due to the motion of the source of light and observer. It is also useful to note that equation (22) in Bertotti & Giampieri's paper [Bertotti and Giampieri, 1992] for the Doppler shift by a gravitational lens contains a misprint of algebraic sign in front of the term depending on the velocity of observer. The error has been corrected in paper [Iess *et al.*, 1999, equation 8] so that their final result coincides precisely with the last two terms in the right side of equation (7.231).

7.7 Relativistic Astrometry in the Solar System

7.7.1 Near-zone and far-zone astrometry

Basic theoretical principles of relativistic astrometry in the solar system were based for a long time on the post-Newtoninan approximate solution of the Einstein field equations in the barycentric reference frame of the solar system [Brumberg, 1972, 1991,?; Kaplan, 1998; Soffel, 1989; Will, 1993]. The metric tensor of the post-Newtonian solution is given in section 4.3.2 of the present book and represents an instantaneous function of coordinate time t because it is found from the Poisson-type equations. The metric also depends on the field point, $\boldsymbol{x}$, the coordinates, $\boldsymbol{x}_a(t)$, and velocities, $\boldsymbol{v}_a(t)$, of the gravitating bodies. The post-Newtonian solution of the Einstein equations consists of an asymptotic series of terms expanded with respect to a small parameter $\epsilon \sim r/\lambda_{gr}$, where λ_{gr} is a characteristic wavelength of gravitational waves emitted by the solar system and r is a distance of the field point from the center of mass of the solar system [Fock, 1964; Misner *et al.*, 1973]. The wavelength $\lambda_{gr} \simeq cP_b$ where c is the speed of gravity [12], P_b is a characteristic orbital period of the solar system bodies that can be viewed as the period of one revolution of Jupiter around the Sun.

The post-Newtonian series must be convergent which demands $r \leq \lambda_{gr}$, making the series valid only inside the near zone of the solar system bounded by radius $\lambda_{gr} \simeq 1.7$ pc. Hence, the post-Newtonian solution of the Einstein equations can be self-consistently applied for consideration of propagation of light rays only inside the near zone of the solar system. This limitation can be completely removed for astrometric observations having precision at the level of 1 millarcsecond (mas). In this case the gravitational field of planets can be ignored and the Sun can be considered as static, spherically-symmetric body with a Schwarzschild metric [Misner *et al.*, 1973] extending up to infinity. However, current astrometric techniques have already achieved precision in ten microarcsecond (μas) [Fomalont and Kopeikin, 2003; Hagiwara *et al.*, 2009], and will be approaching towards 1 μas threshold in a foreseeable future [Fomalont and Reid, 2004; Johnston *et al.*, 2000]. For this reason, one can not ignore non-stationarity of the gravitational field of the solar system and should work out appropriate theoretical tools for adequate mathematical description of propagation of light from any source of light both inside and outside of the near zone of the solar system to an observer located somewhere inside the solar system..

Planets move slowly compared with the speed of light and gravity. Hence, if the source of light is inside the near zone of the solar system one can find solution of the relativistic equation of light propagation (7.46) by means of expanding positions and velocities of the solar system bodies entering the metric tensor, in a Taylor

12) The speed of gravity equals to the speed of light as postulated in general relativity. This postulate is a matter of experimental testing by various techniques - the most notable are pulsar timing [Kramer and Wex, 2009] and VLBI [Fomalont and Kopeikin, 2003; Kopeikin, 2004].

series around some fixed instant of time, substituting this expansion into equation (7.46), and integrating them with respect to time. Such approach can be dubbed as a *near-zone astrometry* as it works only for description of astrometric measurements done within the solar system like radar ranging [Pitjeva, 1993; Reasenberg *et al.*, 1979; Shapiro, 1964] and lunar laser ranging [Battat *et al.*, 2007; Dickey *et al.*, 1994; Kopeikin *et al.*, 2008; Nordtvedt, 1999; Williams *et al.*, 1996] experiments as well as for interpretation of the Doppler tracking of satellites and interplanetary spacecrafts [Anderson *et al.*, 2004; Bertotti and Giampieri, 1992; Bertotti and Iess, 1985; Bertotti *et al.*, 2003, 1999; Iess *et al.*, 1999; Krisher *et al.*, 1993a,b]. Theoretical problem which is characteristic for the near-zone astrometry is how to determine the fiducial instant of time around which coordinates and velocities of the moving gravitating bodies should be expanded, in order to get the correct astrometric predictions for gravitational time delay of light, its deflection angle, and the frequency shift.

The origin of this difficulty is that the post-Newtonian approximations deal with the instantaneous gravitational potentials while light propagates along the characteristics of the null cone. However, gravity in general relativity propagates with the same speed as light and it must interact with light not instantaneously but with a finite speed irrespectively of the approximation scheme. The instantaneous potentials does not allow us to make full use of the true hyperbolic character of the Einstein field equations which necessitates making some further assumptions about the initial values of positions and velocities of the massive bodies in order to integrate the equations of light propagation (see Figure 7.3 for more details). One of the most reasonable choices is to fix the coordinates and velocities of the body at the moment of the closest approach of light ray to it as it was implemented in papers by Hellings [Hellings, 1986a,b] and Will [Will, 2003]. Klioner & Kopeikin [Klioner and Kopeikin, 1992] and Klioner [Klioner, 2003a] had argued that taking the bodies at the instant of the closest approach minimizes the magnitude of the residual terms of the post-Newtonian solution of the equations of light propagation. The same assumption is used in the standards of the International Earth Rotation Service (IERS) [McCarthy and Petit, 2004]. Actual progress in solution of this problem can be achieved only with making use of the post-Minkowskian approximations as described in previous sections. Its application is discussed in more detail later.

So far, one has been discussing the near-zone astrometry where the post-Newtonian approximations could be used as a reasonable theoretical tool for calculating relativistic perturbations in the propagation of light rays. However, many sources of light are lying at much larger distances from the solar system far outside of its near zone ($r > \lambda_{gr}$). Calculation of relativistic perturbations of a light ray coming from such a source to the solar system is a subject of a*far zone astrometry*. There are many interesting theoretical aspects of this calculation because it deals with the perturbation of the light ray by propagating gravitational waves [Kopeikin and Korobkov, 2005; Kopeikin *et al.*, 2006]. The perturbation is rather weak, definitely less than 1 μas, but it may be feasible for observation in some future by powerful radio and/or optical telescopes (see, for example [Dewdney *et al.*, 2009]). For this reason, astronomers are making observational proposals to detect these effects already now [Jenet *et al.*, 2005; Kopeikin and Gwinn, 2000; Kopeikin *et al.*, 1999; Manchester, 2008; Pyne *et al.*, 1996]. The post-Newtonian approach has apparent inconsistencies when one tries to apply it to integration of equations of light propagation in time-dependent gravitational field of the far zone. In this case the point of emission, $\boldsymbol{x}_0$, of the light ray and that of observation, $\boldsymbol{x}$, are separated by distance r which is much larger than the characteristic wavelength of gravitational waves emitted by the solar system. The first breakdown of the post-Newtonian approach comes from the fact that the post-Newtonian expansion of the metric tensor diverges as the distance r of the field point from the system increases [Anderson and Decanio, 1975; Chandrasekhar and Esposito, 1970; Damour and Schäfer, 1985; Ehlers *et al.*, 1976; Ohta *et al.*, 1973]. This divergency has been basically ignored by previous researches who used for the integration of the equations of light rays the truncated post-Newtonian form of the metric tensor

$$g_{00}(t,\boldsymbol{x}) = -1 + \frac{2U(t,\boldsymbol{x})}{c^2} + O(c^{-4}) , \tag{7.232}$$

$$g_{0i}(t,\boldsymbol{x}) = -\frac{4U^i(t,\boldsymbol{x})}{c^3} + O(c^{-5}) , \tag{7.233}$$

$$g_{ij}(t,\boldsymbol{x}) = \delta_{ij}\left[1 + \frac{2U(t,\boldsymbol{x})}{c^2}\right] + O(c^{-4}) , \tag{7.234}$$

where the instantaneous, Newtonian-like potentials are given by the expressions

$$U(t,\boldsymbol{x}) = \sum_{a=1}^{N} \frac{m_a}{|\boldsymbol{x} - \boldsymbol{x}_a(t)|} , \tag{7.235}$$

$$U^i(t,\boldsymbol{x}) = \sum_{a=1}^{N} \frac{m_a v_a^i(t)}{|\boldsymbol{x} - \boldsymbol{x}_a(t)|} , \tag{7.236}$$

and all higher order terms have been omitted. From a formal point of view the expressions (7.232)-(7.234) are not divergent when the distance r goes to infinity but the omitted residual terms in the metric tensor diverge. The divergent residual terms in the metric tensor affect propagation of the light ray and limit the applicability of equations (7.232)-(7.234). The divergency means that the perturbation of the light ray will become arbitrary large as it goes from the source of light to observer in the solar system. This makes a certain interpretational difficulty for astronomers who are going to measure relativistic effects of time-dependent gravitational field produced by the planets of the solar system and/or similar gravitational systems in the sky like a binary star

Another problem with the application of the near-zone expansion of the metric tensor to far-zone relativistic astrometry relates to the retarded nature of the propagation of the gravitational interaction as predicted by general relativity. Expressions (7.232)-(7.234) are instantaneous functions of time and do not show this property of the retardation at all. One can naively ignore the wave nature of gravity and continue to operate with the post-Newtonian metric (7.232)-(7.234) for integration of equations of light rays, at least formally, because the integration will give a convergent result. However, one may expect that the trajectory of light ray, obtained by solving the equations of propagation of light with the instantaneous potentials, will deviate from that obtained by making use of the metric perturbations expressed as the retarded *Liénard-Wiechert* potentials 7.12. The deviation can get, in principle, comparable with the main term of the relativistic deflection of light and time delay, thus, making interpretation of precise astrometric observations of the relativistic effects insensible.

One more, purely technical problem relates to the method of performing the time integration of the instantaneous potentials along the light ray trajectory from the time of emission of light, t_0, to the time of observation, t. It emerges because coordinates and velocities of bodies are not linear functions of time in the most general case of a gravitationally bounded system. Even in the case of the body being on a circular orbit one has a problem of solving integrals of the type

$$\int_{t_0}^{t} U(t,\boldsymbol{x})dt = \sum_{a=1}^{N} m_a \int_{t_0}^{t} \frac{dt}{|\boldsymbol{x}_0 + \boldsymbol{k}\,(t-t_0) - A_a\left[\boldsymbol{e}_1 \sin(\omega_a t + \varphi_a) + \boldsymbol{e}_2 \cos(\omega_a t + \varphi_a)\right]|}\,, \tag{7.237}$$

where A_a, ω_a, and φ_a are the radius, the angular frequency, and the initial phase of the orbit of the a-th body respectively, and $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ are the orthogonal unit vectors lying in the orbital plane. It is easy to check that the given integral can not be performed analytically so that one has to resort to numerical methods. In case of elliptical motion of the bodies calculations will be even more complicated. It was usually assumed that the main contribution to the integral (7.237) comes from that part of the trajectory of the light ray which passes near the body deflecting the light ray so that one is allowed to fix position and velocity of the body at some instant of time, t_a, which is close to the moment t^* of the closest approach of the light ray to the body. Analytic errors of such approximation were discussed by Klioner & Kopeikin [Klioner and Kopeikin, 1992], Klioner [Klioner, 2003a] and Will [Will, 2003] for the case of $t_a = t^*$. They proved that such an approximation is good enough for most practical cases. However, more precise evaluation shows [Kopeikin and Mashhoon, 2002; Kopeikin and Schäfer, 1999] that the time t_a must be identified with the retarded time $s = t - r_a/c$ taken by gravity to propagate from the massive body to the observer (see section 7.9 for more detail). In any case, time series expansion of the body's coordinate and velocity works for calculation of the integral (7.237) only when photon moves inside the near zone of the gravitating system. Far outside of the system another method of solving integral (7.237) is required as shown by Klioner and Kopeikin [1992] and further elaborated on by Will [2003].

Indeed, if one tries to perform a global integration of integral (7.237) using the Taylor time-series expansion of the body' coordinates, the correct logarithmic behavior of the integral can be reproduced only if the first two terms in the expansion are taken into account which is physically equivalent to the case of the bodies moving uniformly along straight lines. Accounting for the third term in the expansion of the coordinates (accelerated motion of the body) changes the logarithmic behavior of the integral for large time intervals of integration exceeding the characteristic Keplerian period of the system, thus, making prediction for the gravitational time delay

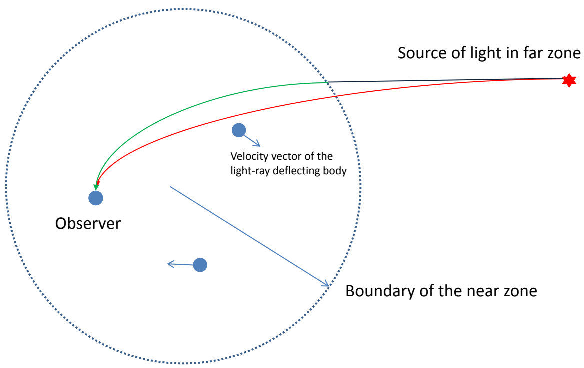

Figure 7.7 Source of light is lying in a far zone beyond the boundary of the near zone of the solar system. Light-ray trajectory is split in two asymptotic parts. Each part is found differently by integrating the static Schwarzschild metric in the far zone (blue line) and the post-Newtonian metric in the near zone (green line) under assumption that the bodies move along straight lines with constant velocities. Real trajectory of the light ray is shown in red color. It is obtained by integrating the post-Minkowski metric in the near and far zones without splitting the light-ray path in two pieces.

and the total angle of deflection of light erroneous. This is one more reason why the Taylor time-series expansion is invalid for finding numerical value of the integral (7.237) in the case when the source of light is beyond the limit of the near zone of the solar system.

Recognizing these difficulties one has proposed [Kopeikin, 1990] to use matching asymptotic technique for finding the gravitationally perturbed trajectory of the light ray going from a remote source of light like a pulsar or a quasar to the solar system. This technique was further developed by Klioner & Kopeikin [Klioner and Kopeikin, 1992] and by Will [Will, 2003]. Graphical presentation of this technique is shown in Figure 7.7. Technically, the whole spacetime was separated in two domains - the near and far zones lying correspondingly inside and outside of the distance $r \simeq \lambda_{gr}$ being approximately equal to the characteristic length of the gravitational waves emitted by the solar system. The internal solution of the equations of light rays (the blue line in Figure 7.7) within the near zone have been obtained by expanding coordinates and velocities of the bodies in the Taylor time series, and then integrating the light geodesic equations. The external solution of the equations (the green line in Figure 7.7) has been found by decomposing the metric tensor in gravitational multipoles and accounting only for the first monopole term (mass) which corresponds the case of static, spherically symmetric field of the Sun. A global solution was obtained by matching of the internal and external solutions at the buffer region in order to reach the required astrometric accuracy of 1 μas.

Unfortunately, the matching asymptotic technique is rather cumbersome to implement in analytic calculations. Moreover, it does not help very much for clarification of the question concerning the fiducial instant of time used in the Taylor expansion of positions and velocities of the massive bodies when integrating equations of light propagation inside the near zone. In addition, the matching technique does not give any recipe how to integrate equations of light propagation in the far zone if the higher, time-dependent gravitational multipoles should be taken into account and what magnitude of the perturbations and/or residual terms one should expect. In any case, the global solution obtained by the matching of the two asymptotics consists of two pieces making visualization of the light ray trajectory difficult and astrometric implementation of the method impractical.

For these reasons, the present book does not rely upon the matching technique but resort to the method of integration of light-ray equations based on the retarded *Liénard-Wiechert* potentials as advocated in papers [Kopeikin *et al.*, 2006; Kopeikin and Mashhoon, 2002; Kopeikin and Schäfer, 1999]. This method allows us to construct a smooth and unique global solution (red curve in Figure 7.7) of the light propagation equations from arbitrary distant source of light to observer located inside the solar system. One is able to handle the integration of the equations more easily and can estimate the magnitude of all residual terms. Proceeding in this way one also gets a unique prediction for that moment of time at which coordinates and positions of gravitating bodies should be fixed for correct and adequate calculation of the light deflection angle and other astrometric effects. Three kinds of observations will be considered - pulsar timing, very long baseline interferometry (VLBI) of quasars, and optical astrometric observations of stars. Astrometric experiments with massive planets will be considered as well.

7.7.2
Pulsar timing

Pulsar timing is the regular monitoring of the rotation of the neutron star by tracking the times of arrival (TOA) of the radio pulses. Pulsar timing unambiguously accounts for every single rotation of the neutron star over long periods (from years to decades) of time. This unambiguous and very precise tracking of rotational phase allows pulsar astronomers to probe the interior physics of neutron stars, make extremely accurate astrometric measurements, and test gravitational theories in the strong-field regime in unique ways. In the proper inertial frame of a pulsar, the intrinsic rotational period of the pulsar is nearly constant, so the time-dependent phase $\phi(T)$ of a pulsar can be approximated by a Taylor expansion with respect to the pulsar proper time T

$$\phi(T) = \phi_0 + fT + \frac{1}{2}\dot{f}T^2 + O(T^3) , \tag{7.238}$$

where ϕ_0 is arbitrary reference phase for each pulsar, f and $\dot{f}$ are the pulsar's rotational frequency and its derivative (caused by emission of the electromagnetic energy), and the residual term $\sim T^3$ describes instability of the pulsar's rotation (timing noise) [Kopeikin, 1997b; Lorimer and Kramer, 2004] In order to measure $\phi(T)$ many transformations have to be applied to the observed TOAs first. The proper time T of the pulsar should be converted to the barycentric time of the solar system, the time of emission must be related to the time of observation by the time delay equation (7.72), and the barycentric TOAs must be converted to the proper time of observer τ correspondingly [Lorimer and Kramer, 2004]. It is out of the scope of this section to describe all corrections to the timing formula - it has been already done in a handbook on pulsar astronomy [Lorimer and Kramer, 2004] to which the reader is referred. One will focus on the discussion of a particular gravitational effect - time delay in the gravitational field of moving solar system bodies - the massive planets and Sun.

Mathematical description of the time delay is based on equations (7.72)–(7.74). Taking in the post-Minkowski equation (7.74) only the terms up to the order v_a/c inclusively, one obtains

$$B_a(s, s_0) = -\ln\left[\frac{r_a(s) - \boldsymbol{k}\cdot\boldsymbol{r}_a(s)}{r_a(s_0) - \boldsymbol{k}\cdot\boldsymbol{r}_a(s_0)}\right] - \int_{s_0}^{s} \frac{\boldsymbol{k}\cdot\boldsymbol{\beta}_a(\zeta)\, d\zeta}{t^* + \frac{1}{c}\boldsymbol{k}\cdot\boldsymbol{x}_a(\zeta) - \zeta} + O\left(\beta_a^2\right) , \quad (7.239)$$

where the retarded times s and s_0 should be calculated from equations (7.13) and (7.75) respectively, $\boldsymbol{r}_a(s) = \boldsymbol{x} - \boldsymbol{x}_a(s)$, $\boldsymbol{r}_a(s_0) = \boldsymbol{x}_0 - \boldsymbol{x}_a(s_0)$, and one assumes that the observation is made at the point with the spatial barycentric coordinates $\boldsymbol{x}$ at the instant of time t, and the pulsar's pulse is emitted at the moment t_0 from the point $\boldsymbol{x}_0$. The distance of pulsar from the solar system is, typically, more than 100 pc. One emphasizes how simple is the integration of the retarded *Liénard-Wiechert* gravitational potentials along the null cone instead of that of the instantaneous post-Newtonian potentials, which bring about integrals like those in equation (7.237) having no simple solution.

In principle, the first (logarithmic) term in this formula is enough to treat timing data for any pulsar with an accuracy being sufficient for practical purposes. The denominator in the argument of the logarithmic function is $r_a(s_0) - \boldsymbol{k}\cdot\boldsymbol{r}_a(s_0) \simeq 2R$, where R is the distance between the barycenter of the solar system and the pulsar. The logarithm of $2R$ is a function which is nearly constant but can have a secular change because of the slow relative (proper) motion of the pulsar with respect to the solar system. All slowly-changing terms can be written down as a polynomial of time that is absorbed to the pulsar's rotational phase, ϕ, by re-definition of the pulsar's rotational parameters. Thus, such terms can not be observed directly. For this reason, the denominator in the logarithmic term of equation (7.239) will be omitted. One emphasizes that positions of the solar system bodies in the numerator of the logarithmic term are to be taken at the moment of the retarded time which is found by iterations of the gravity null-cone equation $s = t - c^{-1}|\boldsymbol{x} - \boldsymbol{x}_a(s)|$. It makes calculation of the gravitational time delay in the solar system theoretically consistent and practically more precise.

There is a small difference between the logarithmic term in (7.239) and the corresponding logarithmic term in timing formulas suggested by Hellings [Hellings, 1986a,b] and Doroshenko & Kopeikin [Doroshenko and Kopeikin, 1990, 1995] where the position of the a-th body is taken at the moment of the closest approach of the radio pulse to the body instead of the retarded time. It is, however, not so important in current practice of timing observations as they are not yet precise enough to distinguish the retarded time from that of the closest approach. Indeed, the maximal time difference is expected to be of order of $(4GM_a/c^3)(d_a/r_a)(v_a/c)$, where M_a and v_a are mass of the body a and its barycentric velocity respectively, and d_a is the impact parameter of the light ray from the pulsar with respect to the body. For the Sun $v_\odot$ is less than 15 m/s and $d_a \leq R_\odot$. This difference has maximal value for the light ray grazing the solar limb and is many orders of magnitude less than 10 nanosecond - the current accuracy of the most precise timing observations [Edwards *et al.*, 2006; Hobbs *et al.*, 2006]. Similar estimates hold for major planets of the solar system - Jupiter and Saturn.

The integral in equation (7.239) can not be calculated analytically if the trajectory of motion of the bodies is not simple. Similar integral were calculated in section 7.6.1.1 which was devoted to discussion of timing formula of binary pulsars. However, in case of a binary pulsar the electromagnetic pulse passes by the pulsar's companion almost immediately after it was emitted by the pulsar. It makes the difference $s - s_0$ between the retarded times s and s_0 corresponding to the times of observation, t, and emission, t_0, of the pulse, very small - of the order of the time the light and/or gravity needs to cross the binary system as evident from equation (7.148) and caption to Figure 7.2. More exact, light propagates from the binary system in the same direction as the gravitational waves emitted by it, so that the gravitational field of the binary system is almost "frozen in-time" for the outgoing photon. Conversely, when light propagates from distant source towards the solar system, it moves opposite to the direction of propagation of gravitational waves emitted by the moving bodies of the solar system. For this reason, the difference $s - s_0$ becomes very large, effectively almost equal to $\simeq 2R/c$ - a double distance between the source of light and the solar system. Thus, one is not allowed, as it was in the case of derivation of timing formula for binary pulsars, to use expansion of coordinates and velocities of the solar system bodies in a Taylor series with respect to time for calculation of the integral. Nevertheless, the integral still can be evaluated analytically. Let us describe the main idea of this calculation as it clarifies how variable gravitational fields affect propagation of light on its long travel from the source of light to observer.

First of all, notice that the orbital plane of any of the solar system bodies lies very close to ecliptic and can be approximated fairly well by a circular motion up to the first order correction with respect to the orbital eccentricity, which is usually small. Orbital motion of the Sun with respect to the barycenter of the solar system may be described as a sum of harmonics corresponding to gravitational perturbations from Jupiter, Saturn, and other large planets [Hardorp, 1985]. Thus, one assumes that $\boldsymbol{x}_a$ is given in the ecliptic plane as follows

$$\boldsymbol{x}_a(t) = a\left[\cos(nt)\,\boldsymbol{e}_1 + \sin(nt)\,\boldsymbol{e}_2\right] , \qquad (7.240)$$

where a is the radius of the cicular orbit of the massive body, $n = 2\pi/P$ is the orbital frequency with P - the orbital period, $\boldsymbol{e}_1$ is a unit vector directed to the point of the vernal equinox, $\boldsymbol{e}_2$ is a unit vector being orthogonal to $\boldsymbol{e}_1$ and lying in the ecliptic plane. The unit vector $\boldsymbol{k}$ is defined in ecliptic coordinates as

$$\boldsymbol{k} \quad = \quad -\cos b \cos l\, \boldsymbol{e}_1 - \cos b \sin l\, \boldsymbol{e}_2 - \sin b\, \boldsymbol{e}_3 \,, \tag{7.241}$$

where b and l are the ecliptic spherical coordinates of the pulsar. Now, one solves equation (7.77) for $y = t^* + \frac{1}{c}\boldsymbol{k}\cdot\boldsymbol{x}_a(\zeta) - \zeta$, by iterations with respect to ζ. One can do it because each body in the solar system moves slowly with respect to the speed of light and gravity, and the difference between $\boldsymbol{x}_a(\zeta)$ and $\boldsymbol{x}_a(y)$ is of the order of $v_a/c \ll 1$. Thus, the approximate solution of equation (7.77) is $\zeta = t^* - y + \boldsymbol{k}\cdot\boldsymbol{x}_a(y)$. Making use of this variable and equation (7.240) in the integral of equation (7.239) reduces it to the following form

$$\int_{s_0}^{s} \frac{\boldsymbol{k}\cdot\boldsymbol{\beta}_a(\zeta)\, d\zeta}{t^* + \dfrac{1}{c}\boldsymbol{k}\cdot\boldsymbol{x}_a(\zeta) - \zeta} = \frac{2\pi a}{cP}\int_{y_0}^{y} \frac{\boldsymbol{k}\cdot\boldsymbol{e}_1 \sin(ny - nt^*) + \boldsymbol{k}\cdot\boldsymbol{e}_2 \cos(ny - nt^*)}{y} dy\,. \tag{7.242}$$

The integral is reduced to the integral sine and cosine special functions. Remembering that for a circular orbit $\beta_a = 2\pi a/cP$, one obtains

$$\begin{aligned} \int_{s_0}^{s} \frac{\boldsymbol{k}\cdot\boldsymbol{v}_a(\zeta)d\zeta}{t^* + \dfrac{1}{c}\boldsymbol{k}\cdot\boldsymbol{x}_a(\zeta) - \zeta} \quad &= \quad \beta_a \boldsymbol{k}\cdot\boldsymbol{x}_a(t^*)\left\{\mathbf{Si}\left[\frac{2\pi}{cP}(r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a)\right] - \mathbf{Si}\left[\frac{2\pi}{cP}(r_{0a} - \boldsymbol{k}\cdot\boldsymbol{r}_{0a})\right]\right\} \\ &+ \quad \boldsymbol{k}\cdot\boldsymbol{\beta}_a(t^*)\left\{\mathbf{Ci}\left[\frac{2\pi}{cP}(r_a - \boldsymbol{k}\cdot\boldsymbol{r}_a)\right] - \mathbf{Ci}\left[\frac{2\pi}{cP}(r_{0a} - \boldsymbol{k}\cdot\boldsymbol{r}_{0a})\right]\right\}, \end{aligned} \tag{7.243}$$

where

$$\mathbf{Si}(z) \quad = \quad \int_0^z \frac{\sin x}{x} dx\,, \tag{7.244}$$

$$\mathbf{Ci}(z) \quad = \quad -\int_z^\infty \frac{\cos x}{x} dx\,, \tag{7.245}$$

are sine and cosine integrals [Arfken and Weber, 2001].

Taking into account the asymptotic behavior of sine and cosine integrals for large and small values of their arguments in relationship (7.243) one obtains the final formula for the gravitational time delay in the solar system

$$\Delta(t) \quad = \quad -\frac{2G}{c^3}\sum_{a=1}^{N} m_a \left[1 - \boldsymbol{k}\cdot\boldsymbol{\beta}_a(t^*)\right] \ln\left[r_a(s) - \boldsymbol{k}\cdot\boldsymbol{r}_a(s)\right] + O\left(\frac{Gm_a}{c^3}\beta_a^2\right), \tag{7.246}$$

where the time t^* must be understood as the time of the closest approach of light ray to the body a. If one takes numerical values of masses and velocities of the solar system bodies one finds that such residual terms are much smaller than the precision of timing measurements. One concludes that these residual terms can not be detected by the present day pulsar timing techniques. Nonetheless, terms, which are linear with respect to velocities of the solar system bodies, are shown for completeness of equation (7.246).

7.7.3
Very long baseline interferometry

Very long baseline interferometry (VLBI) measures the time difference in the arrival of a microwave signal from a radio source received at two or more radio observatories making up an array of radio telescopes. [Sovers *et al.*, 1998]. Data received at each antenna in the array is paired with timing information, usually from a local atomic clock, and then stored for later analysis on magnetic tape or hard disk. At that later time, the data are correlated with data from other antennas similarly recorded, to produce the resulting image. The resolution achievable using interferometry is proportional to the observing frequency and the distance between the antennas farthest apart in the array. VLBI is most well-known for imaging distant cosmic radio sources, spacecraft tracking, and for various applications in astrometry [Mantovani and Kus, 2004]. However, since the VLBI technique measures the time differences between the arrival of radio waves at separate antennas, it can also be used to perform the Earth rotation studies, map movements of tectonic plates within the accuracy of a millimeter, and conduct other types of geodesy [McCarthy and Petit, 2004]. Using VLBI in this manner requires large numbers of time difference measurements from distant extragalactic sources (such as quasars) observed with a global network of antennas over a period of time.

Generally, geodetic observing sessions run for 24 hours and observe a number of different radio sources distributed across the sky. The observatories can be widely separated; the sensitivity of the observations to variations in the orientation of the Earth increases with the size of the VLBI network. VLBI is the only technique capable of measuring all components of the Earth's orientation accurately and simultaneously. Currently, VLBI determinations of Earth-rotation variations, and of the coordinates of terrestrial sites and celestial objects are made routinely and regularly with estimated angular accuracies of about ±0.1 milliarcsecond or better [Schlüter and Behrend, 2007; Schlüter and Vandenberg, 2003]. Such a high precision of observations requires an extremely accurate accounting for different physical effects in propagation of light from radio sources to observer including relativistic gravitational time delay, which is the main concern in this section.

There have been many papers dealing with relativistic effects which must be accounted for in VLBI data processing software (see, e.g., [Kaplan, 1998; Klioner, 1991a; Kopeikin and Schäfer, 1999; Kopeikin *et al.*, 1999], and references therein). The common efforts of many researchers in this area have resulted in the creation of a standard algorithm of VLBI data processing called a consensus model [McCarthy and Petit, 2004; Petit and McCarthy, 2005] that emerged from a workshop on *relativistic models for use in space geodesy* held at USNO in 1990. The consensus model is based on a combination of several relativistic algorithms developed by various groups around the world. The accuracy limit chosen for the consensus VLBI relativistic time delay model is fixed at the level 10^{-12} seconds (one picosecond) of differential VLBI delay for baselines less than two Earth radii in length. As it is stated, all terms in the model of the order of 10^{-13} seconds or larger were included to ensure that the final result is accurate at the picosecond level [McCarthy and Petit, 2004]. By definition, astrometric coordinates of an extragalactic source derived from the consensus model should have no apparent motions due to the solar system relativistic effects at the picosecond level. Though the picosecond accuracy is still quite satisfactory for many scientific purposes, a great improvement in observational VLBI technologies [Fomalont and Reid, 2004; Middelberg and Bach, 2008] accompanied by invention of a powerful phase referencing technique whose goal is a 10 microarcsecond accuracy for astrometric measurements [Fomalont, 2005; Fomalont and Kopeikin, 2002], make it clear that one should develop much better understanding of the relativistic effects that must be included to VLBI data processing algorithm.

In what follows one considers propagation of radio signal through the solar gravitational field in the barycentric coordinates $x^\alpha = (t, \boldsymbol{x})$. Precise definition of the measuring procedure applied in VLBI data processing software requires, however, derivation of relativistic relationship between the proper (atomic) time of observer and the barycentric coordinate time, t. It is given in the IERS manual [McCarthy and Petit, 2004] and can be added to the formalism of the present section for adapting it to practical applications. A complete description of the VLBI data processing formalism is too complicated and goes beyond the scope of the present book.

The differential VLBI time delay is explained graphically in Figure 7.8. It is calculated as the time, t_2, of arrival of the front of a radio wave at station 2 minus the time of arrival, t_1, of the same front at station 1. The time of arrival at station 1 serves as the time reference for the measurement. In what follows, unless explicitly stated otherwise, all vectors and scalar quantities are assumed to be calculated at the time t_1 except for position of the source of light, $\boldsymbol{x}_0$, which is always calculated at the of time of the light emission, t_0. Calculation of the VLBI time delay relies upon equations (7.72), (7.73) referred to the barycentric coordinate frame of the solar system. The equations give us

$$t_2 - t_1 = \frac{1}{c}|\boldsymbol{x}_2(t_2) - \boldsymbol{x}_0| - \frac{1}{c}|\boldsymbol{x}_1 - \boldsymbol{x}_0| + \Delta(t_2, t_0) - \Delta(t_1, t_0)\,, \tag{7.247}$$

where $\boldsymbol{x}_0$ are coordinates of the source of light, $\boldsymbol{x}_2(t_2)$ are coordinates of the station 2 at time t_2, $\boldsymbol{x}_1$ are coordinates of the station 1 at time t_1. The differential relativistic time delay is given by

$$\Delta(t_2, t_0) - \Delta(t_1, t_0) = \frac{2G}{c^3}\sum_{a=1}^{N}\left[B_a(s_2, s_0) - B_a(s_1, s_0)\right]\,, \tag{7.248}$$

where the difference of the B_a's up to the linear with respect to velocities of the solar

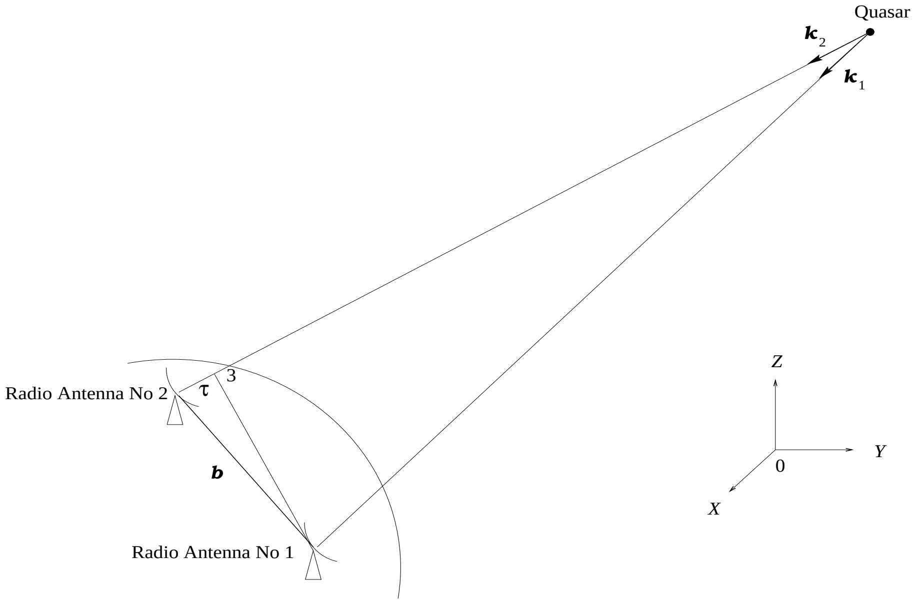

Figure 7.8 Very long baseline interferometry measures the delay τ between times of arrival of the front of a radio wave from a radio source (quasar) at the first and second radio antennas, $\tau = \tau_2 - \tau_1$, located on the Earth. This is equivalent to measuring the light travel time between points 2 and 3. Diurnal rotation and orbital motion of the Earth makes the delay dependent on time. This allows us to determine the baseline vector $\boldsymbol{b}$ between the two antennas, astrometric coordinates of the quasar, position and motion of the Earth's rotational axis (pole) including its precession, nutations, wobble, and many others. Modern data processing of VLBI observations is fully based on general-relativistic conceptions.

system bodies reads (see equation (7.81))

$$\begin{aligned} B_a(s_2, s_0) - B_a(s_1, s_0) &= \ln\frac{r_{1a} - \boldsymbol{k}_1\cdot\boldsymbol{r}_{1a}}{r_{2a} - \boldsymbol{k}_2\cdot\boldsymbol{r}_{2a}} - \ln\frac{r_{0a} - \boldsymbol{k}_1\cdot\boldsymbol{r}_{0a}}{r_{0a} - \boldsymbol{k}_2\cdot\boldsymbol{r}_{0a}} \\ &+ \boldsymbol{k}_2\cdot\boldsymbol{\beta}_a(s_2)\ln(r_{2a} - \boldsymbol{k}_2\cdot\boldsymbol{r}_{2a}) - \boldsymbol{k}_1\cdot\boldsymbol{\beta}_a(s_1)\ln(r_{1a} - \boldsymbol{k}_1\cdot\boldsymbol{r}_{1a}) \\ &+ \int_{s_0}^{s_1} \ln\left[t_1^* + \boldsymbol{k}_1\cdot\boldsymbol{x}_a(\zeta) - \zeta\right]\ \boldsymbol{k}_1\cdot\dot{\boldsymbol{\beta}}_a(\zeta)\,d\zeta \\ &- \int_{s_0}^{s_2} \ln\left[t_2^* + \boldsymbol{k}_2\cdot\boldsymbol{x}_a(\zeta) - \zeta\right]\ \boldsymbol{k}_2\cdot\dot{\boldsymbol{\beta}}_a(\zeta)\,d\zeta\ . \end{aligned} \tag{7.249}$$

Here, s_1 and s_2, are retarded times determined iteratively by solving the null-cone equations

$$s_1 = t_1 - \frac{1}{c}|\boldsymbol{x}_1 - \boldsymbol{x}_a(s_1)|\ , \tag{7.250}$$

$$s_2 = t_2 - \frac{1}{c}|\boldsymbol{x}_2(t_2) - \boldsymbol{x}_a(s_2)|\ , \tag{7.251}$$

the quantity $\boldsymbol{r}_{1a} = \boldsymbol{x}_1 - \boldsymbol{x}_a(s_1)$ is a vector from the a-th body to the station 1, $\boldsymbol{r}_{2a} = \boldsymbol{x}_2(t_2) - \boldsymbol{x}_a(s_2)$ is a vector from the a-th body to the station 2, $r_{1a} = |\boldsymbol{r}_{1a}|$, $r_{2a} = |\boldsymbol{r}_{2a}|$ and

$$t_1^* = t_1 - \frac{1}{c}\boldsymbol{k}_1\cdot\boldsymbol{x}_1\ , \tag{7.252}$$

$$t_2^* = t_2 - \frac{1}{c}\boldsymbol{k}_2\cdot\boldsymbol{x}_2(t_2)\ , \tag{7.253}$$

are times of the closest approach of the light rays 1 and 2 to the barycenter of the solar system.

It will be also helpful in comparing the approach under discussion with the consensus model of IERS [McCarthy and Petit, 2004] to use the moments of the closest approach, t_{1a}^* and t_{2a}^*, of the light rays 1 and 2 to the a-th body. These moments are determined from the condition of minimum of the invariant distance $r_R = -u_a^\alpha r_\alpha$ between the propagating photon and the body, where u_a^α is a four-velocity of the body and $r^\alpha = (r_a, \boldsymbol{r}_a)$ is a four-vector connecting position of the photon with the body along the null cone of the body's gravity field. One has $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{k}(t - t_1)$ $(t < t_1)$ for the first light ray, and $\boldsymbol{x} = \boldsymbol{x}_2 + \boldsymbol{k}(t - t_2)$ $(t < t_2)$ for the second ray. Finding the minimum of the invariance distance r_R with respect to time t yields

$$t_{1a}^* = t_1 - \boldsymbol{k}_1\cdot\boldsymbol{r}_{1a}\ , \tag{7.254}$$

$$t_{2a}^* = t_2 - \boldsymbol{k}_2\cdot\boldsymbol{r}_{2a}\ , \tag{7.255}$$

where one has omitted all terms of the higher order with respect to the velocity of the body. It is worth emphasizing that the definition of times t_{1a}^* and t_{2a}^* is slightly different from the definition of similar quantities given in the consensus model. It relates to the positions of bodies in vectors $\boldsymbol{r}_{1a}$ and $\boldsymbol{r}_{2a}$ that are taken at the retarded times, s_1 and s_2, respectively, while in the consensus model they are taken at the times of arrival to the radio telescopes, t_1 and t_2. This difference yields the terms of the higher order, which were omitted because they go beyond the approximation adopted in the consensus model.

The unit vectors $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ are defined as

$$\boldsymbol{k}_1 = \frac{\boldsymbol{x}_1 - \boldsymbol{x}_0}{|\boldsymbol{x}_1 - \boldsymbol{x}_0|}\ , \qquad \boldsymbol{k}_2 = \frac{\boldsymbol{x}_2(t_2) - \boldsymbol{x}_0}{|\boldsymbol{x}_2(t_2) - \boldsymbol{x}_0|}\ , \tag{7.256}$$

which shows that they have slightly different orientations in space. Let us introduce the barycentric baseline vector at the time of arrival t_1 through the definition $\boldsymbol{B} = \boldsymbol{x}_2(t_1) - \boldsymbol{x}_1(t_1)$. Let us emphasize that the baseline vector lies on the hypersurface of constant time t_1. The barycentric baseline vector must be converted to the geocentric one, $\boldsymbol{b}$. This transformation was derived in [Kopeikin, 1990] and confirmed in later publications [Kaplan, 1998; McCarthy and Petit, 2004; Petit and McCarthy, 2005]. The difference between the barycentric and geocentric baselines in the expression for the gravitational time delay, caused by relativistic effects, will be neglected because it leads to terms of the higher (post-post-Newtonian) order of magnitude. Thus, one assumes $\boldsymbol{B} = \boldsymbol{b}$ in all subsequent equations. The difference between vectors $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ can be found using an expansion with respect to the powers of a small parameter b/R, where R is the distance between the barycenter of the solar system and the source of light. One has

$$\boldsymbol{x}_2 - \boldsymbol{x}_0 = \boldsymbol{x}_1 - \boldsymbol{x}_0 + \boldsymbol{b} + \boldsymbol{v}_2\,(t_2 - t_1) + O\left(\frac{v^2}{c^2}\,b\right)\ , \tag{7.257}$$

$$|\boldsymbol{x}_2 - \boldsymbol{x}_0| = |\boldsymbol{x}_1 - \boldsymbol{x}_0| + \boldsymbol{b}\cdot\boldsymbol{k}_1 + \boldsymbol{v}_2\cdot\boldsymbol{k}_1\,(t_2 - t_1) + O\left(\frac{v^2}{c^2}\,b\right) + O\left(\frac{b^2}{R}\right)\ ,$$

where $\boldsymbol{v}_2$ is the velocity of station 2 with respect to the barycenter of the solar system. These expansions yield

$$\boldsymbol{k}_2 = \boldsymbol{k}_1 + \frac{\boldsymbol{k}_1 \times (\boldsymbol{b} \times \boldsymbol{k}_1)}{R} + O\left(\frac{v}{c}\frac{b}{R}\right) + O\left(\frac{b^2}{R^2}\right)\ , \tag{7.258}$$

and for the time delay (7.247)

$$t_2 - t_1 = \frac{1}{c}\boldsymbol{k}_1\cdot\boldsymbol{b}\,[1 + \boldsymbol{\beta}_2\cdot\boldsymbol{k}_1 + ...] + \Delta(t_2, t_0) - \Delta(t_1, t_0)\ , \tag{7.259}$$

where the ellipsis denote a large number of relativistic corrections which are not considered here, but they can be found in the IERS manual [McCarthy and Petit, 2004]. As a consequence of the previous expansions the following equalities are hold:

$$t_2^* - t_1^* = \frac{(\boldsymbol{b}\times\boldsymbol{k}_1)(\boldsymbol{k}_1\times\boldsymbol{x}_2)}{cR} + O\left(\frac{b\,v^2}{c^3}\right) + O\left(\frac{b^2}{cR}\right)\ , \tag{7.260}$$

$$t_{2a}^* - t_{1a}^* = \frac{1}{c}(\boldsymbol{k}_1\cdot\boldsymbol{b} + r_{1a} - r_{2a})(\boldsymbol{k}_1\cdot\boldsymbol{\beta}_a) + \frac{(\boldsymbol{b}\times\boldsymbol{k}_1)(\boldsymbol{k}_1\times\boldsymbol{r}_{2a})}{cR} + O\left(\frac{b\,v^2}{c^3}\right) + O\left(\frac{b^2}{cR}\right)\ , \tag{7.261}$$

which evidently shows that, e.g. for the Jupiter (index J) and for the source of light at infinity ($R \to \infty$), the time difference $t^*_{2J} - t^*_{1J}$ is of the order $(R_\oplus/c)(R_J/r_{1J})(v_J/c) \simeq$ 7 picoseconds, that is, rather small but still may be important in the analysis of observational errors. The time difference $t^*_2 - t^*_1$ can be considered for extra-solar radio sources as negligibly small since it is of the order $(R_\oplus/c) = 23$ millisecond times the annual parallax of the source of light (pulsar, quasar), which makes it smaller than 1 picosecond. In case of VLBI observations of the solar system radio sources (spacecraft) the time difference $t^*_2 - t^*_1$ can not be ignored and should be taken into account everywhere in appropriate equations.

Now one can simplify formula (7.249). First of all, taking into account the relationship (7.258), one obtains

$$\ln \frac{r_{0a} - \boldsymbol{k}_1 \cdot \boldsymbol{r}_{0a}}{r_{0a} - \boldsymbol{k}_2 \cdot \boldsymbol{r}_{0a}} = -\frac{(\boldsymbol{b} \times \boldsymbol{k}_1)(\boldsymbol{k}_1 \times \boldsymbol{r}_{0a})}{R(r_{0a} - \boldsymbol{k}_1 \cdot \boldsymbol{r}_{0a})} \simeq O\left(\frac{b}{R}\right) , \tag{7.262}$$

which is of the order of the annual parallax of the source of light. This term can be neglected in the delay formula (7.249) since it produces the delay for extra-solar system objects much less than 1 picosecond. However, for VLBI observations of spacecrafts in the solar system the term under discussion is important and must be taken into account. Note that in taking the difference of the two integrals in (7.249) one can equate $\boldsymbol{k}_2 = \boldsymbol{k}_1$ and $t^*_2 = t^*_1$ in their integrands in correspondence with (7.260) and (7.261). This allows us to perform the integration and reduce the difference to the form

$$\begin{aligned}\int_{s_2}^{s_1} \ln\left[t^*_1 + \frac{1}{c}\boldsymbol{k}_1 \cdot \boldsymbol{x}_a(\zeta) - \zeta\right] \boldsymbol{k}_1 \cdot \dot{\boldsymbol{\beta}}_a(\zeta) d\zeta &= \frac{1}{c}\boldsymbol{k}_1 \cdot \dot{\boldsymbol{\beta}}_a(s_1)\Big\{(r_{2a} - \boldsymbol{k}_1 \cdot \boldsymbol{r}_{2a}) \ln(r_{2a} - \boldsymbol{k}_1 \cdot \boldsymbol{r}_{2a}) \\ &\quad - (r_{1a} - \boldsymbol{k}_1 \cdot \boldsymbol{r}_{1a}) \ln(r_{1a} - \boldsymbol{k}_1 \cdot \boldsymbol{r}_{1a}) + r_{1a} - r_{2a} - \boldsymbol{k}_1(\cdot \boldsymbol{r}_{1a} - \boldsymbol{r}_{2a})\Big\} ,\end{aligned} \tag{7.263}$$

that is, after multiplication by the factor $2Gm_a/c^3$, much less than 1 picosecond even in observation of the solar system objects. Hence, one can drop the two integrals from the expression (7.249) for VLBI delay $\Delta(t_1, t_2) = \Delta(t_2, t_0) - \Delta(t_1, t_0)$.

Finally, taking into account equation (7.258) and denoting $\boldsymbol{k}_1 \equiv -\boldsymbol{K}$ one gets for the time delay

$$\Delta(t_1, t_2) = \frac{2G}{c^3} \sum_{a=1}^{N} m_a (1 + \boldsymbol{K} \cdot \boldsymbol{\beta}_a) \left[\ln \frac{r_{1a} + \boldsymbol{K} \cdot \boldsymbol{r}_{1a}}{r_{2a} + \boldsymbol{K} \cdot \boldsymbol{r}_{2a}} - \frac{\boldsymbol{b} \cdot \boldsymbol{r}_{2a} - (\boldsymbol{K} \cdot \boldsymbol{r}_{2a})(\boldsymbol{b} \cdot \boldsymbol{K})}{R\,(r_{2a} + \boldsymbol{K} \cdot \boldsymbol{r}_{2a})}\right] , \tag{7.264}$$

where $\boldsymbol{\beta}_a = \boldsymbol{v}_a(s_1)/c$, $r_{1a} = |\boldsymbol{r}_{1a}|$, $r_{2a} = |\boldsymbol{r}_{2a}|$, and the null-cone distances

$$\boldsymbol{r}_{1a} = \boldsymbol{x}_1(t_1) - \boldsymbol{x}_a(s_1) , \qquad \boldsymbol{r}_{2a} = \boldsymbol{x}_2(t_2) - \boldsymbol{x}_a(s_2) . \tag{7.265}$$

The second term in square barckets of equation (7.264) is a correction for the annual parallax of the radio source. It is inessential if the radio source is lying outside of the solar system farther than a few parsec. However, precise deep-space navigation of spacecrafts demands to take these finite-distance correction into account in calculating the time delay due to the gravitational field of the Sun.

Let us emphasize that formula (7.264) includes the first-order correction for the velocity of the massive bodies deflecting the light ray. Moreover, there is a difference between the definition of vectors $\boldsymbol{r}_{1a}$, $\boldsymbol{r}_{2a}$ in equation (7.264) and that adopted in the consensus model as defined in [McCarthy and Petit, 2004, Chapter 12]. In the model proposed here, coordinates of stations $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ are taken at the instants t_1, t_2 respectively, and coordinates of the light-ray deflecting bodies are calculated at the retarded times s_1, s_2 defined in equations (7.250), (7.251) which are null-cone equations of the linearized theory of general relativity. The consensus model accepts that coordinates of stations are taken also at the instants t_1, t_2 but coordinates of the a-th body are calculated at the time of the closest approach, t^*_{1a}, defined in equation (7.254). This postulate was introduced in [Hellings, 1986a,b; Treuhaft and Lowe, 1991] on the basis of intuitive guess. Such a guess gives a rather good approximation for practical purposes but can not be adopted as a self-consistent theoretical recommendation as it obscures the causal nature of gravity associated with its null-cone characteristics. This null-cone structure of the gravitational field is currently a matter of experimental testing by making use of precise VLBI astrometry [Fomalont and Kopeikin, 2008, 2003; Kopeikin, 2001].

Let us omit the parallactic term in the right side of equation (7.264) and denote by Δt_{grav} the VLBI delay in the consensus model described in the IERS conventions [McCarthy and Petit, 2004, Chapter 12] and [Kaplan, 1998, equation 5], where one takes the PPN parameter $\gamma = 1$. Then, one gets a general-relativistic difference between the Lorentz-covariant expression for the time delay in the model under discussion and the expression Δt_{grav} in the consensus model:

$$\begin{aligned}\Delta(t_1,t_2) &= \Delta t_{grav} + \frac{2G}{c^3}\sum_{a=1}^{N} m_a\,(\boldsymbol{K}\cdot\boldsymbol{\beta}_a)\ln\frac{r_{1a}+\boldsymbol{K}\cdot\boldsymbol{r}_{1a}}{r_{2a}+\boldsymbol{K}\cdot\boldsymbol{r}_{2a}} \\ &- \frac{2G}{c^3}\sum_{a=1}^{N} m_a\frac{(\boldsymbol{\beta}_a\times\boldsymbol{r}_{1a})(\boldsymbol{b}\times\boldsymbol{r}_{1a})}{r^3_{1a}} + \ldots\,,\end{aligned}\tag{7.266}$$

where ellipsis denote the residual terms. One can easily evaluate that the third term in the right side of equation (7.266) is so small that can be neglected for any observational configuration of the source of light and the deflecting body including the Earth. Expansion of the second term in the right side of equation (7.266) with respect to powers b/d_a, where d_a is the impact parameter of the light ray with respect to the a-th light-deflecting body, yields

$$\begin{aligned}2\frac{Gm_a}{c^3}(\boldsymbol{K}\cdot\boldsymbol{\beta}_a)\ln\frac{r_{1a}+\boldsymbol{K}\cdot\boldsymbol{r}_{1a}}{r_{2a}+\boldsymbol{K}\cdot\boldsymbol{r}_{2a}} &= -2\frac{Gm_a}{c^3}(\boldsymbol{K}\cdot\boldsymbol{\beta}_a)\frac{\boldsymbol{b}\cdot(\boldsymbol{n}_{1a}+\boldsymbol{K})}{r_{1a}+\boldsymbol{K}\cdot\boldsymbol{r}_{1a}} \\ &= -4\frac{Gm_a}{c^3}(\boldsymbol{K}\cdot\boldsymbol{\beta}_a)\frac{\boldsymbol{b}\cdot(\boldsymbol{n}_{1a}+\boldsymbol{K})}{d_a}\frac{r_{1a}}{d_a}\,,\end{aligned}\tag{7.267}$$

where the unit vector $\boldsymbol{n}_{1a} = \boldsymbol{r}_{1a}/r_{1a}$. This term is approximated as $(Gm_a/c^3)(v_a/c)(b/r_{1a})$ and is much smaller than 1 picosecond for any possible configuration of the radio source and the gravitating body inside the solar system including the Sun. Hence, the conclusion is that the relativistic model of VLBI data processing adopted in the IERS conventions [McCarthy and Petit, 2004] is practically good enough for applications at the level of accuracy about one picosecond for astronomical objects with negligibly small parallaxes.

7.7.4
Relativistic space astrometry

Space astrometry is a new branch of fundamental astrometry. Ground-based optical telescopes working in the interferometric regime may reach the angular resolution not better than 0.001 arcseconds. This limits one's ability to create a fundamental inertial system on the sky [Lieske and Abalakin, 1990] with the accuracy required for better understanding of the laws of orbital and rotational motions of celestial bodies both inside and outside of the solar system. The epoch of the space astrometry began in 1989 when the Hipparcos satellite was successfully launched by Ariane 4 rocket of the European Space Agency on 8 August 1989. Despite the unexpected failure to put the satellite on the intended geostationary orbit at 36,000 km from Earth the astrometric program has been completely fulfilled [Kovalevsky, 1998]. As a result, the new astrometric catalog of all stars up to 13-th stellar magnitude was obtained [Mignard, 1997]. It includes about 120,000 stars and has a precision of around 0.002 arcseconds. Unfortunately, such high precision can not be retained longer than 10 years because of statistical errors in determination of the proper motions of stars. For this reason the second analogous mission having the same or better astrometric accuracy should be launched in a near future to maintain the accuracy of the Hipparcos catalogu.

Rapid industrial development of space technologies allows us to hope that in the next several years the precision of astrometric satellites in the determination of positions, proper motions, and parallaxes of celestial objects will reach a few microarcseconds or even better. Besides, the photometric sensitivity of measuring devices will be substantially improved. The most advanced project of this type is a mission of the European Space Agency dubbed Gaia from an original abbreviation for Galactic Astrometric Interferometer for Astrophysics [Lindegren, 2009; Lindegren and Perryman, 1996]. In the framework of this project positions, proper motion, and parallaxes of about 1000 million stars up to 20 stellar magnitude are to be measured with accuracy better than 100 microarcsecond. It means that practically almost all stars in the Milky Way will be observed and registered. It will give us an essential clue to understanding dynamics of the Milky Way, its origin and evolution. Similar project is under development in the US Naval Observatory [Johnston *et al.*, 2006].

NASA astrometric satellite the SIM Lite Astrometric Observatory was a planned space astrometry telescope [Edberg *et al.*, 2007] which was canceled by NASA in 2010. It would have determined the positions and distances of stars several hundred times more accurately than any previous program. This accuracy would have allowed SIM to determine the distances to stars throughout the galaxy and to probe nearby stars for Earth-sized planets. This breakthrough in capabilities would be possible because SIM would use optical interferometry. Pioneered by Albert Michelson, optical interferometry can fulfill its full potential only outside the distorting effects of Earth's atmosphere. There, it can combine light from two or more telescopes as if they were pieces of a single, gigantic telescope mirror. Developed for use in space with SIM, this technique will eventually lead to the development of telescopes powerful enough to take images of Earth-like planets orbiting distant stars.

Such extremely difficult optical observations can not be processed adequately without taking into account numerous relativistic corrections as explained in paper [Kaplan, 1998]. Indeed, the gravitational deflection of light caused by the Sun is not less than 1 milliarcsecond at any direction in the sky. Major planets yield gravitational deflection of light about 1 microarcsecond at the angular distances from 1 to 90 degrees outside the planet [Brumberg *et al.*, 1990; Turyshev, 2009b]. It is worth emphasizing that the relativistic deflection of light produced by the Earth reaches a maximal value of about 550 microarcseconds and should be accounted for any position of a star with respect to the Earth [Gould, 1993]. In addition, the reduction of astrometric observations made on the moving platform will require an extremely careful consideration of relativistic aberration and classic parallax terms [Klioner, 2003a; Krivov, 1994] in order to reduce the measurements to the solar system barycenter - the point to which the origin of the fundamental inertial reference system is attached. This brief consideration makes it evident that the data processing of astrometric observations from modern space satellites should be fully based on general relativistic conceptions rather than on a classical approach in which the relativistic corrections are considered as additive and are taken into account at the very last stage of the reduction procedure.

Brumberg [Brumberg, 1981] was perhaps, the first who recognized the importance of development of such, self-consistent relativistic approach to the data processing of astrometric observations. It required a new approach to the theory of the reference frames in the solar system that was formulated in paper [Brumberg and Kopejkin, 1989b] and further explored in [Brumberg and Kopejkin, 1989a; Klioner, 2003a; Klioner and Kopeikin, 1992; Kopeikin, 1989a,b, 1990; Kopejkin, 1988a]. A supplementary approach was presented in papers [Damour *et al.*, 1991, 1992, 1993]. One global and several local reference frames around major planets of the solar system have been constructed by solving the Einstein equations for gravitational field. The global frame is the barycentric reference frame of the solar system with the origin at the center of mass of the solar system. As for the local frames, the most important for practical applications is the geocentric frame with the origin at the center of mass of the Earth (geocenter) and the proper reference frame of an observer on the Earth [Kopejkin, 1991a] or at the satellite in case of a space mission [Klioner, 2004]. All reference frames are harmonic that is specified by four differential conditions (7.2). They have been build on the spacetime manifold in such a way that minimizes spurious perturbations of gravitational field caused by inappropriate use of coordinate freedom in general relativity. Proceeding in this way a significant progress was achieved in describing relativistic aberration, parallax, and proper motion corrections [Klioner, 2003a; Klioner and Kopeikin, 1992]. Subsequent work solved the problem of propagation of light rays from distant sources of light to an observer in the non-stationary gravitational field of any isolated astronomical system including emission of gravitational waves [Kopeikin and Korobkov, 2005; Kopeikin *et al.*, 2006; Kopeikin and Mashhoon, 2002; Kopeikin and Schäfer, 1999; Kopeikin *et al.*, 1999].

The astrometric quantity which are a matter of concern in this section, is the direction, $\boldsymbol{s}$, towards the source of light (star, quasar) as measured by observer at the point having coordinates $(t, \boldsymbol{x})$. For simplicity, one supposes that the observer is at rest. This eliminates a multitude of relativistic terms accounting for aberration of light. They can be treated separately and the algorithm for this treatment is well-known [Klioner, 2003a; Klioner and Kopeikin, 1992]. This direction is given by equation (7.109) and can be explicitly written as follows

$$\begin{aligned} s^i(\tau, \boldsymbol{\xi}) = K^i \quad &+ \quad \frac{2G}{c^2}\sum_{a=1}^{N}\frac{m_a}{\sqrt{1-\beta_a^2}}\frac{(1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a)^2}{r_a-\boldsymbol{k}\cdot\boldsymbol{r}_a}\frac{P^i{}_j\, r_a^j}{r_a-\boldsymbol{\beta}_a\cdot\boldsymbol{r}_a} \\ &- \quad \frac{2G}{c^2}\sum_{a=1}^{N}\frac{m_a}{\sqrt{1-\beta_a^2}}\frac{2-\boldsymbol{k}\cdot\boldsymbol{\beta}_a}{r_a-\boldsymbol{\beta}_a\cdot\boldsymbol{r}_a}P^i{}_j\,\beta_a^j \\ &- \quad \frac{2G}{c^2R}\sum_{a=1}^{N}\frac{m_a}{\sqrt{1-\beta_a^2}}\frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a}{r_a-\boldsymbol{k}\cdot\boldsymbol{r}_a}P^i{}_j r_a^j \\ &- \quad \frac{4G}{c^2R}\sum_{a=1}^{N}\frac{m_a}{\sqrt{1-\beta_a^2}}\ln\left(\frac{r_a-\boldsymbol{k}\cdot\boldsymbol{r}_a}{2R}\right)P^i{}_j\beta_a^j + \ldots\,, \end{aligned} \tag{7.268}$$

where positions and velocities of the solar system bodies that deflect the light ray, are calculated at the retarded time $s = t - r_a/c$, $R = |\boldsymbol{x} - \boldsymbol{x}_0|$ is the distance from the source of light to observer, $\beta_a^i = v_a^i/c$, and ellipsis denote residual terms depending on accelerations of the bodies which have been neglected because of their insignificant numerical value in practical observations inside the solar system. Further simplification of equation (7.268) is possible if one remembers that the velocities of bodies, $\boldsymbol{v}_a$, comprising the solar system are small in comparison with the fundamental speed c, and the distance R to to the source of light (star, quasar) is very large compared to the size of the solar system. For this reason, the terms being quadratic with respect to velocity $\boldsymbol{v}_a$ and the last term in the right side of equation (7.268) that is proportional to logarithm, are omitted. It yields

$$\begin{aligned} s^i(\tau, \boldsymbol{\xi}) = K^i \quad &+ \quad \frac{2G}{c^2}\sum_{a=1}^{N} m_a\left(1 - 2\boldsymbol{k}\cdot\boldsymbol{\beta}_a + \boldsymbol{n}_a\cdot\boldsymbol{\beta}_a - \frac{r_a}{R}\right)\frac{\boldsymbol{k}\times(\boldsymbol{n}_a\times\boldsymbol{k})}{r_a-\boldsymbol{k}\cdot\boldsymbol{r}_a} \\ &- \quad \frac{4G}{c^2}\sum_{a=1}^{N} m_a\frac{\boldsymbol{k}\times(\boldsymbol{\beta}_a\times\boldsymbol{k})}{r_a}\,, \end{aligned} \tag{7.269}$$

where the unit vector $\boldsymbol{n}_a = \boldsymbol{r}_a/r_a$, symbol $\times$ denotes the usual Euclidean vector product between two vectors. Equation (7.269) was derived in [Kopeikin and Schäfer, 1999] and independently confirmed by Klioner [Klioner, 2003b] who obtained it by the technique of the Lorentz transformation.

The leading order term in equation (7.269) gives us the well-known expression for the angle of gravitational deflection of light ray in the field of a static, spherically symmetric body. The velocity dependent terms in (7.269) describe small relativistic corrections, which can be important in the data analysis of the high-precision astrometric space missions like SIM. The very last term in the large round brackets in equation (7.269) depends on the distance R between the observer and the source of light. It changes the magnitude of the angle of the gravitational deflection for the celestial objects which are lying close to or within the solar system, and must be taken into account in reducing astrometric observations of such objects.

After the gravitational deflection of light is taken into account in the data reduction procedure, the classic astrometric effects - parallax and aberration of light - must be removed from the observed direction s^i. Parallactic corrections to the observed direction s^i are extracted from the unit vector $\boldsymbol{K}$ by its expansion in powers of the ratio (barycentric distance to observer)/(barycentric distance to a star). Accounting for the aberrational corrections is made by means of the Lorentz transformation of the vector s^i, from the static frame to the new direction observed by a moving observer. It is worth emphasizing that correcting for the aberration of light must be done before correcting for the parallax. In other words, the parallactic and aberrational corrections are not commutative. Other particular details of the parallactic and aberrational calculations can be found in [Klioner, 2003a; Klioner and Kopeikin, 1992].

7.8 Doppler Tracking of Interplanetary Spacecrafts

Doppler tracking of deep space probes is central to spacecraft navigation and many radio science investigations. The most sensitive Doppler observations to date were taken when tracking the Cassini spacecraft using the NASA/JPL Deep Space Network antenna DSS 25, which is a 34 m diameter beam-waveguide dish instrumented with simultaneous X- and K-band uplink and tropospheric scintillation calibration equipment [Armstrong *et al.*, 2008]. Those observations achieved Doppler fractional frequency stability (Doppler frequency fluctuation, $\delta\nu$, divided by center frequency, ν_0) $\delta\nu/\nu_0 \simeq 10^{-15}$ at 1000 s integration time. Examples of the scientific studies done with the Doppler tracking include determinations of planetary masses and mass moments [Rappaport *et al.*, 1997], measurements of planetary atmospheres/ionospheres/rings [Kliore *et al.*, 2004, 2008], studies of the solar wind [Noci *et al.*, 1987], and solar system tests of relativistic gravity [Anderson *et al.*, 2004; Bertotti and Giampieri, 1992; Bertotti *et al.*, 2003; Iess *et al.*, 1999; Shang *et al.*, 2009].

The Doppler tracking of interplanetary spacecrafts is the only method presently available to search for gravitational waves in the milliHertz frequency range (10^{-5} - 0.01 Hz) [Armstrong *et al.*, 1997; Bertotti and Iess, 1985; Bertotti *et al.*, 1999; Estabrook and Wahlquist, 1975; Kopeikin *et al.*, 2007; Smarr *et al.*, 1983]. Several experiments to search for gravitational waves have been carried out so far, for instance, VOYAGER, PIONEER, ULYSSES, GALILEO and MARS-OBSERVER [Armstrong, 2006]. The space-probe CASSINI represents the next step in such gravitational Doppler experiments [Iess and Asmar, 2007]. The CASSINI was launched in October 15, 1997 with a primary goal to study Saturn's atmosphere, magnetic fields, rings and icy moons. The spacecraft carries on board advanced radio-transponders and performed three long (40 days each) dedicated data acquisition runs in 2002, 2003 and 2004 searching for gravitational waves with sensitivity about twenty times better than that having been achieved so far [Armstrong, 2006; Armstrong *et al.*, 2003]. Adequate interpretation of gravitational experiments and detection of gravitational waves require precise knowledge of the Doppler frequency shift caused by the solar system bodies lying near an observer's line of sight to spacecraft as shown in Figure 7.9.

Practically important implementation of the Doppler tracking is the Global Positioning System (GPS) which provides reliable positioning, navigation, and timing services to civilian users on a continuous worldwide basis [U.S. Navigation Center, 2010]. The GPS is made up of three parts: satellites orbiting the Earth; control and monitoring stations on Earth; and the GPS receivers owned by users.

- The space segment consists of a nominal constellation of 24 operating satellites that transmit one-way signals that give the current GPS satellite position and time.
- The control segment consists of worldwide monitor and control stations that maintain the satellites in their proper orbits through occasional command maneuvers, and adjust the satellite clocks. It tracks the GPS satellites, uploads updated navi-

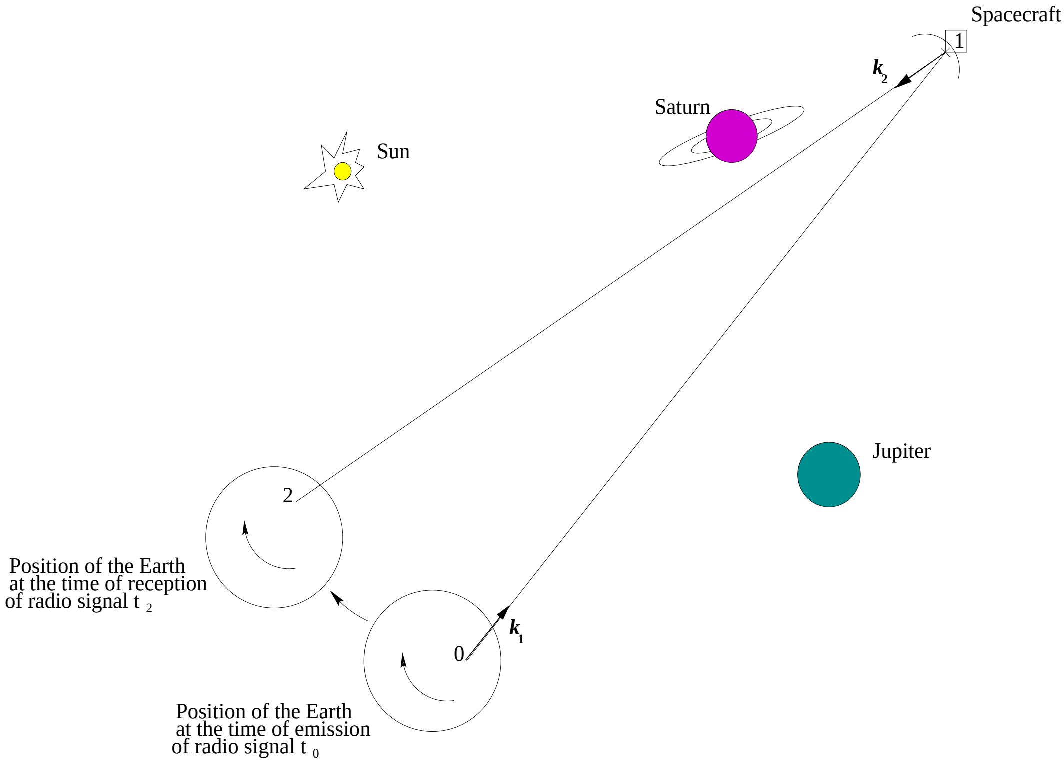

Figure 7.9 Spacecraft Doppler tracking experiment in deep space. Radio signal is transmitted at the time t_0 and at the point $\boldsymbol{x}_0$ on the Earth along a direction defined by the unit vector $\boldsymbol{k}_1$. The radio signal reaches spacecraft at the moment t_1 and at the point $\boldsymbol{x}_1$ and is instantaneously re-emitted back to the Earth along the unit vector $\boldsymbol{k}_2$ which, in general, is not anti-parallel to $\boldsymbol{k}_1$. The signal from spacecraft arrives to a receiver on the Earth located at point $\boldsymbol{x}_2$ at the time t_2. During the time of the round-trip of the radio signal the Earth rotates around its own axis and moves along its orbit. Hence, the barycentric position and velocity of the transmitter is different from the barycentric position and velocity of the receiver but their topocentric positions on the Earth can coincide depending on the Doppler system design. When the impact parameter of the signal's trajectory is small the gravitational Doppler shift of the transmitted frequency with respect to the reference frequency, ν, is roughly estimated as $\delta\nu/\nu = 2\alpha(v_\oplus/c)\cos\varphi$, where α is the gravitational deflection angle of the corresponding light ray, $v_\oplus$ is the orbital velocity of the Earth, and φ is the angle between $v_\oplus$ and the impact parameter of the light ray.

gational data, and maintains health and status of the satellite constellation.

- The user segment consists of the GPS receiver equipment, which receives the signals from the GPS satellites and uses the transmitted information to calculate the user's three-dimensional position and time.

GPS satellites broadcast signals from space that are picked up and identified by GPS receivers. Each GPS receiver then provides three-dimensional location (latitude, longitude, and altitude) plus the time. The GPS uses accurate, stable atomic clocks in satellites and on the ground These clocks have relativistic frequency shifts which are so large that, without accounting for numerous relativistic effects, the system would not function [Ashby, 2002, 2003].

Quite recently, the Europeian Space Agency (ESA) has adopted a new program aimed at achieving an even better precision in measuring time and frequency in spacetime observations. The program is called the Atomic Clock Ensemble in Space (ACES) and will be carried out on board of the International Space Station (ISS). The principal idea is to use a cold atom clock in weightless condition which will outperform the fountains clock on the ground with the potential accuracy of 5×10^{-16} in space [Laurent *et al.*, 2007, 1998]. The European Union and ESA have also decided to build a global navigation satellite system called GALILEO. The project is an alternative and complementary to the GPS and the Russian GLONASS. Current agreement is that it should be operational by 2013. GLONASS is a Russian system that is very similar to GPS. The satellites are at slightly lower altitudes, and orbit the Earth 17 times while the GPS satellites orbit 16 times [Ashby, 2003]. BEIDOU is a satellite navigation system being developed and deployed by the People's Republic of China.

Adequate treatment and further development of the satellite navigation systems, deep space navigation of satellites, and gravitational experiments require advanced analytic algorithm which will properly account for all relativistic terms of the order of 10^{-16} and smaller in the classic and gravitational Doppler shifts between the transmitted and received electromagnetic frequencies caused by the relative motion of the spacecraft with respect to observer and time-dependent gravitational field of the solar system bodies. This section discusses the basic principles of the Doppler tracking observations and gives the most important relationships for calculation of various relativistic effects. However, complete theory of the whole operation of the navigational systems goes beyond the scope of the present book as it includes too many particular details. Hence, the discussion will be focused mainly on the Doppler tracking of spacecrafts in deep space in application to precise gravitational experiments in the field of major planets and the Sun.

7.8.1
Definition and calculation of the Doppler shift

Let us assume (see Figure 7.9) that an electromagnetic signal has been transmitted with frequency ν_0 at the barycentric time t_0 from the point with barycentric coordinates $\boldsymbol{x}_0$ located on the Earth. The signal travels to the interplanetary spacecraft and is received on its board with frequency ν_1 at the barycentric time t_1 at the point with the barycentric coordinates $\boldsymbol{x}_1$. Spacecraft's transponder instantaneously transmits the signal back to the Earth at exactly the same frequency ν_1 where one observes this signal with frequency ν_2 at the barycentric time t_2 at the point with the barycentric coordinates $\boldsymbol{x}_2$. One notices that because of the orbital and rotational motion of the Earth during the light travel time of the signal, the frequency ν_2 will be different from the emitted frequency ν_0 even if the transmitter and receiver are located at the same point on the Earth's surface.

The reading of the clock measuring the proper time of transmitter at the point of emission of a radio signal will be denoted by $\mathcal{T}_0$, and the reading of the same clock at the point of reception of the radio signal will be denoted by $\mathcal{T}_2$. Proper time of spacecraft's transponder will be denoted by $\mathcal{T}_1$. Barycentric coordinate time t at the emission point will be denoted t_0, at the point of reception - t_2, and at the spacecraft's position - t_1. Further calculations are similar to those presented in section 7.5.3. The emitted electromagnetic frequency of the radio signal is ν_0, the frequency received by the spacecraft - ν_1, and the frequency received back on the ground - ν_2. The spectral shift of frequency ν_0 with respect to ν_1 is given by equation

$$1 + z_1 \equiv \frac{\nu_0}{\nu_1} = \frac{d\mathcal{T}_1}{dt_1}\frac{dt_1}{dt_0}\frac{dt_0}{d\mathcal{T}_0} , \tag{7.270}$$

and the shift of frequency, ν_1, with respect to ν_2 is described by a similar relationship

$$1 + z_2 \equiv \frac{\nu_1}{\nu_2} = \frac{dt_1}{d\mathcal{T}_1}\frac{dt_2}{dt_1}\frac{d\mathcal{T}_2}{dt_2} . \tag{7.271}$$

Here the time derivatives $d\mathcal{T}_1/dt_1$ and $dt_1/d\mathcal{T}_1$ are calculated at the spacecraft's position, $dt_0/d\mathcal{T}_0$ is calculated at the point of emission, and calculation of $d\mathcal{T}_2/dt_2$ is done at the point of reception of the radio signal. Time derivatives dt_1/dt_0 and dt_2/dt_1 are two-point functions, which are calculated by differentiation of the solution of equation of propagation of the electromagnetic signal in a time-dependent gravitational field of the solar system (7.72). This equation establishes theoretical description of the transmitter-to-spacecraft uplink, and spacecraft-to-receiver downlink.

In practice, when Doppler tracking observations are made, the frequency ν_2 of the receiver is kept fixed. It relates to the fact that the frequency band of the receiver is made rather narrow to supress the level of environmental radio interference and to increase the sensitivity of receiver to detection of a very weak radio signal transmitted back to the Earth from the spacecraft. On the other hand, technical limitations on the range of the transmitted frequency ν_0 are not so restrictive and it can be changed smoothly in a very broad band according to a prescribed frequency-modulation law. This law of the frequency modulation is chosen to ensure the receiving of the transmitted signal from the spacecraft exactly at the pre-calculated frequency ν_2. It requires to know the precise ephemerides of the transmitter, the observer, and the spacecraft as well as the law of propagation of electromagnetic signal on its round-trip journey. Hence, one needs to calculate the frequency fluctuation

$\delta\nu/\nu_2$, where $\delta\nu = \nu_0 - \nu_2$. From equations (7.270), (7.271) one has

$$\frac{\delta\nu}{\nu_2} = \frac{\nu_0}{\nu_2} - 1 = \frac{dt_0}{d\mathcal{T}_0}\frac{dt_1}{dt_0}\frac{dt_2}{dt_1}\frac{d\mathcal{T}_2}{dt_2} - 1 \,. \tag{7.272}$$

As one can see from this equation there is no need to know explicitly the transformation between the proper time of the spacecraft, $\mathcal{T}_1$, and the barycentric coordinate time of the solar system, t_1, if the spacecraft's transponder has no intrinsic time delay [13].

Accounting for relationship (7.125) and expression (7.12) for the metric tensor yields at the point of emission

$$\frac{dt_0}{d\mathcal{T}_0} = \left[\left(1-\beta_0^2\right)\left(1+\frac{2G}{c^2}\sum_{a=1}^{N}\frac{m_a\sqrt{1-\beta_{a0}^2}}{r_{0a}-\boldsymbol{\beta}_{a0}\cdot\boldsymbol{r}_{0a}}\right)-\frac{4G}{c^2}\sum_{a=1}^{N}\frac{m_a}{\sqrt{1-\beta_{a0}^2}}\frac{(1-\boldsymbol{\beta}_0\cdot\boldsymbol{\beta}_{a0})^2}{r_{0a}-\boldsymbol{\beta}_{a0}\cdot\boldsymbol{r}_{0a}}\right]^{-1/2}, \tag{7.273}$$

where $\boldsymbol{\beta}_0 = \boldsymbol{\beta}_0(t_0)$ is the (normalized to c) barycentric velocity of emitter, $\boldsymbol{\beta}_{a0} = \boldsymbol{\beta}_a(s_0)$ is the (normalized to c) barycentric velocity of the a-th gravitating body taken at a retarded time s_0, $r_{0a} = |\boldsymbol{r}_{0a}|$, $\boldsymbol{r}_{0a} = \boldsymbol{x}_0(t_0) - \boldsymbol{x}_a(s_0)$, and $s_0 = t_0 - r_{0a}/c$ is the retarded time corresponding to the time of emission, t_0, of the radio signal.

Similar arguments give us at the point of reception of the radio signal the next relationship

$$\frac{d\mathcal{T}_2}{dt_2} = \left[\left(1-\beta_2^2\right)\left(1+\frac{2G}{c^2}\sum_{a=1}^{N}\frac{m_a\sqrt{1-\beta_{a2}^2}}{r_{2a}-\boldsymbol{\beta}_{a2}\cdot\boldsymbol{r}_{2a}}\right)-\frac{4G}{c^2}\sum_{a=1}^{N}\frac{m_a}{\sqrt{1-\beta_{a2}^2}}\frac{(1-\boldsymbol{\beta}_2\cdot\boldsymbol{\beta}_{a2})^2}{r_{2a}-\boldsymbol{\beta}_{a2}\cdot\boldsymbol{r}_{2a}}\right]^{1/2}, \tag{7.274}$$

where $\boldsymbol{\beta}_2 = \boldsymbol{\beta}_2(t_2)$ is the (normalized to c) barycentric velocity of emitter, $\boldsymbol{\beta}_{a2} = \boldsymbol{\beta}_a(s_2)$ is the (normalized to c) barycentric velocity of the a-th gravitating body taken at a retarded time s_2, $r_{2a} = |\boldsymbol{r}_{2a}|$, $\boldsymbol{r}_{2a} = \boldsymbol{x}_2(t_2) - \boldsymbol{x}_a(s_2)$, and $s_2 = t_2 - r_{2a}/c$ is the retarded time corresponding to the time, t_2, of the signal's reception.

For up- and down- radio links the time delay relationship (7.129) yields respectively

$$\frac{dt_1}{dt_0} = \frac{1+\boldsymbol{K}_1\cdot\boldsymbol{\beta}_0+\dfrac{2G}{c^2}\displaystyle\sum_{a=1}^{N}m_a\left[\frac{\partial s_1}{\partial t_0}\frac{\partial}{\partial s_1}+\frac{\partial s_0}{\partial t_0}\frac{\partial}{\partial s_0}+\frac{\partial t_1^*}{\partial t_0}\frac{\partial}{\partial t_1^*}+\frac{\partial k_1^i}{\partial t_0}\frac{\partial}{\partial k_1^i}\right]B_a(s_1,s_0,t_1^*,\boldsymbol{k}_1)}{1+\boldsymbol{K}_1\cdot\boldsymbol{\beta}_1-\dfrac{2G}{c^2}\displaystyle\sum_{a=1}^{N}m_a\left[\frac{\partial s_1}{\partial t_1}\frac{\partial}{\partial s_1}+\frac{\partial s_0}{\partial t_1}\frac{\partial}{\partial s_0}+\frac{\partial t_1^*}{\partial t_1}\frac{\partial}{\partial t_1^*}+\frac{\partial k_1^i}{\partial t_1}\frac{\partial}{\partial k_1^i}\right]B_a(s_1,s_0,t_1^*,\boldsymbol{k}_1)}\,, \tag{7.275}$$

13) The intrinsic time delay of the transponder can be introduced for technical purposes but it is not considered here to keep calculations simple,

and

$$\frac{dt_2}{dt_1} = \frac{1 + \boldsymbol{K}_2 \cdot \boldsymbol{\beta}_1 + \dfrac{2G}{c^2} \displaystyle\sum_{a=1}^{N} m_a \left[\frac{\partial s_2}{\partial t_1} \frac{\partial}{\partial s_2} + \frac{\partial s_1}{\partial t_1} \frac{\partial}{\partial s_1} + \frac{\partial t_2^*}{\partial t_1} \frac{\partial}{\partial t_2^*} + \frac{\partial k_2^i}{\partial t_1} \frac{\partial}{\partial k_2^i} \right] B_a(s_2, s_1, t_2^*, \boldsymbol{k}_2)}{1 + \boldsymbol{K}_2 \cdot \boldsymbol{\beta}_2 - \dfrac{2G}{c^2} \displaystyle\sum_{a=1}^{N} m_a \left[\frac{\partial s_2}{\partial t_2} \frac{\partial}{\partial s_2} + \frac{\partial s_1}{\partial t_2} \frac{\partial}{\partial s_1} + \frac{\partial t_2^*}{\partial t_2} \frac{\partial}{\partial t_2^*} + \frac{\partial k_2^i}{\partial t_2} \frac{\partial}{\partial k_2^i} \right] B_a(s_2, s_1, t_2^*, \boldsymbol{k}_2)} . \tag{7.276}$$

Here the retarded time s_1 comes out from the relation $s_1 = t_1 - r_{1a}/c$, where $r_{1a} = |\boldsymbol{x}_1(t_1) - \boldsymbol{x}_a(s_1)|$,

$$\boldsymbol{k}_1 = -\boldsymbol{K}_1 = \frac{\boldsymbol{x}_1(t_1) - \boldsymbol{x}_0(t_0)}{|\boldsymbol{x}_1(t_1) - \boldsymbol{x}_0(t_0)|} , \tag{7.277}$$

$$\boldsymbol{k}_2 = -\boldsymbol{K}_2 = \frac{\boldsymbol{x}_2(t_2) - \boldsymbol{x}_1(t_1)}{|\boldsymbol{x}_2(t_2) - \boldsymbol{x}_1(t_1)|} , \tag{7.278}$$

are the unit vectors which define the direction of propagation of the radio signal in uplink and downlink respectively, and

$$t_1^* = t_0 - \frac{1}{c} \boldsymbol{k}_1 \cdot \boldsymbol{x}_0 , \qquad t_2^* = t_1 - \frac{1}{c} \boldsymbol{k}_2 \cdot \boldsymbol{x}_1 . \tag{7.279}$$

are the times of the closest approach of the radio signal to the origin of the barycentric coordinates of the solar system from uplink and downlink.

7.8.2
The null-cone partial derivatives

The relationships (7.133)-(7.136) allow us to write down the partial time derivatives for the retarded times s_0, s_1, and s_2. One has to carefully distinguish between the derivatives for the up- and down- radio links. For the transmitter-spacecraft uplink one has

$$\frac{\partial s_1}{\partial t_1} = \frac{r_{1a} - \boldsymbol{k}_1 \cdot \boldsymbol{r}_{1a}}{r_{1a} - \boldsymbol{\beta}_{a1} \cdot \boldsymbol{r}_{1a}} - \frac{(\boldsymbol{k}_1 \times \boldsymbol{\beta}_1) \cdot (\boldsymbol{k}_1 \times \boldsymbol{r}_{1a})}{r_{1a} - \boldsymbol{\beta}_{a1} \cdot \boldsymbol{r}_{1a}} , \tag{7.280}$$

$$\frac{\partial s_1}{\partial t_0} = \frac{(1 - \boldsymbol{k}_1 \cdot \boldsymbol{\beta}_0)(\boldsymbol{k}_1 \cdot \boldsymbol{r}_{1a})}{r_{1a} - \boldsymbol{\beta}_{a1} \cdot \boldsymbol{r}_{1a}} , \tag{7.281}$$

$$\frac{\partial s_0}{\partial t_0} = \frac{r_{0a} - \boldsymbol{\beta}_0 \cdot \boldsymbol{r}_{0a}}{r_{0a} - \boldsymbol{\beta}_{a0} \cdot \boldsymbol{r}_{0a}} , \tag{7.282}$$

$$\frac{\partial s_0}{\partial t_1} = 0 . \tag{7.283}$$

These formulas are to be used in equation (7.275). For the spacecraft-receiver downlink one obtains

$$\frac{\partial s_2}{\partial t_2} = \frac{r_{2a} - \boldsymbol{k}_2 \cdot \boldsymbol{r}_{2a}}{r_{2a} - \boldsymbol{\beta}_{a2} \cdot \boldsymbol{r}_{2a}} - \frac{(\boldsymbol{k}_2 \times \boldsymbol{\beta}_2) \cdot (\boldsymbol{k}_2 \times \boldsymbol{r}_{2a})}{r_{2a} - \boldsymbol{\beta}_{a2} \cdot \boldsymbol{r}_{2a}} , \tag{7.284}$$

$$\frac{\partial s_2}{\partial t_0} = \frac{(1 - \boldsymbol{k}_2 \cdot \boldsymbol{\beta}_1)(\boldsymbol{k}_2 \cdot \boldsymbol{r}_{2a})}{r_{2a} - \boldsymbol{\beta}_{a2} \cdot \boldsymbol{r}_{2a}} , \tag{7.285}$$

$$\frac{\partial s_1}{\partial t_1} = \frac{r_{1a} - \boldsymbol{\beta}_1 \cdot \boldsymbol{r}_{1a}}{r_{1a} - \boldsymbol{\beta}_{a1} \cdot \boldsymbol{r}_{1a}} , \tag{7.286}$$

$$\frac{\partial s_1}{\partial t_2} = 0 . \tag{7.287}$$

These formulas are to be used in equation (7.276). One points out that the time derivative (7.280) is different from that given by equation (7.286) although the both derivatives are calculated at one and the same point of space (spacecraft's location). At the first sight it may look surprising. However, if one remembers that the derivative (7.280) is calculated along the transmitter-spacecraft light path and that (7.286) along the spacecraft-receiver light path, which have opposite and slightly different directions, the difference becomes making more sense.

The other set of time derivatives required in the subsequent calculations is as follows,

$$\frac{1}{c}\frac{\partial k_1^i}{\partial t_1} = \frac{(\boldsymbol{k}_1 \times (\boldsymbol{\beta}_1 \times \boldsymbol{k}_1))^i}{R_{01}} , \qquad \frac{1}{c}\frac{\partial k_1^i}{\partial t_0} = -\frac{(\boldsymbol{k}_1 \times (\boldsymbol{\beta}_0 \times \boldsymbol{k}_1))^i}{R_{01}} , \tag{7.288}$$

$$\frac{1}{c}\frac{\partial k_2^i}{\partial t_2} = \frac{(\boldsymbol{k}_2 \times (\boldsymbol{\beta}_2 \times \boldsymbol{k}_2))^i}{R_{21}} , \qquad \frac{1}{c}\frac{\partial k_2^i}{\partial t_1} = -\frac{(\boldsymbol{k}_2 \times (\boldsymbol{\beta}_1 \times \boldsymbol{k}_2))^i}{R_{21}} , \tag{7.289}$$

$$\frac{\partial t_1^*}{\partial t_0} = 1 - \boldsymbol{k}_1 \cdot \boldsymbol{\beta}_0 + \frac{\boldsymbol{\beta}_0 \cdot \boldsymbol{\xi}_1}{R_{01}} , \qquad \frac{\partial t_1^*}{\partial t_1} = -\frac{\boldsymbol{\beta}_1 \cdot \boldsymbol{\xi}_1}{R_{01}} , \tag{7.290}$$

$$\frac{\partial t_2^*}{\partial t_1} = 1 - \boldsymbol{k}_2 \cdot \boldsymbol{\beta}_1 + \frac{\boldsymbol{\beta}_1 \cdot \boldsymbol{\xi}_2}{R_{21}} , \qquad \frac{\partial t_2^*}{\partial t_2} = -\frac{\boldsymbol{\beta}_2 \cdot \boldsymbol{\xi}_2}{R_{21}} , \tag{7.291}$$

where $R_{01} = |\boldsymbol{x}_0 - \boldsymbol{x}_1|$ is the radial distance between the emitter on the Earth and the spacecraft, $R_{21} = |\boldsymbol{x}_2 - \boldsymbol{x}_1|$ is the radial distance between the receiver on the Earth and the spacecraft, and the impact parameters of the uplink and downlink are defined by equations $\boldsymbol{\xi}_1 = \boldsymbol{k}_1 \times (\boldsymbol{x}_1 \times \boldsymbol{k}_1)$, and $\boldsymbol{\xi}_2 = \boldsymbol{k}_2 \times (\boldsymbol{x}_1 \times \boldsymbol{k}_2)$ respectively.

Partial derivatives of functions $B_a(s_1, s_0, t_1^*, k_1^i)$ and $B_a(s_2, s_1, t_2^*, k_2^i)$ can be found by making use of relationships (7.138)-(7.143). This yields

$$\frac{\partial B_a(s_1,s_0,t_1^*,k_1^i)}{\partial s_1} = \frac{c}{\sqrt{1-\beta_{a1}^2}}\frac{(1-\boldsymbol{k}_1\cdot\boldsymbol{\beta}_{a1})^2}{r_{1a}-\boldsymbol{k}_1\cdot\boldsymbol{r}_{1a}}\,, \tag{7.292}$$

$$\frac{\partial B_a(s_1,s_0,t_1^*,k_1^i)}{\partial s_0} = -\frac{c}{\sqrt{1-\beta_{a0}^2}}\frac{(1-\boldsymbol{k}_1\cdot\boldsymbol{\beta}_{a0})^2}{r_{0a}-\boldsymbol{k}_1\cdot\boldsymbol{r}_{0a}}\,, \tag{7.293}$$

$$\frac{\partial B_a(s_1,s_0,t_1^*,k_1^i)}{\partial t_1^*} = -\frac{c}{\sqrt{1-\beta_{a1}^2}}\frac{1-\boldsymbol{k}_1\cdot\boldsymbol{\beta}_{a1}}{r_{1a}-\boldsymbol{k}_1\cdot\boldsymbol{r}_{1a}} + \frac{c}{\sqrt{1-\beta_{a0}^2}}\frac{1-\boldsymbol{k}_1\cdot\boldsymbol{\beta}_{a0}}{r_{0a}-\boldsymbol{k}_1\cdot\boldsymbol{r}_{0a}}\,, \tag{7.294}$$

$$\begin{aligned}\frac{\partial B_a(s_1,s_0,t_1^*,k_1^i)}{\partial k_1^i} = & -\frac{1-\boldsymbol{k}_1\cdot\boldsymbol{\beta}_{a1}}{\sqrt{1-\beta_{a1}^2}}\frac{x_a^j(s_1)}{r_{1a}-\boldsymbol{k}_1\cdot\boldsymbol{r}_{1a}} + \frac{1-\boldsymbol{k}_1\cdot\boldsymbol{\beta}_{a0}}{\sqrt{1-\beta_{a0}^2}}\frac{x_a^j(s_0)}{r_{0a}-\boldsymbol{k}_1\cdot\boldsymbol{r}_{0a}} \\ & +\frac{2v_{a1}^j}{\sqrt{1-\beta_{a1}^2}}\ln(r_{1a}-\boldsymbol{k}_1\cdot\boldsymbol{r}_{1a}) - \frac{2v_{a0}^j}{\sqrt{1-\beta_{a0}^2}}\ln(r_{0a}-\boldsymbol{k}_1\cdot\boldsymbol{r}_{0a})\,,\end{aligned} \tag{7.295}$$

and

$$\frac{\partial B_a(s_2,s_1,t_2^*,k_2^i)}{\partial s_2} = \frac{c}{\sqrt{1-\beta_{a2}^2}}\frac{(1-\boldsymbol{k}_2\cdot\boldsymbol{\beta}_{a2})^2}{r_{2a}-\boldsymbol{k}_2\cdot\boldsymbol{r}_{2a}}\,, \tag{7.296}$$

$$\frac{\partial B_a(s_2,s_1,t_2^*,k_2^i)}{\partial s_1} = -\frac{c}{\sqrt{1-\beta_{a1}^2}}\frac{(1-\boldsymbol{k}_2\cdot\boldsymbol{\beta}_{a1})^2}{r_{1a}-\boldsymbol{k}_2\cdot\boldsymbol{r}_{1a}}\,, \tag{7.297}$$

$$\frac{\partial B_a(s_2,s_1,t_2^*,k_2^i)}{\partial t_2^*} = -\frac{c}{\sqrt{1-\beta_{a2}^2}}\frac{1-\boldsymbol{k}_2\cdot\boldsymbol{\beta}_{a2}}{r_{2a}-\boldsymbol{k}_2\cdot\boldsymbol{r}_{2a}} + \frac{c}{\sqrt{1-\beta_{a1}^2}}\frac{1-\boldsymbol{k}_2\cdot\boldsymbol{\beta}_{a1}}{r_{1a}-\boldsymbol{k}_2\cdot\boldsymbol{r}_{1a}}\,, \tag{7.298}$$

$$\begin{aligned}\frac{\partial B_a(s_2,s_1,t_2^*,k_2^i)}{\partial k_2^i} = & -\frac{1-\boldsymbol{k}_2\cdot\boldsymbol{\beta}_{a2}}{\sqrt{1-\beta_{a2}^2}}\frac{x_a^j(s_2)}{r_{2a}-\boldsymbol{k}_2\cdot\boldsymbol{r}_{2a}} + \frac{1-\boldsymbol{k}_2\cdot\boldsymbol{\beta}_{a1}}{\sqrt{1-\beta_{a1}^2}}\frac{x_a^j(s_1)}{r_{1a}-\boldsymbol{k}_2\cdot\boldsymbol{r}_{1a}} \\ & +\frac{2v_{a2}^j}{\sqrt{1-\beta_{a2}^2}}\ln(r_{2a}-\boldsymbol{k}_2\cdot\boldsymbol{r}_{2a}) - \frac{2v_{a1}^j}{\sqrt{1-\beta_{a1}^2}}\ln(r_{1a}-\boldsymbol{k}_2\cdot\boldsymbol{r}_{1a})\,.\end{aligned} \tag{7.299}$$

One has neglected in formulas (7.294),(7.295) and (7.298), (7.299) all terms depending explicitly on the orbital accelerations of the solar system bodies. If the body moved with constant velocities these equations would be exact.

7.8.3
Doppler effect in spacecraft-planetary conjunctions

The relationships (7.272) - (7.299) constitute the basic elements of the Lorentz-invariant algorithm for the Doppler tracking of spacecrafts in the deep space of the solar system. They are sufficient to calculate the Doppler response for any conceivable relative configuration of transmitter, spacecraft, and the solar system bodies including the effects of special relativity and gravity in the weak field approximation. One is going to consider in this section only the case when spacecraft is behind a massive body (Sun, Jupiter or Saturn) and the impact parameters of up- and down-radio links are small compared with distances from the body to transmitter, receiver, and the spacecraft. Calculation of gravitational shift of frequency will be the primary goal. This approximation is similar to that used in section 7.6.2.2 for calculation of the frequency shift in gravitational lens. Making use of this approximation, one can neglect all terms of the order of Gm_a/c^2r_{0a}, Gm_a/c^2r_{1a}, Gm_a/c^2r_{2a}, Gm_a/c^2R_{01}, m_a/R_{21} as well as the terms that are quadratic with respect to the velocity v_a of the massive body. It is worthwhile to point out that the travel time for a round trip of a radio signal from emitter to spacecraft and back to receiver, is much shorter than the orbital period of any of the solar system body. For this reason, all functions with the retarded time argument entering the equations for calculation of the Doppler shift of frequency can be expanded around the time of transmission of the radio signal from the Earth, which is precisely determined by atomic clocks at the point of the emission. Taking into account these remarks and making use of relationship (7.231) one obtains

$$\left(\frac{\delta\nu}{\nu_2}\right)_{gr} = 2\left(\boldsymbol{\beta}_a - \frac{r_1}{R}\boldsymbol{\beta}_0 - \frac{r_0}{R}\boldsymbol{\beta}_1\right)\cdot\boldsymbol{\alpha}(\boldsymbol{\xi}_a)\,, \qquad \alpha^i(\boldsymbol{\xi}_a) = \frac{4Gm_a}{c^2d_a^2}\,\xi_a^i\,, \tag{7.300}$$

where $\boldsymbol{\beta}_0 = \boldsymbol{v}_0/c$ is the barycentric velocity of the transmitter, $\boldsymbol{\beta}_1 = \boldsymbol{v}_1/c$ is the barycentric velocity of the spacecraft, $\boldsymbol{\beta}_a = \boldsymbol{v}_a/c$ is velocity of the a-th gravitating body deflecting trajectory of the emitted radio signal at the total angle α^i, $d_a = |\boldsymbol{\xi}_a|$ is the length of the impact parameter of the light ray with respect to the a-th body, r_1 is the distance between the transmitter and the light-ray deflecting body, r_0 is the distance between the spacecraft and the light-ray deflecting body, and $R = |\boldsymbol{x}_0 - \boldsymbol{x}_1| \simeq r_0 + r_1$ due to the fact that outgoing and ingoing light rays move along almost the same direction. Formula (7.300) for the Doppler shift by a moving gravitational lens depends on velocities of the transmitter, the spacecraft, and the light-ray deflecting body. It was derived by Kopeikin and Schäfer [1999]. They generalized the result obtained previously by Bertotti and Giampieri [1992] who considered the gravitational shift of frequency only for a static gravitational lens. The effect of motion of a massive body on the gravitational shift of frequency is detectable in the solar system [Kopeikin *et al.*, 2007].

The approximate value of the gravitational shift of frequency of the radio signal on its round trip is determined by the expression $\delta\nu/\nu_2 = 2\alpha_a(v_\oplus/c)\cos\varphi$, where α_a is the total deflection angle of the light ray by a-th massive body, $v_\oplus$ is the magnitude of the orbital velocity of the Earth, and φ is the angle between the Earth's velocity, $\boldsymbol{v}_\oplus$, and the direction of the impact parameter. For the Sun the minimal value of the deflection angle $\alpha_\odot$ is not less than 1 milliarcsecond or $\simeq 5\cdot 10^{-9}$ radians. The ratio of $v_\oplus/c \simeq 10^{-4}$. These numbers demonstrate that the gravitational shift of frequency in the Doppler tracking of interplanetary spacecraft caused by the Sun is not less than $\simeq 5\cdot 10^{-13}$ for any location of the spacecraft in the sky. If the path of the radio link grazes the solar limb, the Doppler shift can reach the value of 10^{-9}. The same kind of evaluation applied to the radio waves grazing the limbs of Jupiter and Saturn yields for the gravitational shift of frequency by Jupiter $8 \cdot 10^{-12}$, and that by Saturn $3 \cdot 10^{-12}$, which can be measured in practice.

Our Lorentz-invariant formalism for derivation of the gravitational shift of frequency in the Doppler tracking of interplanetary spacecrafts can be compared with other approaches based on application of the post-Newtonian approximation scheme [Bertotti and Giampieri, 1992; Kopeikin, 1990]. The advantage of the Lorentz-invariant formalism is that it accounts for all relativistic effects related to velocities of the gravitating bodies through the compact expression for the *Liénard-Wiechert* gravitational potentials. The post-Newtonian scheme is also quite successful but makes calculation longer and presents the result of the calculation in the form of a post-Newtonian expansion [Blanchet *et al.*, 2001b].

7.8.4
The Doppler Effect Revisited

It is instructive from the methodological point of view to compare calculation of the *Doppler effect in terms of frequency* used in the present book, with that done *in terms of energy* (see section 7.5.3 for definition) used, e.g., by Bertotti and Giampieri [1992]. Let us introduce definitions of the four-velocity of observer $u^\alpha = u^0(1, \beta^i)$, the four-velocity of source of light $u_0^\alpha = u_0^0(1, \beta_0^i)$, the four-momentum of photon at the point of emission $\mathcal{K}_0^\alpha = \mathcal{K}_0^0(1, p_0^i)$ and that at the point of observation $\mathcal{K}^\alpha = \mathcal{K}^0(1, p^i)$, where $u^0 = dt/d\mathcal{T}$, $u_0^0 = dt_0/d\mathcal{T}_0$, $p_0^i = c^{-1}\dot{x}^i(t_0)$, $p^i = c^{-1}\dot{x}^i(t)$, $\mathcal{K}_0^0 = dt_0/d\lambda_0$ and $\mathcal{K}^0 = dt/d\lambda$ with λ and λ_0 being values of the affine parameter along the light geodesics at the points of emission and observation. Then, using definition (7.126) of the *Doppler effect in terms of energy*, it is not difficult to show that equation (7.126) can be recast in the form

$$\frac{\nu}{\nu_0} = \frac{u^0\mathcal{K}^0\left\{g_{00}(t,\boldsymbol{x}) + g_{0i}(t,\boldsymbol{x})\left[p^i + \beta^i\right] + g_{ij}(t,\boldsymbol{x})p^i\beta^i\right\}}{u_0^0\mathcal{K}_0^0\left\{g_{00}(t_0,\boldsymbol{x}_0) + g_{0i}(t_0,\boldsymbol{x}_0)\left[p_0^i + \beta_0^i\right] + g_{ij}(t_0,\boldsymbol{x}_0)p_0^i\beta_0^i\right\}} . \tag{7.301}$$

Calculation of the time component $\mathcal{K}^0$ of the four-momentum of photon in equation (7.301) can be done if one knows the relationship of the affine parameter λ along the light geodesic and the coordinate time t. This is found by solution of the time component of the equation for the light geodesic

$$\frac{d^2t}{d\lambda^2} = -c\left(\Gamma^0_{00} + 2\Gamma^0_{0i}p^i + \Gamma^0_{ij}p^ip^j\right)\left(\frac{dt}{d\lambda}\right)^2 . \tag{7.302}$$

Using decomposition (7.1) of the metric tensor and parametrization (7.36) of the unperturbed light ray trajectory, equation (7.302) can be written as

$$\frac{d^2t}{d\lambda^2} = -\left(\frac{1}{2}k^\alpha k^\beta \partial_t h_{\alpha\beta} - k^\alpha \partial_\tau h_{0\alpha}\right)\left(\frac{dt}{d\lambda}\right)^2 , \tag{7.303}$$

where the constant vector $k^\alpha = (1, k^i) = (1, \boldsymbol{k})$, and the substitution for the unperturbed trajectory of the light ray in the metric perturbation $h_{\alpha\beta}$ is done after taking the partial derivative with respect to coordinate time t. Solution of equation (7.303) can be found by iterations using the following expansion

$$\lambda = E^{-1}\left[t + \mathcal{F}(t)\right] , \tag{7.304}$$

where E is the constant photon's energy at past null infinity measured by a fictitious observer being at rest, and function $\mathcal{F}(t)$ is of the order of $h_{\alpha\beta}$. It is obtained by solution of equation (parameter τ was defined in (7.33))

$$\frac{d^2\mathcal{F}}{d\tau^2} = \frac{1}{2}k^\alpha k^\beta \partial_t h_{\alpha\beta} - k^\alpha \partial_\tau h_{0\alpha} , \tag{7.305}$$

which is a direct consequence of equations (7.303), (7.304). Solving the ordinary differential equation (7.305) one finds

$$\mathcal{K}^0 \equiv \mathcal{K}^0(\tau) = E\left[1 - \dot{\mathcal{F}}(\tau)\right] , \tag{7.306}$$

$$\mathcal{K}^0_0 \equiv \mathcal{K}^0(\tau_0) = E\left[1 - \dot{\mathcal{F}}(\tau_0)\right] , \tag{7.307}$$

and

$$\dot{\mathcal{F}}(\tau) = \frac{1}{2}k^\alpha k^\beta \int_{-\infty}^{\tau} \left[\frac{\partial h_{\alpha\beta}(t,\boldsymbol{x})}{\partial t}\right]_{\substack{t=\sigma+t^* \\ \boldsymbol{x}=c\boldsymbol{k}\sigma+\boldsymbol{\xi}}} d\sigma - k^\alpha h_{0\alpha}(\tau) , \tag{7.308}$$

$$\dot{\mathcal{F}}(\tau_0) = \frac{1}{2}k^\alpha k^\beta \int_{-\infty}^{\tau_0} \left[\frac{\partial h_{\alpha\beta}(t,\boldsymbol{x})}{\partial t}\right]_{\substack{t=\sigma+t^* \\ \boldsymbol{x}=c\boldsymbol{k}\sigma+\boldsymbol{\xi}}} d\sigma - k^\alpha h_{0\alpha}(\tau_0) . \tag{7.309}$$

Examination of structure of the integrands in expressions (7.308), (7.309) reveals that

$$\left[\frac{\partial h_{\alpha\beta}(t,\boldsymbol{x})}{\partial t}\right]_{\substack{t=\sigma+t^* \\ \boldsymbol{x}=c\boldsymbol{k}\sigma+\boldsymbol{\xi}}} = \frac{\partial h_{\alpha\beta}(\sigma + t^*, \boldsymbol{k}\sigma + \boldsymbol{\xi})}{\partial t^*} . \tag{7.310}$$

A remarkable property of equality (7.310) is that the parameter t^* is independent from the argument σ of the integrand in (7.308), (7.309) and, for this reason, the partial derivative with respect to t^* is commutative with the integration along the light-ray trajectory. This property allows us to take the partial time derivative in integrals (7.308), (7.309) out of the sign of the integrals. For example, integral (7.308) becomes

$$\int_{-\infty}^{\tau} \left[\frac{\partial h_{\alpha\beta}(t,\boldsymbol{x})}{\partial t}\right]_{\substack{t=\sigma+t^* \\ \boldsymbol{x}=c\boldsymbol{k}\sigma+\boldsymbol{\xi}}} d\sigma = \frac{\partial}{\partial t^*}\int_{-\infty}^{\tau} h_{\alpha\beta}(\sigma + t^*, c\boldsymbol{k}\sigma + \boldsymbol{\xi}) d\sigma . \tag{7.311}$$

Using the *Liénard-Wiechert* solution (7.12) for $h_{\alpha\beta}$ and relationship (7.52) relating the total differentials of the coordinate time σ and the retarded time ζ, one obtains

$$\frac{\partial}{\partial t^*}\int_{-\infty}^{\tau} h_{\alpha\beta}(\sigma + t^*, c\boldsymbol{k}\sigma + \boldsymbol{\xi}) d\sigma = \frac{4G}{c^5}\sum_{a=1}^{N}\left[\frac{\partial}{\partial t^*}\int_{-\infty}^{s(\tau,t^*)} \frac{T^a_{\alpha\beta}(\zeta) - \frac{1}{2}\eta_{\alpha\beta}T^{\lambda}_{a\lambda}(\zeta)}{t^* + c^{-1}\boldsymbol{k}\cdot\boldsymbol{x}_a(\zeta) - \zeta} d\zeta\right] ,$$

(7.312)

where the upper limit $s(\tau, t^*)$ of the integral in the right side of this equation is calculated by means of solution of the gravity null-cone equation, $s+c^{-1}|ck\sigma+\xi-x_a(s)| = \tau + t^*$, and depends on time τ and the instant of the closest approach t^* considered as a parameter. One again emphasizes that among the three variables t, τ, and t^* only two can be considered as independent because of the definition $\tau = t - t^*$ introduced in equation (7.35). The same assertion is valid for the set t_0, τ_0, and t^*.

Because the upper limit of the integral in the right side of equation (7.312) depends on t^*, the derivative $\partial/\partial t^*$ of the integral is taken both from the integrand of the integral and from its upper limit. It yields

$$\frac{\partial}{\partial t^*}\int_{-\infty}^{s(\tau,t^*)}\frac{T^a_{\alpha\beta}(\zeta)-\frac{1}{2}\eta_{\alpha\beta}T^\lambda_{a\lambda}(\zeta)}{t^*+c^{-1}\boldsymbol{k}\cdot\boldsymbol{x}_a(\zeta)-\zeta}d\zeta = \frac{\partial}{\partial t^*}\int_{-\infty}^{s}\frac{T^a_{\alpha\beta}(\zeta)-\frac{1}{2}\eta_{\alpha\beta}T^\lambda_{a\lambda}(\zeta)}{t^*+c^{-1}\boldsymbol{k}\cdot\boldsymbol{x}_a(\zeta)-\zeta}d\zeta + \frac{T^a_{\alpha\beta}(s)-\frac{1}{2}\eta_{\alpha\beta}T^\lambda_{a\lambda}(s)}{r_a(s)-\boldsymbol{k}\cdot\boldsymbol{r}_a(s)}\,\frac{cr_a}{r_a(s)-\boldsymbol{\beta}_a(s)\cdot\boldsymbol{r}_a(s)}\,, \tag{7.313}$$

where the second term in the right hand side is a partial derivative of the upper limit of the integral in equation (7.312) with respect to t^* [14], s is the retarded time defined by the gravity null-cone equation (7.13) treated as independent from t^*. Finally, one has

$$\frac{1}{2}k^\alpha k^\beta\int_{-\infty}^{\tau}\left[\frac{\partial h_{\alpha\beta}(t,\boldsymbol{x})}{\partial t}\right]_{\substack{t=\sigma+t^*\\ \boldsymbol{x}=c\boldsymbol{k}\sigma+\boldsymbol{\xi}}}d\sigma = \frac{1}{2}k^\alpha k^\beta h_{\alpha\beta}(\tau)-\frac{2G}{c^2}\sum_{a=1}^{N}m_aC_a(s) + \frac{2G}{c^2}\sum_{a=1}^{N}\frac{m_a}{\sqrt{1-\beta_a^2}}\frac{(1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a)^2}{r_a-\boldsymbol{k}\cdot\boldsymbol{r}_a}\frac{\boldsymbol{k}\cdot\boldsymbol{r}_a}{r_a-\boldsymbol{\beta}_a\cdot\boldsymbol{r}_a}\,, \tag{7.314}$$

where function $C_a(s)$ is defined in equation (7.116). Similar arguments give

$$\frac{1}{2}k^\alpha k^\beta\int_{-\infty}^{\tau_0}\left[\frac{\partial h_{\alpha\beta}(t,\boldsymbol{x})}{\partial t}\right]_{\substack{t=\sigma+t^*\\ \boldsymbol{x}=c\boldsymbol{k}\sigma+\boldsymbol{\xi}}}d\sigma = \frac{1}{2}k^\alpha k^\beta h_{\alpha\beta}(\tau_0)-\frac{2G}{c^2}\sum_{a=1}^{N}m_aC_a(s_0) + \frac{2G}{c^2}\sum_{a=1}^{N}\frac{m_a}{\sqrt{1-\beta_{a0}^2}}\frac{(1-\boldsymbol{k}\cdot\boldsymbol{\beta}_{a0})^2}{r_{0a}-\boldsymbol{k}\cdot\boldsymbol{r}_{0a}}\frac{\boldsymbol{k}\cdot\boldsymbol{r}_{0a}}{r_{0a}-\boldsymbol{\beta}_{a0}\cdot\boldsymbol{r}_{0a}}\,. \tag{7.315}$$

Going back to formula (7.301) of the *Doppler effect in terms of energy* one can see that it can be factorized in three terms

$$\frac{\nu}{\nu_0} = \mathcal{S}_1\cdot\mathcal{S}_2\cdot\mathcal{S}_3\,, \tag{7.316}$$

14) See formula (7.51) for calculation of the partial derivative of the retarded time s with respect to the parameter t^*.

where

$$\mathcal{S}_1 \equiv \frac{u^0}{u^0_0} = \frac{1-\beta_0^2 - h_{00}(t_0,\boldsymbol{x}_0) - 2h_{0i}(t_0,\boldsymbol{x}_0)\beta_0^i - h_{ij}(t_0,\boldsymbol{x}_0)\beta_0^i\beta_0^j}{1-\beta^2 - h_{00}(t,\boldsymbol{x}) - 2h_{0i}(t,\boldsymbol{x})\beta^i - h_{ij}(t,\boldsymbol{x})\beta^i\beta^j}\,, \tag{7.317}$$

$$\mathcal{S}_2 \equiv \frac{\mathcal{K}^0}{\mathcal{K}^0_0} = \frac{1-\dot{\mathcal{F}}(\tau)}{1-\dot{\mathcal{F}}(\tau_0)}\,, \tag{7.318}$$

$$\mathcal{S}_3 = \frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}-\boldsymbol{\beta}\cdot\dot{\boldsymbol{\Xi}}(\tau) - k^\alpha h_{0\alpha}(t,\boldsymbol{x}) - k^\alpha\beta^j h_{\alpha j}(t,\boldsymbol{x})}{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_0-\boldsymbol{\beta}_0\cdot\dot{\boldsymbol{\Xi}}(\tau_0) - k^\alpha h_{0\alpha}(t_0,\boldsymbol{x}_0) - k^\alpha\beta_0^j h_{\alpha j}(t_0,\boldsymbol{x}_0)}\,. \tag{7.319}$$

Here, functions $\dot{\boldsymbol{\Xi}}(\tau)$ and $\dot{\boldsymbol{\Xi}}(\tau_0)$ must be calculated from equation (7.64) taken at the instants τ and τ_0 respectively.

On the other hand, the above-given result for calculation of the *Doppler shift in terms of frequency* obtained in section 7.5.3 had the following form

$$\frac{\nu}{\nu_0} = \mathcal{S}_1\frac{dt_0}{dt}\,. \tag{7.320}$$

Thus, in order to have an agreement with calculation of the *Doppler shift in terms of energy* one must prove that

$$\frac{dt_0}{dt} = \mathcal{S}_2\cdot\mathcal{S}_3\,. \tag{7.321}$$

One can recast the right side of this equation in another form by accounting for equations (7.308), (7.309), (7.314), (7.315). It yields

$$\mathcal{S}_2\cdot\mathcal{S}_3 = \frac{\mathcal{A}(\tau)}{\mathcal{A}(\tau_0)}\frac{\mathcal{B}(\tau)}{\mathcal{B}(\tau_0)}\,, \tag{7.322}$$

where

$$\mathcal{A}(\tau) = 1-\boldsymbol{k}\cdot\boldsymbol{\beta} - \frac{1}{2}\beta^i k^\alpha k^\beta \hat{\partial}_i B_{\alpha\beta}(\tau) - \frac{1}{2}k^\alpha k^\beta h_{\alpha\beta}(\tau)\,, \tag{7.323}$$

$$\mathcal{A}(\tau_0) = 1-\boldsymbol{k}\cdot\boldsymbol{\beta}_0 - \frac{1}{2}\beta_0^i k^\alpha k^\beta \hat{\partial}_i B_{\alpha\beta}(\tau_0) - \frac{1}{2}k^\alpha k^\beta h_{\alpha\beta}(\tau_0)\,, \tag{7.324}$$

$$\mathcal{B}(\tau) = 1+\frac{2G}{c^2}\sum_{a=1}^{N}\left[m_a C_a(s) - \frac{m_a}{\sqrt{1-\beta_a^2}}\frac{(1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a)^2}{r_a-\boldsymbol{k}\cdot\boldsymbol{r}_a}\frac{\boldsymbol{k}\cdot\boldsymbol{r}_a}{r_a-\boldsymbol{\beta}_a\cdot\boldsymbol{r}_a}\right], \tag{7.325}$$

$$\mathcal{B}(\tau_0) = 1+\frac{2G}{c^2}\sum_{a=1}^{N}\left[m_a C_a(s_0) - \frac{m_a}{\sqrt{1-\beta_{a0}^2}}\frac{(1-\boldsymbol{k}\cdot\boldsymbol{\beta}_{a0})^2}{r_{0a}-\boldsymbol{k}\cdot\boldsymbol{r}_{0a}}\frac{\boldsymbol{k}\cdot\boldsymbol{r}_{0a}}{r_{0a}-\boldsymbol{\beta}_{a0}\cdot\boldsymbol{r}_{0a}}\right] \tag{7.326}$$

where the partial derivatives $\hat{\partial}_i B_{\alpha\beta}(\tau)$ and $\hat{\partial}_i B_{\alpha\beta}(\tau_0)$ are calculated as defined in equation (7.58). With equations (7.322) - (7.326) it is straightforward to confirm that equation (7.321) is true, if one notes that up to the second order of the post-Minkowskian approximation scheme

$$\mathcal{B}^{-1}(\tau) = 1-\frac{2G}{c^2}\sum_{a=1}^{N}\left[m_a C_a(s) - \frac{m_a}{\sqrt{1-\beta_a^2}}\frac{(1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a)^2}{r_a-\boldsymbol{k}\cdot\boldsymbol{r}_a}\frac{\boldsymbol{k}\cdot\boldsymbol{r}_a}{r_a-\boldsymbol{\beta}_a\cdot\boldsymbol{r}_a}\right], \tag{7.327}$$

so that equation (7.322) can be re-written as follows

$$\mathcal{S}_2 \cdot \mathcal{S}_3 = \frac{\mathcal{A}(\tau)}{\mathcal{A}(\tau_0)\mathcal{B}^{-1}(\tau)\mathcal{B}(\tau_0)} . \tag{7.328}$$

Farther straightforward calculations confirm that the numerator and denominator in equation (7.328) coincide precisely with those of equation (7.146) used for calculation of the *Doppler shift in terms of frequency* and, for this reason, equation (7.321) is valid. This finalizes the proof of the equivalence of the two mathematical techniques for calculation of the Doppler effect in gravitational field of moving bodies.

In conclusion of this section one would like to point out that the method of calculation of integrals in formulas (7.308), (7.309) explained in a sequence of equations (7.310) - (7.315) can be successfully applied for calculation of the red shift of the cosmic microwave background radiation (CMBR) caused by a cosmic string moving with ultra-relativistic velocity [Stebbins, 1988].

7.8.5
The Explicit Doppler Tracking Formula

In view of practical applications it is useful to give the explicit formula that can be used for the Doppler tracking navigation of satellites in deep space and for high-precise gravitational experiments. It is derived in the present section for one-way propagation of an electromagnetic signal emitted at point $\boldsymbol{x}_0$ at time t_0, and received at point $\boldsymbol{x}$ at time t. The Doppler shift of the observed frequency ν with respect to the emitted frequency ν_0 is given by equation (7.316) which is convenient to transformed to another form to separate the effects of special and general relativity. Formula (7.316) is factorized as follows

$$\frac{\nu}{\nu_0} = \frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}}{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_0}\left[\frac{1-\beta_0^2}{1-\beta^2}\right]^{1/2}\left[\frac{\mathrm{a}(\tau_0)}{\mathrm{a}(\tau)}\right]^{1/2}\frac{\mathrm{b}(\tau)}{\mathrm{b}(\tau_0)} , \tag{7.329}$$

where the first two terms in the right side describe the special relativistic Doppler effect, and the next two terms give the general relativistic contribution. The unit vector $\boldsymbol{k}$ is given at past null infinity and relates to the unit vector $\boldsymbol{K}$ of the boundary value problem through the relationship (7.67) which has the following explicit form

$$\begin{aligned} k^i &= -K^i + \frac{2G}{c^2R}\sum_{a=1}^{N} m_a\left[\frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a}{\sqrt{1-\beta_a^2}}\frac{r_a^i - k^i(\boldsymbol{k}\cdot\boldsymbol{r}_a)}{r_a-\boldsymbol{k}\cdot\boldsymbol{r}_a} - \frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_{a0}}{\sqrt{1-\beta_{a0}^2}}\frac{r_{0a}^i - k^i(\boldsymbol{k}\cdot\boldsymbol{r}_{0a})}{r_{0a}-\boldsymbol{k}\cdot\boldsymbol{r}_{0a}}\right] \\ &+ \frac{4G}{c^2R}\sum_{a=1}^{N} m_a\left[\frac{P^i_j\beta^j_a}{\sqrt{1-\beta_a^2}}\ln(r_a-\boldsymbol{k}\cdot\boldsymbol{r}_a) - \frac{P^i_j\beta^j_{a0}}{\sqrt{1-\beta_{a0}^2}}\ln(r_{0a}-\boldsymbol{k}\cdot\boldsymbol{r}_{0a})\right] , \end{aligned} \tag{7.330}$$

where $R = |\boldsymbol{x}-\boldsymbol{x}_0|$ is the distance between the point of emission and observation.

Explicit expressions for functions $\mathrm{a}(\tau)$ and $\mathrm{a}(\tau_0)$ are derived using equation (7.273) which leads to

$$
\begin{aligned}
\mathrm{a}(\tau) &= 1 + \frac{2G}{c^2}\sum_{a=1}^{N}\frac{m_a\sqrt{1-\beta_a^2}}{r_a - \boldsymbol{\beta}_a\cdot\boldsymbol{r}_a} \\
&- \frac{4G}{c^2}(1-\beta^2)^{-1}\sum_{a=1}^{N}\frac{m_a}{\sqrt{1-\beta_a^2}}\frac{(1-\boldsymbol{\beta}\cdot\boldsymbol{\beta}_a)^2}{r_a-\boldsymbol{\beta}_a\cdot\boldsymbol{r}_a}\,,
\end{aligned}
\tag{7.331}
$$

$$
\begin{aligned}
\mathrm{a}(\tau_0) &= 1 + \frac{2G}{c^2}\sum_{a=1}^{N}\frac{m_a\sqrt{1-\beta_{a0}^2}}{r_{0a} - \boldsymbol{\beta}_{a0}\cdot\boldsymbol{r}_{0a}} \\
&- \frac{4G}{c^2}(1-\beta_0^2)^{-1}\sum_{a=1}^{N}\frac{m_a}{\sqrt{1-\beta_{a0}^2}}\frac{(1-\boldsymbol{\beta}_0\cdot\boldsymbol{\beta}_{a0})^2}{r_{0a}-\boldsymbol{\beta}_{a0}\cdot\boldsymbol{r}_{0a}}\,.
\end{aligned}
\tag{7.332}
$$

Let us recall that $\boldsymbol{v}_0 = \boldsymbol{v}(t_0)$ is the barycentric velocity of emitter, $\boldsymbol{v}_{a0} = \boldsymbol{v}_a(s_0)$ is the barycentric velocity of a-th gravitating body at the retarded time $s_0 = t_0 - r_{0a}$, $r_{0a} = |\boldsymbol{r}_{0a}|$, $\boldsymbol{r}_{0a} = \boldsymbol{x}_0(t_0) - \boldsymbol{x}_a(s_0)$. Besides, $\boldsymbol{v} = \boldsymbol{v}(t)$ is the barycentric velocity of receiver, $\boldsymbol{v}_a = \boldsymbol{v}_a(s)$ is the barycentric velocity of the a-th gravitating body at the retarded time $s = t - r_a/c$, $r_a = |\boldsymbol{r}_a|$, $\boldsymbol{r}_a = \boldsymbol{x}(t) - \boldsymbol{x}_a(s)$.

Omitting all acceleration-dependent terms in equation (7.119) for the integral C_a, and reducing similar terms, one obtains functions in the last factor of the basic relationship (7.329) the following explicit result

$$
\begin{aligned}
\mathrm{b}(\tau) &= 1 + \frac{2G}{c^2}\sum_{a=1}^{N}\frac{m_a}{\sqrt{1-\beta_a^2}}\frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a}{r_a-\boldsymbol{\beta}_a\cdot\boldsymbol{r}_a} \\
&\times\left[\frac{(1-\boldsymbol{k}\cdot\boldsymbol{\beta}_a)(\boldsymbol{k}\times\boldsymbol{\beta})\cdot(\boldsymbol{k}\times\boldsymbol{r}_a)}{r_a-\boldsymbol{k}\cdot\boldsymbol{r}_a} - \frac{(\boldsymbol{k}\times\boldsymbol{\beta}_a)\cdot(\boldsymbol{k}\times\boldsymbol{r}_a)}{r_a-\boldsymbol{k}\cdot\boldsymbol{r}_a} + \boldsymbol{k}\cdot\boldsymbol{\beta}_a\right],
\end{aligned}
\tag{7.333}
$$

$$
\begin{aligned}
\mathrm{b}(\tau_0) &= 1 + \frac{2G}{c^2}\sum_{a=1}^{N}\frac{m_a}{\sqrt{1-\beta_{a0}^2}}\frac{1-\boldsymbol{k}\cdot\boldsymbol{\beta}_{a0}}{r_{0a}-\boldsymbol{\beta}_{a0}\cdot\boldsymbol{r}_{0a}} \\
&\times\left[\frac{(1-\boldsymbol{k}\cdot\boldsymbol{\beta}_{a0})(\boldsymbol{k}\times\boldsymbol{\beta}_0)\cdot(\boldsymbol{k}\times\boldsymbol{r}_{0a})}{r_{0a}-\boldsymbol{k}\cdot\boldsymbol{r}_{0a}} - \frac{(\boldsymbol{k}\times\boldsymbol{\beta}_{a0})\cdot(\boldsymbol{k}\times\boldsymbol{r}_{0a})}{r_{0a}-\boldsymbol{k}\cdot\boldsymbol{r}_{0a}} + \boldsymbol{k}\cdot\boldsymbol{\beta}_{a0}\right].
\end{aligned}
\tag{7.334}
$$

Formulas (7.329) -(7.334) describe the Doppler shift of the radio signal transmitted from observer to spacecraft. The Doppler shift of the radio signal transmitted back to the observer is described by a similar set of equations with assignment to all quantities instants of the signal's reflection from the spacecraft and its reception correspondingly. In case when the light ray grazes limb of a gravitating body, formula (7.329) gives, of course, the result shown already in equation (7.300).

7.9 Astrometric Experiments with the Solar System Planets

7.9.1 Motivations

Attaining the level of a microarcsecond (μas) positional accuracy and better will completely revolutionize fundamental astrometry by merging it with relativistic gravitational physics. Beyond the microarcsecond threshold, one will be able to observe a new range of celestial physical phenomena caused by gravitational waves from the early universe and various localized astronomical sources, spacetime topological defects, moving gravitational lenses, time variability of gravitational fields of super-massive binary black holes located in quasars, and many others [Kopeikin and Gwinn, 2000; Kopeikin and Makarov, 2006; Kopeikin *et al.*, 1999]. Furthermore, this will allow us to test general theory of relativity in the solar system in a dynamic regime, that is in case when velocity- and acceleration-dependent components of gravitational field (the metric tensor) of the Sun and major planets give rise to observable relativistic effects in the light deflection, time delay and frequency shift. Some of these effects have been already observed [Fomalont and Kopeikin, 2008, 2003; Kopeikin and Fomalont, 2002; Kopeikin *et al.*, 2007; Kopeikin, 2009] and will be discussed later in more detail.

Early calculations [Brumberg *et al.*, 1990; Schuh *et al.*, 1988] revealed that major planets of the solar system are sufficiently massive to make flybys of photon sensitive to their gravitational pull. It was also noticed [Brumberg *et al.*, 1990; French *et al.*, 1993; Hubbard *et al.*, 1993; Klioner, 1991b; Klioner and Kopeikin, 1992; Turyshev, 2009b] that gravitational field of the planets have a complicated multipolar structure so that besides the monopole component of the field, the quadrupole, J_2, and higher-order multipoles can deflect light at the level of a few microarcseconds [15], in contrast to the Sun whose quadrupole moment is only $J_{2\odot} \leq 2.3 \times 10^{-7}$ [Pijpers, 1998; Pireaux *et al.*, 2007; Pitjeva, 2005b]. Moreover, in the case of photon propagating near planet the interaction between the gravitational field and the photon can no longer be considered as static, because the planet moves around the Sun as the photon traverses across the solar system [Kopeikin and Mashhoon, 2002; Kopeikin and Schäfer, 1999]. The optical interferometer designed for the space astrometric mission SIM [Edberg *et al.*, 2007; SIM, 2010] is capable of observing background optical sources fairly close to planetary limbs with a microarcsecond accuracy. Similar resolution can be achieved for radio sources (quasars) with the Square Kilometer Array (SKA) [Carilli and Rawlings, 2004; Dewdney *et al.*, 2009] if it is included to the inter-continental baseline network of VLBI stations [Fomalont and Reid, 2004]. The European space astrometry mission Gaia [Lindegren, 2009; Lindegren *et al.*, 1995; Lindegren and Perryman, 1996] and the OBSS [Johnston *et al.*, 2006] astrometric project represent another alternative path to microarcsecond astrometry

15) The most recent data on physical and geometric parameters of major planets of the solar system are available at NASA JPL website http://ssd.jpl.nasa.gov/?gravity_fields_op

[Klioner, 2003a; Vecchiato *et al.*, 2003].

New generation of microarcsecond astrometry satellites: SIM and a cornerstone mission of ESA - Gaia, requires a novel approach for an unambiguous interpretation of astrometric data obtained from the on-board optical instruments. SIM and Gaia complement one another. Both SIM and Gaia have a potential to approach the accuracy of 1 μas. Gaia will observe all stars ($\sim 10^9$) between magnitude 6 and 20. The accuracy of Gaia is about 5 μas for the optimal stars (magnitude between 6 and 13). SIM was planned to observe 10000 stars with magnitude up to 20. The accuracy of SIM was expected to be a few μas for any object brighter than about 20 provided that sufficient observing time would be allocated for that object. At this level the problem of propagation of light rays must be treated with taking into account relativistic effects generated by non-static part of the gravitational field of the solar system and binary stars [Kopeikin and Gwinn, 2000]. It is a challenge for the SIM and SKA interferometers as well as for Gaia and OBSS to measure the gravitational bending of light caused by various planetary multipoles and the orbital motion of the planets. This measurement, if successful, will be a cornerstone step in further deployment of theoretical principles of general relativity to fundamental astrometry and the solar system navigation at a new, exciting technological level.

The first attempt to observe gravitational deflection of radio waves by a major planet - Jupiter - was undertaken by H. Schuh et al. [Schuh *et al.*, 1988; Yoshino *et al.*, 1989] when quasar P0201+113 passed within 200 arcsec of Jupiter on March 21, 1988. An expected apparent slight position change during the approach of P0201+113 to Jupiter was not been clearly detected, perhaps due to insufficient accuracy of the source position. The other team from NASA JPL was conducting the same experiment with the technique of differential VLBI, over two Deep Space Network (DSN) California-Australia baselines and it was more lucky [Treuhaft and Lowe, 1991]. The deflection term associated with the monopole field of Jupiter was approximately 300 μas and it was determined to an accuracy of $\simeq 15\%$ in agreement with Einstein's general relativity theory. Later on, the Hubble Space Telescope was used to measure the gravitational deflection of light of the bright star HD 148898 as it passed within a few seconds of arc near Jupiter's limb on 24 September 1995 [Whipple *et al.*, 1996]. S. Kopeikin [Kopeikin, 2001] proposed to use Jupiter's orbital motion in order to measure the retardation effect in the time of propagation of gravitational force from Jupiter to observed photon. The retardation effect of gravity appears in observations as a small excess to the static Shapiro time delay due to the gravitomagnetic change in the direction of the gravitational force exerted on photon by the planet as it moves along its orbit. The proposal was executed experimentally with Very Long Baseline Array (VLBA) in 2002 September 8, and the speed of propagation of gravity (as compared with the speed of light) was measured to $\simeq 20\%$ accuracy [Fomalont and Kopeikin, 2003]. This measurement is in agreement with general relativity. Physical interpretation of the gravity propagation in the light-ray deflection experiments has several aspects emphasizing various parts of general relativity involved to the experiment that were summarized by Kopeikin and Fomalont [2006]. There are also authors who believe that the gravitational light-ray deflection experiments with moving massive planets the speed of light from the quasar [Will, 2006]. This point of view reflects misunderstanding of the underlying principles of general relativity and the way how it relates to special relativity [16]. Section 7.9.4.3 comments on the "speed of light" interpretation.

The speed-of-gravity experiment conducted in the gravity field of moving Jupiter stimulated researchers to take a next step in exploring gravitational bending of light by the major planets. Most notably, Crosta & Mignard [Crosta and Mignard, 2006] proposed to measure the deflection of light associated with the axisymmetric quadrupolar part of Jupiter's gravitational field. Their work was aimed at converting the earlier theoretical calculations [French *et al.*, 1993; Hubbard *et al.*, 1993; Klioner, 1991b; Kopeikin, 1997c; Le Poncin-Lafitte and Teyssandier, 2005] of light bending by gravitational multipoles into a practical algorithm for Gaia data processing, thus, extending the relativistic techniques of astrometric data reduction having been worked out in a number of previous papers [de Felice *et al.*, 2004, 2006; Klioner, 2003a; Klioner and Kopeikin, 1992; Kopeikin and Mashhoon, 2002; Kopeikin and Schäfer, 1999]. It is worth noting that the quadrupolar field of planets and their natural satellites can be independently measured by the Doppler tracking of spacecraft orbiting the planet [Rappaport *et al.*, 1997]. The quadrupolar field modulates a microwave beam transmitted from a spacecraft orbiting the planet, thus, making the received frequency of the beam a complicated function of the planetary quadrupole deformation that can be measured. This effect is related not to the process of propagation of the radio signal but to the effect of gravitational red shift originating from the transformation from the proper time of the spacecraft to the barycentric time of the solar system that are described by equation (3.246). Detection and precise measurement of the quadrupolar deflection of light by a planet of the solar system differs from the Doppler tracking technique and is important for providing an independent experimental support for the theory of cosmological gravitational lenses formed by a galaxy or a cluster of galaxies. The cluster's gravitational potential (that includes the invisible dark matter) is reconstructed from the observed distortion of images of background quasars under assumption that the multipolar field of the gravitational lens deflects light exactly as predicted in general relativity [Courbin *et al.*, 2002; Schneider *et al.*, 1992]. This is the best assumption for doing physically meaningful interpretation of observational data but its experimental confirmation is highly needed and will be crucial for getting an unbiased estimate of the amount of dark matter in the universe.

The work [Crosta and Mignard, 2006] was substantially extended in several directions in paper by Kopeikin and Makarov [Kopeikin and Makarov, 2007]. Crosta & Mignard [Crosta and Mignard, 2006] assumed that star light propagates in the static field of Jupiter, thus, ignoring that the planet moves along its orbit as the light traverses the solar system toward observer. A question arises about what position of Jupiter is to be used for calculation of the gravitational light deflection in the data processing algorithm. Another question is about the displacement of the center of

16) It is common to refer to the fundamental speed c in Einstein's field equations as to the speed of light. However, general relativity is a fundamental theory about the nature of gravity and has nothing to do with the physical speed of light. The speed c in general relativity is the speed of gravity!

mass of Jupiter with respect to the origin of the inertial coordinates in the sky used for measuring the stellar positions that are affected by the gravitational deflection of light by Jupiter. An assumption that the center of mass of the planet coincides with the origin of the coordinates makes the dipole moment, I^i, of the gravitational field of Jupiter vanish identically [17)], which definitely simplifies theoretical calculation of the light bending. However, the position of Jupiter's center of mass is not known precisely but with some error due to the finite accuracy of the Jupiter's celestial ephemeris currently limited to a few hundred kilometers [Pireaux *et al.*, 2007; Pitjeva, 2005a,b]. This inaccuracy in the position of Jupiter should be taken into account in data processing of microarcsecond astrometric observations by including a non-zero value of Jupiter's dipole moment to the multipolar expansion of the gravitational field of the planet. In other words, any realistic set of measurements can only be adequately interpreted within a certain model of the relativistic deflection of light, which includes all parameters accounting for a possible shift of the true position of the planet's center of mass from the origin of the local inertial frame used for theoretical calculation of the light deflection. The dipole moment can introduce a spurious deflection of light which must be understood and sorted out. The problem is that the data processing algorithm can reduce the dipole moment of the gravitational field to zero, if and only if, general relativity is compatible with special relativity, that is the speed of gravity (the fundamental speed in general relativity) equals the speed of the Minkowski spacetime [Kopeikin and Fomalont, 2007; Kopeikin and Makarov, 2007]. Thus, the problem of measuring the dipolar light deflection component acquires a great significance of testing this fundamental principle.

If the effect of the gravitational dipole is not properly removed from observations it will forge a quadrupolar deflection of light because of the change in the planetary moments of inertia due to the effect known in physics as the parallel-axis or Steiner's theorem [Arnold, 1995] (see equation (7.350) below). This translation-induced quadrupolar distortion of the light-ray deflection pattern of the background stars should be clearly discerned from that caused by the intrinsic quadrupole moment of the planet J_2. One is also going to discuss theoretical and observational aspects of monopolar, dipolar, and quadrupolar light-ray deflections and to investigate how the spurious deflections can be separated from the physical ones caused by the intrinsic quadrupole moment of planet and how to use the measurement of the dipolar anisotropy to test the relativistic effects caused by the time-dependent component of the gravitational field deflecting light rays. One will employ the mathematical technique similar to that having been used for calculation of light propagation in the field of the *Liénard-Wiechert* potentials of point-like masses. This technique will be extended to the case of deflection of light by higher-order multipoles.

One further stipulates that general relativity is valid, which implies that, at least in linearized approximation, gravity operates on the null cone and the force of gravitational interaction propagates on the spacetime manifold with the fundamental speed c [Low, 1999]. It is suggested that each photon incoming to the solar system from a source of light propagates in vacuum, and its physical speed is equal to c. It simplifies

17) That is, $I^i = 0$ at any instant of time.

calculations and avoids discussion of the dispersive effects in plasma. One adopts, as in any other part of the book, that Latin indices take values 1,2,3, and the Greek ones run from 0 to 3. The Kronecker symbol (a diagonal unit matrix) is denoted $\delta_{ij} = \mathrm{diag}(1,1,1)$, and the fully anti-symmetric symbol of Levi-Civita, ε_{ijk}, is defined in such a way that $\varepsilon_{123} = +1$. The Minkowski metric is $\eta_{\alpha\beta} = \mathrm{diag}(-1,1,1,1)$. Greek indices are raised and lowered with the Minkowski metric $\eta_{\alpha\beta}$. By convention, Latin indices are raised and lowered with the Kronecker symbol δ_{ij} which makes no difference between super- and sub-script Latin indices. Repeated indices indicate the Einstein summation rule. Bold letters denotes a spatial vector, for instance $\boldsymbol{x} = x^i = (x^1, x^2, x^3)$. A dot or a cross between two spatial vectors denote the Euclidean scalar or vector products respectively: $\boldsymbol{A}\cdot\boldsymbol{B} = A_i B^i = A_i B_i = A^i B^i$, and $\boldsymbol{A}\times\boldsymbol{B} = (\boldsymbol{A}\times\boldsymbol{B})^i = \varepsilon_{ijk}A^j B^k$. Angular brackets around a pair of the Latin indices of a spatial tensor of rank two denote its symmetric and trace-free (STF) part [Thorne and Hartle, 1985], for example,

$$I^{<ij>} \equiv \frac{1}{2}\left(I^{ij} + I^{ji}\right) - \frac{1}{3}\delta^{ij} I^{kk}\ , \tag{7.335}$$

where $I^{kk} = \delta^{kp} I^{kp}$ (see Section 1.2.2 for more detail). Partial derivatives with respect to four-dimensional coordinates x^α are denoted with ∂_a so that for any differentiable function $F(t,\boldsymbol{x})$ one has $\partial_a F \equiv \partial F/\partial x^\alpha$. Partial derivatives of $F(t,\boldsymbol{x})$ with respect to spatial coordinates x^i are denoted as $\partial_i F \equiv \partial F/\partial x^i \equiv \boldsymbol{\nabla} F$. Partial derivative of function $F(t,\boldsymbol{x})$ with respect to time is denoted $\partial_t F \equiv \partial F/\partial t$. Total derivative of $F(t,\boldsymbol{x})$ with respect to time is denoted with an overdot appearing above function, that is

$$\dot{F} \equiv \frac{dF}{dt} = \frac{\partial F}{\partial t} + \frac{d\boldsymbol{x}}{dt}\cdot\boldsymbol{\nabla} F\ . \tag{7.336}$$

Notice that, in general, $\dot{F} \neq \partial F/\partial t$ except when F is a function of time only.

One will repeat derivation of the solution of light-ray propagation equation for the case of light propagating in the field of a moving massive planet with a quadrupole component of its gravitational field taken into account. Mathematical technique will be similar to that used in section 7.3 but significantly adapted to the particular situation of the astrometric experiment under discussion. For this reason, some details of the technique are repeated here to make this section self-sufficient. One starts from the discussion of the unperturbed trajectory of light ray.

7.9.2
The Unperturbed Light-ray Trajectory

Let us introduce a global coordinate system $x^\alpha = (x^0, x^i) = (ct, x^i)$ covering the entire spacetime and coinciding with an inertial frame at infinity where gravitational field is absent. One will assume that observer does not move with respect to the coordinate system x^α. Section 7.9.4 will prove that the results are gauge-independent and Lorentz-invariant and, in fact, are valid in any other frame with observer moving with arbitrary velocity. It is further assumed that the origin of the coordinate system x^α is close to the center of mass of the planet, which is the only source of the

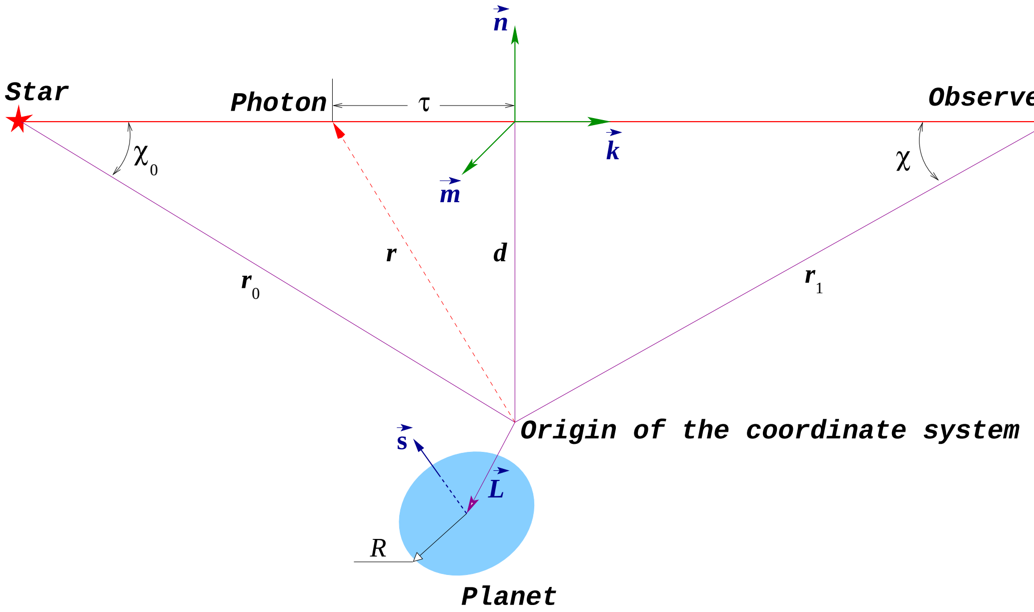

Figure 7.10 Light propagates from a star to observer in the direction of the unit vector $\boldsymbol{k}$. The planet is displaced from the origin of the coordinate system by a vector $\boldsymbol{L} = \boldsymbol{x}_P$ that is a parameter of the data analysis algorithm. The planet's rotation axis is specified by the unit vector $\boldsymbol{s}$. The impact parameter of the light ray is d. Two unit vectors, $\boldsymbol{n}$ and $\boldsymbol{m}$, are orthogonal to the vector $\boldsymbol{k}$ and form a "plane of the sky" that is perpendicular to the line of sight of the observer. Vector $\boldsymbol{s}$ has arbitrary direction in space not associated with the plane of the sky. Letter χ denotes the angle between the star and the planet.

gravitational field. This assumption can be relaxed in the linearized approximation of general relativity where the gravitational field of masses is additive [Misner *et al.*, 1973]. One prefers to simplify the consideration and does not include other masses because the main concern is the propagation of light in the field of a particular planet. The light deflection by other bodies of the solar system is calculated on the basis of the same equations and can be added later if necessary.

Let us consider a bundle of light-rays emitted by a source of light (star, quasar) simultaneously and propagating as a narrow beam along parallel lines toward the solar system. In the absence of gravitational field each light particle (photon) from the bundle propagates in the coordinate system x^α along a straight line

$$x^i = x_0^i + ck^i(t - t_0) \,, \tag{7.337}$$

where t_0 and $x_0^i = x^i(t_0)$ are the time and space coordinates of the photon at the time of emission, and k^i is the unit vector along the unperturbed photon's trajectory as shown in Figure 7.10. One assumes that the photon hits detector (is observed) at time t_1 when its coordinate $x_1^i = x^i(t_1)$. Let us denote the time of the closest approach of the photon to the origin of the coordinate system as [18)]

$$t_* = t_0 - c^{-1}\boldsymbol{k}\cdot\boldsymbol{x}_0 \ . \tag{7.338}$$

It is mathematically convenient to introduce parameter τ along the unperturbed light ray

$$\tau = t - t_* \ , \tag{7.339}$$

and a constant (impact-parameter) vector

$$\xi^i = P^{ij}x^j = P^{ij}x_0^j \ , \tag{7.340}$$

that points out from the origin of the coordinate system to the point of the closest approach of the unperturbed light ray (see Figure 7.10). Here $P^{ij} = \delta^{ij} - k^i k^j$ is the operator of projection onto the plane of the sky which is orthogonal to the vector k^i. By definition, $P^{ij}P^{jk} = P^{ik}$. Notice that the impact parameter vector ξ^i has nothing to do with the impact parameter of the light ray with respect to the planet, which is not located, in general, at the origin of the coordinate system. The role of the impact parameter ξ^i is supplementary as it will not enter any equation describing the observable effects. This is natural as one has chosen the origin of the coordinate system arbitrary and the result of physical observations can not depend on this choice.

Parametrization of the unperturbed light-ray trajectory given by equations (7.338)-(7.340) converts equation (7.337) to

$$x^i = ck^i\tau + \xi^i \ , \tag{7.341}$$

where τ and ξ^i are independent of each other, and can be considered as coordinates in the 2+1 manifold of the light-ray bundle because the projection (7.340) makes ξ^i have only two independent spatial components lying in the plane of the sky. At each instant of time $t = \tau + t^*$, distance $r = r(\tau)$ of the photon from the origin of the coordinate system is

$$r = \sqrt{c^2\tau^2 + d^2} \ , \tag{7.342}$$

where $d = |\boldsymbol{\xi}|$ is the absolute value of the impact parameter of the photon, that is constant for each light ray from the parallel beam of the light-ray bundle moving in the direction $\boldsymbol{k}$. Notice that definition (7.339) implies that $\tau = 0$ when $t = t_*$, $\tau < 0$ when $t < t_*$, and positive otherwise (see Figure 7.10).

7.9.3 The Gravitational Field

7.9.3.1 *The field equations*

The solar system is considered as isolated and spacetime as asymptotically flat which means there are no other masses outside of the solar system. The gravitational field

18) Notice that t_* changes its value from one inertial frame to another in accordance with the law of the Lorentz transformation. Equation (7.338) gives t_* in a static, non-moving frame.

of the solar system is produced by the Sun and the planets which curve spacetime and deflect light by their gravitational fields. In what follows, one takes into account the gravitational field of two bodies only - the Sun and a planet moving around the barycenter of the solar system. Furthermore, one considers the solar gravitational field as spherically-symmetric in its own, proper reference frame because in this frame the dipole and quadrupole moments of the Sun are negligible. Gravitational deflection of light by the Sun is well-known [Brumberg, 1991] and the discussion will be focused of the light bending by a moving axisymmetric planet.

The planet moves around the barycenter of the solar system as a light ray propagates from a star toward observer. Position of the planet with respect to the origin of the coordinates x^α at time t is defined by vector $\boldsymbol{x}_P = \boldsymbol{x}_P(t)$, its velocity is denoted $\boldsymbol{v}_P = d\boldsymbol{x}_P/dt$, and acceleration $\boldsymbol{a}_P = d\boldsymbol{v}_P/dt$. One stipulates that the planet's distribution of mass is axisymmetric around the unit vector $\boldsymbol{s} = \boldsymbol{s}(t)$ that defines the rotational axis of the planet at time t. This vector can change its orientation in space due to precession. Our calculation method is general enough, and it does not need to assume that the parameters characterizing translational and rotational motion of the planet are constant or equal to zero. In other words, one does not assume that $\boldsymbol{x}_P = 0$, nor that $\boldsymbol{v}_P = \boldsymbol{a}_P = 0$ and $\boldsymbol{s}$ are constant vectors. One calculates gravitational deflection of light by the axisymmetric gravitational field of the planet but ignores the relativistic effects in light propagation caused by its gravitomagnetic field due to the intrinsic rotation of planet [Kopeikin and Mashhoon, 2002], since these effects are negligibly small and can not be detected with the microarcsecond astrometric resolution [Brumberg *et al.*, 1990; Turyshev, 2009b].

The gravitational field of the solar system is described by the metric tensor

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} \ , \tag{7.343}$$

where $\eta_{\alpha\beta}$ is the constant Minkowski metric, and $h_{\alpha\beta} = h_{\alpha\beta}(t, \boldsymbol{x})$ is its perturbation which is associated in general relativity with gravitational potentials. Let us impose the harmonic gauge condition 7.2 on the potentials which is convenient and mathematically powerful choice for solving Einstein's gravity field equations [Fock, 1957, 1964]. Outside the planet, in vacuum, and in the harmonic gauge (7.2), the linearized Einstein equations for the field $h_{\alpha\beta}$ are homogeneous wave equations [Blanchet and Damour, 1986, 1989; Thorne, 1980]

$$\left(-\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \boldsymbol{\nabla}^2\right) h_{\alpha\beta} = 0 \ . \tag{7.344}$$

A question can arise about the universality of the harmonic gauge and physical interpretation of the solution of the wave equation (7.344). Indeed, Einstein's theory of general relativity is formulated in covariant tensor form while the wave equation (7.344) is valid in a particular harmonic gauge only. First of all, one notices that the harmonic gauge does not pick up a single coordinate system but admits a whole class of both global and local harmonic coordinates related to each other by the coordinate transformations which do not violate the harmonic gauge condition (7.2). The class of the global harmonic coordinates consists of the reference frames which are moving with respect to each other with constant velocities. These asymptotically-inertial reference frames are connected to each other through the Lorentz transformation as was shown by Fock [Fock, 1957] for a generic case including non-linearity of Einstein's equations. Local harmonic coordinates were introduced to gravitational physics by Thorne and Hartle [Thorne and Hartle, 1985]. The law of transformation between the local harmonic coordinates extends the Lorentz transformation to that making a class of harmonic polynomials that are solutions of the homogeneous wave equation [Brumberg and Kopejkin, 1989a,b; Damour *et al.*, 1991; Klioner and Voinov, 1993; Kopeikin and Vlasov, 2004; Kopejkin, 1988a]. They have a great practical value for modern fundamental astronomy [Kopeikin, 2007; Soffel *et al.*, 2003]. Second important observation is that the class of the harmonic coordinates is actually much more general than physicists used to think. In particular, papers [Kopeikin and Korobkov, 2005; Kopeikin *et al.*, 2006, 1999] showed that the standard Arnowitt-Deser-Misner (ADM) gauge [Arnowitt *et al.*, 1962, 2008] widely used for dynamic formulation of general relativity, may be viewed as a sub-class of the harmonic gauge in vacuum (see also Section 3.10.4 of the present book). Final remark is that data processing algorithms always operates with observable quantities - deflection angle of a light ray, time delay, frequency shift, etc. These quantities are invariant with respect to gauge transformations and, hence, all results obtained in this and other sections of the present book are valid in arbitrary gauge. The harmonic gauge is simply a convenient mathematical tool facilitating calculations and physical interpretation of the observed relativistic effects.

7.9.3.2 ***The planet's gravitational multipoles***

A general solution of equation (7.344) is given in the form of a multipolar expansion (3.544), (3.545a) and (3.560) depending on spatial coordinates and time-dependent parameters which are linear combinations of the intrinsic multipole moments of the planet and displacement vector $\boldsymbol{x}_P$ of its center of mass from the origin of the coordinate system. This section omits the relativistic effects of the planet's spin and focuses primarily on the effects caused by its mass monopole, dipole, and quadrupole fields in the multipolar expansion of the gravitational field of the planet. Relativistic effects caused by the spin and higher-order mass multipoles are discussed exhaustively in [Kopeikin and Korobkov, 2005; Kopeikin *et al.*, 2006; Kopeikin and Mashhoon, 2002; Kopeikin, 1997c].

Because equation (7.344) is linear, one can consider gravitational field of the solar system as a linear superposition of the individual fields. For each massive body the solution of equation (7.344) in the quadrupolar approximation is

$$h_{00}(s,\boldsymbol{x}) = \frac{2GM}{c^2 r} - \frac{\partial}{\partial x^i}\left[\frac{2GI^i(s)}{c^2 r}\right] + \frac{\partial^2}{\partial x^i \partial x^j}\left[\frac{GI^{<ij>}(s)}{c^2 r}\right], \tag{7.345}$$

$$h_{0i}(s,\boldsymbol{x}) = -\frac{4G\dot{I}^i(s)}{c^3 r} + \frac{\partial}{\partial x^j}\left[\frac{2G\dot{I}^{<ij>}(s)}{c^3 r}\right], \tag{7.346}$$

$$h_{ij}(s,\boldsymbol{x}) = \delta_{ij}h_{00} + \frac{2G\ddot{I}^{<ij>}(s)}{c^4 r}, \tag{7.347}$$

where $r = |\boldsymbol{x}|$ is the distance from the origin of the coordinate system to the field point $x^\alpha = (ct, x^i)$, and the angular brackets around two indices denote the STF tensor which structure has been explained in equation (1.49) and in appendix A. Notice that the second time derivative of the dipole moment I^i is absent due to the law of conservation of the linear momentum of the body [Blanchet and Damour, 1986, 1989; Thorne, 1980].

The gravitational field described by equations (7.345)-(7.347) should be interpreted as created by a massive planet placed at some distance $\boldsymbol{x}_P$ from the origin of the coordinate system and characterized by several parameters, of which M is a constant mass of the planet, $I^i(s)$ is the dipole, and $I^{<ij>}(s)$ is the quadrupole moment - both taken at the retarded instant of time

$$s = t - \frac{r}{c} \,. \tag{7.348}$$

The retardation in the solution (7.345)-(7.347) of the gravity wave equation (7.344) is a direct consequence of the retarded (causal) nature of the gravitational field in general relativity and its finite speed of propagation, which is equal to c. The null cone corresponding to the causal domain of influence of the gravitational field of planet is shown in Figure 7.11. The planet moves along its own worldline so that observer is expected to see any effect of the planet's gravitational field not instantaneously but with the retardation that is due to the finite speed of propagation of gravity on the future part of the null cone. Figure 7.11 gives a general-relativistic picture of the process of propagation of gravity from the moving planet to observer but it is also valid in a number of alternative theories of gravity [Kopeikin and Fomalont, 2006; Kopeikin and Ni, 2008].

The dipole and quadrupole moments are fully taken into account because they produce a relativistic deflection of light that can be significant for several planets [Brumberg *et al.*, 1990; Crosta and Mignard, 2006; Malkin *et al.*, 2009; Turyshev, 2009b]. In general relativity, all relativistic effects due to the dipole moment of the gravitational field are coordinate-dependent and, hence, can be eliminated if the origin of the coordinate system is placed exactly at the center of mass of the real planet. However, one retains the dipole moment I^i in the multipolar decomposition of the planetary gravitational field (7.345)-(7.347) because in practice the center of mass of the planet is a fitting parameter which is to be determined from observations. Furthermore, even if one were able to make $I^i = 0$ at a particular instant of time, this equality would be violated in later times because the uncertainties of astrometric measurements will lead to the residual velocity of the planet with respect to the coordinate frame and planet's center of mass, x^i_P, can not be kept fixed at the origin of the frame. It is true that the effect of this residual velocity on the value of the dipole may be rather small but it exists from a principal point of view. Therefore, it is important to properly parameterize and evaluate the influence of the dipole moment of the gravitational field on light deflection. More physical reason, for having $I^i \neq 0$ is that planet moves with acceleration around the barycenter of the solar system. If one try to make $I^i = 0$ for any instant of time, it will violate the basic assumption that the coordinate chart used for calculations is asymptotically flat as the accelerated frame does not admit this property [Misner *et al.*, 1973].

If the center of mass of the planet is shifted from the origin of the coordinate

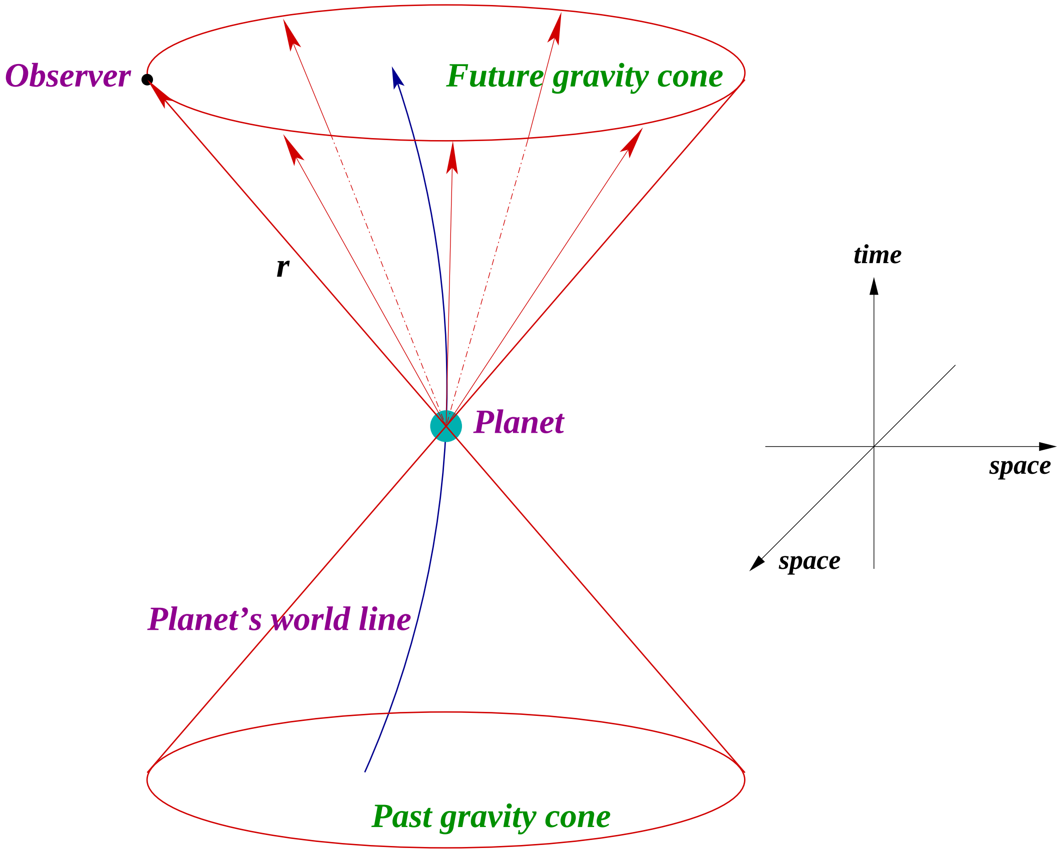

Figure 7.11 Gravitational field of a planet is a retarded solution of the gravity wave equation (7.344). In general relativity, gravitational field propagates on the hypersurface of a null cone from past to future. Directions of the propagation of the gravitational field are null characteristics of the gravity field of the planet and they are shown by arrows. The picture assumes for simplicity that the coordinate-dependent effects associated with the dipole moment I^i are excluded ($I^i = 0$). Observer measures gravitational field at time t, when the planet is located at the retarded position on its orbit at the retarded time $s = t - r/c$, that is equation of a null characteristic of the gravitational field.

system by a spatial vector $L^i = x_P^i$, the dipole and quadrupole moments of the gravitational field entering the metric tensor perturbations (7.345)-(7.347) are defined in the linearized approximation of general relativity by the following equations

$$I^i = Mx_P^i \,, \tag{7.349}$$
$$I^{<ij>} = J^{<ij>} + Mx_P^{<i}x_P^{j>} \,, \tag{7.350}$$

where $J^{<ij>}$ is the intrinsic quadrupole moment of the planet in its own proper reference frame whose origin coincides with the planet's center of mass, and one has used the parallel-axis theorem [Arnold, 1995] to split I^{ij} to J^{ij} and the translational quadrupole term $Mx_P^{<i}x_P^{j>}$. Equations (7.349), (7.350) approximate the post-Newtonian definitions of the multipole moments used in the relativistic celestial mechanics of N-body system as described in section 4.4 of this book. However, the post-Newtonian corrections to equations (7.349), (7.350) yield terms which are much less than 1 μas in light deflection, and hence, can be ignored.

Both the dipole, Mx_P^i, and the translational quadrupole, $Mx_P^{<i}x_P^{j>}$, will affect the post-fit results of the light-ray deflection measurement, and thus, they are directly observable in the light-ray deflection pattern. However, if general relativity is valid, these deflections are spurious, coordinate-dependent effects and the goal of the data analysis system is to suppress their impact on the values of the fitting parameters within the accuracy of astrometric observations. To this end, the dipolar and quadrupolar deflections associated with the translation $L^i = x_P^i$ must be incorporated to the data analysis in order to assume a full control on their influence on the values of the fitting physical parameters like the intrinsic oblateness J_2, etc. If this is not done, the coordinate-dependent effects can not be eliminated and the light deflection due to the dipole, Mx_P^i, and the translational quadrupole, $Mx_P^{<i}x_P^{j>}$, may exceed the physical deflection of light caused by the intrinsic quadrupole moment, J_2, of the planet, thus, making its measurement in the gravitational light-ray deflection experiments impossible.

In what follows, the planet is assumed to be axisymmetric around its rotational axis defined at each instant of time by a unit vector s^i. The planet has equal equatorial moments of inertia $A = B$, and the axial moment of inertia $C \neq A$. The dynamic oblateness of the planet is denoted as $J_2 = (C - A)/A$ [Brumberg, 1991; Murray, 1983]. This definition yields the intrinsic quadrupole moment represented as an STF tensor of the second rank [Murray, 1983]

$$J^{<ij>} = MJ_2R^2\left(s^is^j - \frac{1}{3}\delta^{ij}\right) \,, \tag{7.351}$$

where R is the equatorial radius of the planet. It is immediately seen from equation (7.351) that the intrinsic quadrupole moment is symmetric and trace-free, that is $J^{<ii>} = 0$, in accordance with its definition. One also notices that the first time derivative of the dipole and quadrupole moments

$$\dot{I}^i = Mv_P^i \,, \tag{7.352}$$
$$\dot{I}^{<ij>} = 2Mx_P^{<i}v_P^{j>} \,, \tag{7.353}$$

should be used in calculations of h_{0i} in equation (7.346). The first time derivative of $J^{<ij>}$ in h_{0i}, and the second time derivative of the overall quadrupole moment $I^{<ij>}$ in h_{ij} will be neglected as they lead to higher-order relativistic effects in light deflection. For example, as follows from the subsequent calculations, the light-ray deflections caused by the second derivative of the quadrupole moment $I^{<ij>}$ have magnitude of the order of $\alpha_M(v_P/c)^2$, where α_M is the gravitational bending due to the mass monopole of the planet. In case of Jupiter the maximal value of $\alpha_M \sim$ 16300 μas [Brumberg *et al.*, 1990; Crosta and Mignard, 2006; Turyshev, 2009b], and $v_P/c \sim 4.5\times10^{-5}$ [Kopeikin, 2001], that makes the light deflection angle caused by the second time derivatives $\alpha_M(v_P/c)^2 \ll 1$ μas.

7.9.4
The Light-ray Gravitational Perturbations

7.9.4.1 *The light-ray propagation equation*

The equation of motion of light in gravitational field is given by the light-ray geodesic 7.20 which is

$$\frac{dK^\alpha}{d\lambda} + \Gamma^\alpha_{\mu\nu}K^\mu K^\nu = 0\,, \tag{7.354}$$

where $K^\alpha = dx^\alpha/d\lambda$ is the wave vector of the electromagnetic wave corresponding to the light ray, λ is the affine parameter along the light ray, and

$$\Gamma^\alpha_{\mu\nu} = \frac{1}{2}g^{\alpha\beta}\left(\partial_\nu g_{\beta\mu} + \partial_\mu g_{\beta\nu} - \partial_\beta g_{\mu\nu}\right)\,, \tag{7.355}$$

is the Christoffel symbols (7.24)–(7.29).

When one substitutes the expansion of the metric tensor (7.343) to equation (7.354), and transforms the affine parameter λ in this equation to the coordinate time t, the light geodesic equation (7.354) is reduced to the following three-dimensional form [Brumberg, 1991]

$$\frac{d^2x^i}{dt^2} = F^i(t,\boldsymbol{x})\,, \tag{7.356}$$

where the gravitational perturbation is given explicitly in equation 7.30 and depends on the perturbation of the metric tensor $h_{\alpha\beta}$ and the coordinate velocity of photon $\dot{x}^i$. One notices that making use of time t in place of the parameter λ does not change direction of propagation of the light ray, and is merely a technical tool that allows us to express coordinate of the photon as a function of the same time argument which governs evolution of the multipole moments of the gravitational field.

Equation (7.356) has an unperturbed ($F^i = 0$) solution described in section 7.9.2. In a weak-gravitational field approximation the light-ray geodesic equation (7.356) has a unique solution given by

$$x^i = ck^i\tau + \xi^i + \Xi^i(\tau,\boldsymbol{\xi}) - \Xi^i(\tau_0,\boldsymbol{\xi})\,, \tag{7.357}$$

where $\Xi^i(\tau, \boldsymbol{\xi})$ is a small perturbation of the straight light ray trajectory. In terms of the parameter τ and the projection coordinates ξ^i, equation (7.356) is significantly simplified, so that the perturbation Ξ^i obeys the ordinary differential equation [Kopeikin *et al.*, 2006, 1999]

$$\frac{d^2\Xi^i}{d\tau^2} = \frac{c^2}{2} k^\alpha k^\beta \frac{\partial h_{\alpha\beta}}{\partial \xi^i} - c\frac{d}{d\tau}\left(k^\alpha h_{i\alpha} + \frac{1}{2} k^i h_{00} - \frac{1}{2} k^i k^j k^p h_{jp}\right), \tag{7.358}$$

where $k^\alpha = (1, k^i)$ is a null vector along the unperturbed trajectory of the light ray. Equation (7.358) describes two relativistic effects - gravitational bending and time delay of light which are closely related to each other [Kopeikin and Korobkov, 2005]. This section discusses only the deflection of light as one of the most prominent method of testing general theory of relativity in the solar system..

7.9.4.2 *The null cone integration technique*

One has already described in detail the integration technique of the null geodesics in section 7.3. Nevertheless, since the metric tensor (7.345)-(7.347) has the structure being different from the retarded *Liénard-Wiechert* potentials (7.12), it is instructive to re-introduce this technique to see more clearly how the null cone structure of the gravitational field is involved to the process of propagation of light.

Without imposing any restriction on the ratio of the impact parameter d of the light-ray trajectory to the distance r_1 from the origin of the coordinate system to observer, the total angle of the gravitational deflection of light in the plane of the sky is given by a vector

$$\alpha^i = -\frac{1}{c} P^{ij} \frac{d\Xi^j}{d\tau} + \Delta\alpha^i, \tag{7.359}$$

that should be compared with equation (7.112). Function $\Delta\alpha^i$ in equation (7.359) includes relativistic corrections due to both the finite distance r_0 of the source of light (star, quasar) from the planet, the finite distance r_1 between the observer and the planet, and the difference of the local inertial frame of observer [Klioner, 2004; Ni and Zimmermann, 1978b] from the coordinate system $x^\alpha = (t, x^i)$ introduced for the calculation of light-ray propagation. It is further stipulated that $r_0 \to \infty$, and the observer is at rest at a sufficiently large distance from the light-ray deflecting body which are realistic assumptions [19]. Under these circumstances $\Delta\alpha^i$ is negligible, and can be omitted. Indeed, the largest contribution to $\Delta\alpha^i$ associated with the finite distance r_0 of star from the planet, is smaller than the first term in the right side of equation (7.359) by a factor of r_1/r_0 [Kopeikin and Schäfer, 1999; Kopeikin *et al.*, 1999] that, in case of Jupiter, is about 2×10^{-5} even for the closest star α Centauri, thus, making $\Delta\alpha^i \ll 1$ μas. Notice that the finite-distance effects in the light-ray deflection angle may be important in the solar gravitational bending of light for several nearby stars as discussed in a paragraph following equation (7.269). Contribution

19) Real observer (SIM, Gaia, VLBI station, etc.) moves with respect to the coordinate system x^α with velocity v^i. Relativistic effects of this motion on light are obtained later after calculation of the gravitational deflection of light by making use of the Lorentz transformation.

to $\Delta\alpha^i$ due to the difference between the local inertial and global coordinate frames is smaller than the first term in the right side of equation (7.359) by a factor of $(d/r_1)(v_P/c)^2$ [Kopeikin and Schäfer, 1999; Kopeikin *et al.*, 1999] that also makes $|\Delta\alpha^i| \ll 1\,\mu$as, and one can ignore it. In cases when $\Delta\alpha^i$ may be important, for example, in studying refraction of light by gravitational waves [Kopeikin and Korobkov, 2005; Kopeikin *et al.*, 1999] or in cosmological gravitational lensing where $r_0 \neq 0$, and r_1 is comparable with r_0, the term $\Delta\alpha^i$ is solved in papers [Kopeikin *et al.*, 2006; Kopeikin and Mashhoon, 2002; Kopeikin and Schäfer, 1999].

Integrating equation (7.358) along the unperturbed light ray trajectory and substituting the result to equation (7.359) yields

$$\alpha^i = -\frac{c}{2}\frac{\partial}{\partial\xi^i}\int_{\tau_0}^{\tau_1} k^\alpha k^\beta h_{\alpha\beta}(\tau,\boldsymbol{\xi})d\tau + k^\alpha P^{ij}\Big[h_{j\alpha}(s_1,\boldsymbol{x}_1) - h_{j\alpha}(s_0,\boldsymbol{x}_0)\Big]\,, \tag{7.360}$$

where the retarded times

$$s_0 = t_0 - \frac{r_0}{c}\,, \tag{7.361}$$

$$s_1 = t_1 - \frac{r_1}{c}\,, \tag{7.362}$$

the integrand is taken on the unperturbed light-ray trajectory and can be represented as

$$\begin{aligned}\frac{c^2}{2}k^\alpha k^\beta h_{\alpha\beta}(\tau,\boldsymbol{\xi}) = 2G\Bigg\{&\frac{M}{r} - \frac{\partial}{\partial\xi^j}\left[\frac{I^j(s)}{r}\right] - \frac{d}{d\tau}\left[\frac{k^j I^j(s)}{cr}\right] - \frac{k^j \dot{I}^j(s)}{cr}\\ &+ \frac{\partial^2}{\partial\xi^j\partial\xi^p}\left[\frac{I^{<jp>}(s)}{2r}\right] + \frac{d^2}{d\tau^2}\left[\frac{k^j k^p I^{<jp>}(s)}{c^2 r}\right] + \frac{\partial}{\partial\xi^j}\frac{d}{d\tau}\left[\frac{k^p I^{<jp>}(s)}{cr}\right]\Bigg\}\,,\end{aligned} \tag{7.363}$$

the limits of integration τ_1 and τ_0 are the values of the parameter τ taken at the time of observation, t_1, and emission, t_0, respectively (notice that $\tau_0 < 0$ and $\tau_1 > 0$), and $s = t_* + \tau - r/c$ with $r = r(\tau,\boldsymbol{\xi})$ given by equation (7.342). Integration in the right side of equation (7.360) can be performed easily if one adopts that in general relativity the speed of propagation of gravity and the speed of light are numerically the same, making the light cone hypersurface coinciding with that of the gravity null cone, so that in practical experiments gravity interacts with light in a way shown and explained in Figure 7.12. In some alternative theories of gravity the hypersurfaces of the light and gravity null cones do not coincide as the speed of propagation of gravity and light are different [Bailey and Kostelecký, 2006; Kopeikin, 2004; Mattingly, 2005; Will, 1993, 2006]. One has discussed the gravitational deflection of light caused by moving bodies, in such theories in papers [Kopeikin, 2004; Kopeikin and Fomalont, 2006; Kopeikin and Ni, 2008] which can be used as references.

One has found [Kopeikin and Korobkov, 2005; Kopeikin *et al.*, 2006, 1999] that for any smooth function of the retarded time $F(s)/r$, where $r = \sqrt{c^2\tau^2 + d^2}$ and $s = t - r/c = t_* + \tau - \sqrt{\tau^2 + (d/c)^2}$, the integration can be performed explicitly, and

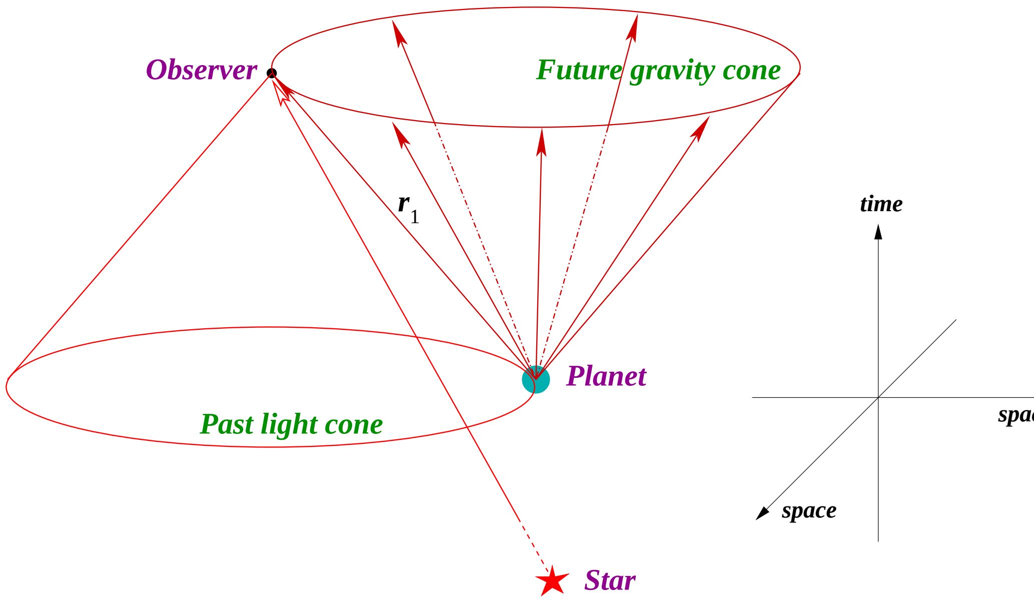

Figure 7.12 Gravitational field of a planet (the case of dipole $I^i = 0$ is shown) affects only the particles lying on the hypersurface of the future gravity null cone. A photon emitted by a star at time t_0 arrives to observer at time t_1 along a null direction of the past light cone with a vertex at the observer. Therefore, the future gravity null cone of the planet and the past light cone of the observer must coincide along the null direction that is a null characteristic of the retarded solution of the gravity-field wave equation (7.344). The photon detected at time t_1, is deflected by planet's gravity force from the planet's retarded position taken at time $s_1 = t_1 - r_1/c$. This effect of the retardation of gravity can be observed by measuring the amount of gravitational deflection of light by a moving planet, and used to measure the numerical value of the fundamental speed in Einstein's equations (the speed of gravity) as discussed at the end of section 7.9.4 and papers [Fomalont and Kopeikin, 2003; Kopeikin and Fomalont, 2002; Kopeikin, 2004; Kopeikin and Fomalont, 2006]. Note that the retardation of gravity would not be measurable had the planet been at rest with respect to observer.

it yields

$$\int_{\tau_0}^{\tau_1} \frac{\partial}{\partial \xi^i}\left[\frac{F(s)}{r}\right] d\tau = \frac{\partial}{\partial \xi^i}\int_{\tau_0}^{\tau_1} \frac{F(s)}{r} d\tau \tag{7.364}$$

$$= -\Big[(1+\cos\chi)F(s_1) - (1-\cos\chi_0)F(s_0)\Big]\frac{\xi^i}{cd^2} ,$$

$$\int_{\tau_0}^{\tau_1} \frac{d}{d\tau}\left[\frac{F(s)}{r}\right] d\tau = \frac{F(s_1)}{r_1} - \frac{F(s_0)}{r_0} , \tag{7.365}$$

where $\cos\chi_0 = |c\tau_0|/r_0$, $\cos\chi = |c\tau_1|/r_1$ are normalized to distances $r_0 = |\boldsymbol{x}(t_0)|$, $r_1 = |\boldsymbol{x}(t_1)|$ of the photon from the origin of the coordinate system taken at the times of emission and observation of light respectively. Signs in the right side of equation (7.364) are valid in case shown in Figure 7.10 when the light-ray deflecting body is located between the source of light and observer. If the observer is located between the source of light and the light-ray deflecting body, equation (7.364) is still applicable after replacing $\cos\chi \to -\cos\chi$ [Kopeikin *et al.*, 1999]. Finally, if $d = r_1$, that is the light ray arrives from the direction perpendicular to the light-ray deflecting body, one must take $\cos\chi = 0$ in equation (7.364). All these situations are discussed in [Kopeikin *et al.*, 1999] in more detail to which the reader is referred for a comprehensive theoretical review. Here one will discuss only a configuration shown in Figure 7.10 because only this case is practically important for observation of the gravitational deflection of starlight by the solar system planets.

It should be noted that the retarded time equation (7.362) describes the null direction connecting the planet and the observer and lying on the future gravity null cone with the planet at its vertex as shown in Figure 7.12. Light moves along a different null direction connecting the star and the observer and lying on the past light cone of the observer. Therefore, equation (7.362) describes the retardation effect in propagation of gravity force from the moving planet to a photon as the photon propagates toward the observer and is subsequently detected at time t_1. This retarded component in the interaction of gravity with light was measured within 20% in the VLBI experiment on September 8, 2002, and led to a direct observational confirmation of the general relativistic postulate that the speed of gravity and light are the same within the observational error [Fomalont and Kopeikin, 2003; Kopeikin, 2001; Kopeikin and Fomalont, 2006].

7.9.4.3 ***The speed of gravity, causality, and the principle of equivalence***

> Render unto Caesar the things which are Caesar's, and unto God the things that are God's
> *Matthew 22:21; Mark 12:17; Luke 20:25*

One should say here a few words about the concept of the speed of gravity because there is a great deal of confusion on this subject in literature. Most notably, Will [2006] argues that the speed of gravity is associated *exclusively* with the second time derivatives of the metric that appear only in the Riemann tensor and in the field equations. This point of view stems from the assumption adopted in Nordtvedt-Will's PPN formalism [Will, 1993], according to which the first-order partial time derivatives of the metric tensor entering the Christoffel symbols are "naturally" normalized to the fundamental speed of the Maxwell equations (that is the speed of light). Because the Christoffel symbols enter operationally the principle of equivalence, the book [Will, 1993, section 10.1] states that the principle of equivalence demands nothing about the speed of gravity. Hence, Will [2006] considers the speed of propagation of gravity from a moving planet to photon as being totally irrelevant in the astrometric light-ray deflection experiments and insists that the relativistic corrections to the gravitational light bending caused by a moving planet can be predicted correctly even if the speed of gravity were infinite.

This is an example of incomprehension of different facets of the fundamental speed c in conceptually different theories of fundamental physics [Ellis and Uzan, 2005] – general relativity and electrodynamics. Both theories operate with the same fundamental constant c, but c that normalizes time derivatives in general relativity is the limiting speed of propagation of gravity while c normalizing time derivatives in equations of electrodynamics is the limiting speed of light [Kopeikin, 2004]. One must also distinguish c that appears in the coupling constant of matter with gravity in the right side of Einstein's equations. Measuring relativistic effects produced by the time derivatives of the metric tensor in general relativistic equations is equivalent to measuring the ultimate speed of gravity. It does not matter whether the measured effect is produced by the first, or second, or third-order time derivative – they are all normalized to one and the same fundamental speed c. Conceptually, such measurement of c in time-dependent gravitational fields of general relativity is fundamentally different from measuring the fundamental speed c in electrodynamics. The question is whether one can work out a theoretical setting in which the effects associated with the speed of gravity (that is the c in time derivatives of general-relativistic equations) can be clearly distinguished from the c associated with the propagation of light. Contrary to what Will [1993, 2006] states, the answer to this question is positive, and the appropriate speed-of-gravity parametrization is based on the Liénard-Wiechert solutions of the linearized Einstein equations [Kopeikin and Fomalont, 2006, 2007]. The speed of gravity appears everywhere in equations in the form of the retarded time $s = t - r/c$ of the *Liénard-Wiechert* potentials. hence, the retarded time s can be considered as a parameter which helps to track down the gravity-propagation effects from the effects associated with the speed of light. This amazing property of the gravitational *Liénard-Wiechert* potentials was noticed by Kopeikin [2001] and led to the first experimental measurement of the numerical value of the fundamental speed c in general relativity [Fomalont and Kopeikin, 2003].

Light is just a test particle propagating in gravity field. Measuring how light is deflected by gravity, how it changes light's frequency, polarization, etc., allows us to study various properties of gravitational field including the speed of its propagation. Gravitational experiments with test particles are not measuring the properties of the particles. These properties must be calibrated and well-known before the experiment, so the test particle parameters are fixed and excluded from the data processing algorithm. The speed of light is a fundamental speed associated with Maxwell's equation for electromagnetic field in vacuum. This is the ultimate speed for propagation of any matter field. Gravity is not matter and is governed by Einstein's theory of general relativity which is physically different from the Maxwell theory of electromagnetic field. Both theories share the same fundamental speed c but one can trace its origin in any kind of gravitational experiments which make use of light as a probe of the gravity field characteristics. It is crucial to realize that the ultimate speed of gravity appears already in the first-order time derivative of the metric tensor, $\partial g_{\alpha\beta}/\partial x^0 = (1/c)\partial g_{\alpha\beta}/\partial t$ on spacetime manifold. The speed of gravitational waves is calculated from the gravity field equations, which involve the second time derivatives of the metric tensor. Since general relativity is a theory of massless gravitational field, the speed of gravitational waves coincide with the ultimate speed of gravity entering the Christoffel symbols [Low, 1999]. The hypothetical difference of the speed of gravitational waves from c could appear if graviton had mass [Babak and Grishchuk, 2003]. At the same time, the ultimate speed of gravity in the Christoffel symbols could be different from c because of non-zero value of nonmetricity $\mathfrak{Q}^{\alpha}_{\mu\nu}$ of the manifold's affine connection [Heinicke *et al.*, 2005]. In this case the fundamental speed of gravity in general relativity could be different from the speed of light, at least, in a some preferred frame [Kopeikin, 2004]. In any case, any experiment confirming that the time derivatives of equations of general relativity are normalized to c, indicates that the speed of gravity is finite. This is why the measurement of the fundamental speed c in the affine connection is of a paramount importance [Kopeikin, 2004; Kopeikin and Fomalont, 2006].

Einstein had no any experimental evidence for the finite speed of gravity. However, he consistently believed in the universal character of the principle of relativity and in the universality of the fundamental speed c. Einstein's letter written to Carl Seelig [Seelig, 1956] in 1955, expressed his point of view on the theory of relativity as follows:

> "Its novelty was to formulate that the Lorentz transformations are of importance beyond the scope of the Maxwell equations and concerns the structure of the space and time. Another new point was the conclusion that the Lorentz invariance is the general demand for any physical theory."

Hence, he postulated that the speed of gravity in all equations of general relativity must be equal to the speed of light in vacuum. However, this postulate is not a dogma but a matter of experimental confirmation.

Going back to PPN formalism [Will, 1993] one notices that it conceptually disentangles the meaning of the fundamental speed c normalizing the first and second partial time derivatives of the metric tensor, and focuses on the parametrization of the second time derivatives which are supposed to deal with the speed of gravitational waves. Will [1993] stipulates that the fundamental speed c in the first partial time derivatives is automatically equal to the speed of light. Hence, Will [2006] claims that any gravitational experiment measuring the first time derivatives of the metric tensor makes no sense because there must be no violation of general relativity associated with the principle of equivalence in time-dependent gravitational fields. In this respect PPN formalism is at fault and unable to test a large class of non-metric theories of gravity predicting different violations of causality principle for gravitational field. It is well-known [Hawking and Ellis, 1975] that the causal behavior of gravitational field is associated not so much with the field equations but with equations of null geodesics in a given spacetime manifold. This behavior can be determined in a close neighborhood of any event from the solution of the equation of geodesics without imposing Einstein's field equations [Wald, 1984]. The null geodesics play a special role as they define: 1) the causal past of observer, that is the region bounded by the past light cone in Figure 7.12, and 2) the causal future of the gravitational field of a moving massive body, that is the region inside the future gravity cone in Figure 7.12. General-relativistic gravitational interaction of photon with the light-ray deflecting body implies that the causal past of the observer must coincide with the causal future of the gravitational field along a null direction [20]. This relationship between the past and future null cones of the two fields would be violated had the speed of gravity been not equal to the speed of light.

The equivalence principle tells us that in a local reference frame light moves along a straight line [Landau and Lifshitz, 1975; Misner *et al.*, 1973]. This implies that in the global reference frame the light-ray trajectory is bent because of two reasons: 1) the local reference frame falls in the gravitational field with acceleration, and 2) the space is curved [Misner *et al.*, 1973]. Gravitational light-ray deflection is not like the Compton scattering of photons in particle physics. It does not occur in a single local frame of reference but gradually accumulates to its integrated value as photon propagates through a continuous sequence of such local frames. Photon's classic propagator is given by the integral from the affine connection, that is the gravity force F^i shown in the right side of equation (7.356), with the integration performed along the null cone. This propagator, after it is projected onto the plane of the sky with operator $P^{ij} = \delta^{ij} - k^i k^j$, yields the integrated value of the deflection angle α^i. The magnitude and direction of the deflection angle are functions not only the wave vector k^i of the observed starlight but the position of the light-ray deflecting body $x^i_a(s)$ with respect to observer as well. If one is able to derive position of the massive body from the precise measurement of the angle α^i and to confirm that the body and the observer are connected by a characteristic of the gravity null cone, it gives a direct proof of the causal nature of the gravitational field and measures the speed of gravity [Kopeikin, 2001, 2004; Kopeikin and Makarov, 2007]. It is clear that this measurement of the speed of gravity is impossible if the light-ray deflecting body does not move. Indeed, in case of a static planet the gravitational field does not change as photon moves from star to observer, making the causal character of the gravitational field hidden because the planet is always at a fixed distance from the observer (see Figure 7.13a).

The experimental situation changes dramatically if the planet moves with respect to observer because it makes the distance between the planet and the observer dependent on time (see Figure 7.13b). In this case, photon traverses through the gravitational field that changes on the light-ray trajectory due to the planetary motion,

20) More generally, on a null hypersurface.

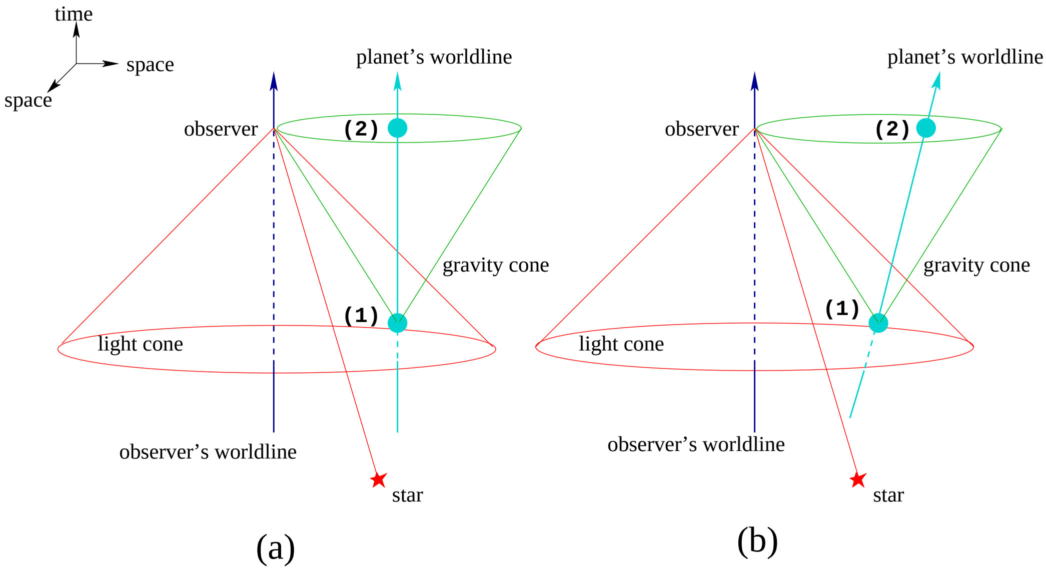

Figure 7.13 Light-ray deflection by a static (a) and moving (b) planet. In case (a) the distance between the planet and observer does not change as light propagates. Thus, measuring the deflection of light does not allow us to determine experimentally whether the gravity force of the planet acts on photon with retardation from position (1), or instantaneously from position (2). In case (b) the distance between the planet and observer varies as photon travels toward the observer. The retarded interaction of gravity with light becomes visible since measuring the angle of the gravitational deflection of light allows us to distinguish between positions (1) and (2) of the planet on its worldline making the causal structure of the gravity null cone measurable.

even if this motion is uniform. If the speed of gravity were different from the speed of light the light cone would be different from the gravity null cone, and the moving planet could not deflect light from the retarded position as predicted by general relativity [21]. For example, the instantaneous propagation of gravity would imply that one could determine current (as opposed to retarded) position of the planet on its orbit from observation of the gravitational deflection of light. This means that gravitational field would transmit information about the planet's orbital position to observer faster than light. This violates the principle of causality. Hence, one concludes that the correct description of the gravitational physics of the light-ray deflection experiment with a moving planet requires taking into account both the light and gravity null cones as demonstrated in Figures 7.11, 7.12, 7.13, and supported by calculations in this book. Additional arguments supporting this interpretation of the retardation-of-gravity experiments are discussed in paper [Kopeikin and Fomalont, 2006] both in the framework of general relativity and in a bi-metric theory of gravity [Carlip, 2004]. One concludes that the principle of equivalence does imply the principle of causality for gravitational field which can be, and has been, confirmed in the solar system experiments of the gravitational deflection of radio waves for quasar by moving planets [Fomalont and Kopeikin, 2008; Fomalont *et al.*, 2009b; Fomalont and Kopeikin, 2003].

7.9.5 Light-ray Deflection Patterns

7.9.5.1 *The deflection angle*

Let us assume for simplicity that the distance r_0 from the solar system to the source of light (star, quasar) is very large compared to the impact parameter d. It allows us to omit all terms, related to the time of emission, from the equations for calculation of the light-ray deflection angle because $1 - \cos\chi_0 \simeq d^2/(2r_0^2) \ll 1$. It will be assumed that only the stars that are sufficiently close to the planetary limb in the plane of the sky are observed during the time of the experiment. This makes the angle $\chi \ll 1$, and hence $\cos\chi \simeq 1 - d^2/2r_1$ and $\sin\chi = d/r_1 \ll 1$. All terms being proportional to $\sin\chi$ will be neglected, and the approximation $\cos\chi = 1$ will be used everywhere but in the calculation of the monopolar deflection, where more exact approximation of $\cos\chi$ is required.

Integrating equation (7.360) with the technique explained in equations (7.364), (7.365), and keeping only the leading terms, one deduces

$$\alpha^i = \alpha_\odot + \alpha^i_M + \alpha^i_D + \alpha^i_Q \, , \tag{7.366}$$

where α^i_M, α^i_D, α^i_Q are the angles of the gravitational deflection of light caused by the planetary mass monopole, dipole, and quadrupole moments respectively, and $\alpha_\odot$ stands for the Eddington's deflection angle of light caused by Sun's gravity [Ken-

21) This is because the Lorentz transformation must transform both the gravity and electromagnetic field, simultaneously.

nefick, 2009]

$$\alpha^i_\odot = \frac{2GM_\odot}{c^2}(1+\cos\chi_\odot)\frac{\xi^i_\odot}{d^2_\odot}\;, \tag{7.367}$$

where $M_\odot$ is mass of the Sun, $\xi^i_\odot$ is the impact parameter of the light ray with respect to the Sun, $d_\odot = |\boldsymbol{\xi}_\odot|$, $\cos\chi_\odot = \sqrt{1-(d_\odot/\rho)^2}$,and ρ is the heliocentric distance to the observer. The other deflection angles are defined with respect to the coordinate system associated with the planet (see Figure 7.10) by the following equations [Kopeikin and Makarov, 2007]

$$\alpha^i_M = \frac{2GM}{c^2}(1+\cos\chi)\frac{\xi^i}{d^2}\;, \tag{7.368}$$

$$\alpha^i_D = -\frac{2G}{c^2}\left\{(1+\cos\chi)k^j\dot{I}^j(s_1)\frac{\xi^i}{cd^2}+\frac{\partial}{\partial\xi^i}\left[(1+\cos\chi)\frac{\xi^j I^j(s_1)}{d^2}\right]\right\}\;, \tag{7.369}$$

$$\alpha^i_Q = \frac{G}{c^2}\frac{\partial^2}{\partial\xi^i\partial\xi^j}\left[(1+\cos\chi)\frac{\xi^p I^{<jp>}(s_1)}{d^2}\right]\;, \tag{7.370}$$

of which equation (7.368) is exact, and in equations (7.369), (7.370) the terms proportional to $\sin\chi$ and higher, have been omitted because of their smallness. The omitted residual terms are of the order of $\alpha_D d/r_1$ and $\alpha_Q d/r_1$, that is negligible, since $\max\{\alpha_D,\alpha_Q\}\simeq 2000$ μas (see Figs. 7.14, 7.15), and $d/r_1\simeq 10^{-4}$ for Jupiter, and about 10^{-5} for Saturn. Taking all the partial derivatives with accounting for strong inequality

$$\frac{\partial\cos\chi}{\partial\xi^i} = -\frac{\xi^i}{d^2}\cos\chi\sin^2\chi \simeq \frac{d}{r_1^2} \ll \frac{1}{d}\;, \tag{7.371}$$

and approximating $\cos\chi = 1$, $\sin\chi = 0$ in the dipolar and the quadrupolar deflections, equations (7.368)-(7.370) are reduced to

$$\alpha^i_M = \frac{2GM}{c^2 d}(1+\cos\chi)n^i\;, \tag{7.372}$$

$$\alpha^i_D = \frac{4G}{c^2}\left[\frac{I^j(s_1)}{d^2}\left(n^i n^j - m^i m^j\right) - \frac{k^j\dot{I}^j(s_1)}{cd}n^i\right]\;, \tag{7.373}$$

$$\alpha^i_Q = \frac{4GI^{<jp>}(s_1)}{c^2 d^3}\left(n^i n^j n^p - n^i m^j m^p - m^i m^j n^p - m^i m^p n^j\right)\;, \tag{7.374}$$

where I^i and I^{ij} are determined by equations (7.349) and (7.350) respectively, the unit vectors $n^i = \xi^i/d$, $m^i = (\boldsymbol{k}\times\boldsymbol{n})^i$, and again all terms proportional to $\sin\chi$ have been neglected. On the other hand, the term with $\cos\chi$ in equation (7.372) must be retained as it reaches a magnitude of the order of 2 μas when $\chi = 90^o$ for Jupiter, and is about 1 μas when $\chi = 16^o$ for Saturn, that is the small-angle approximation in equation (7.372) is invalid for microarcsecond astrometric measurements.

Our calculation reveals that the deflection angles α^i_D and α^i_Q depend on the value of the dipole and quadrupole moments of the planet taken at the retarded time $s_1 = t_1 - r_1/c$, where t_1 is the time of observation. The deflections caused by the

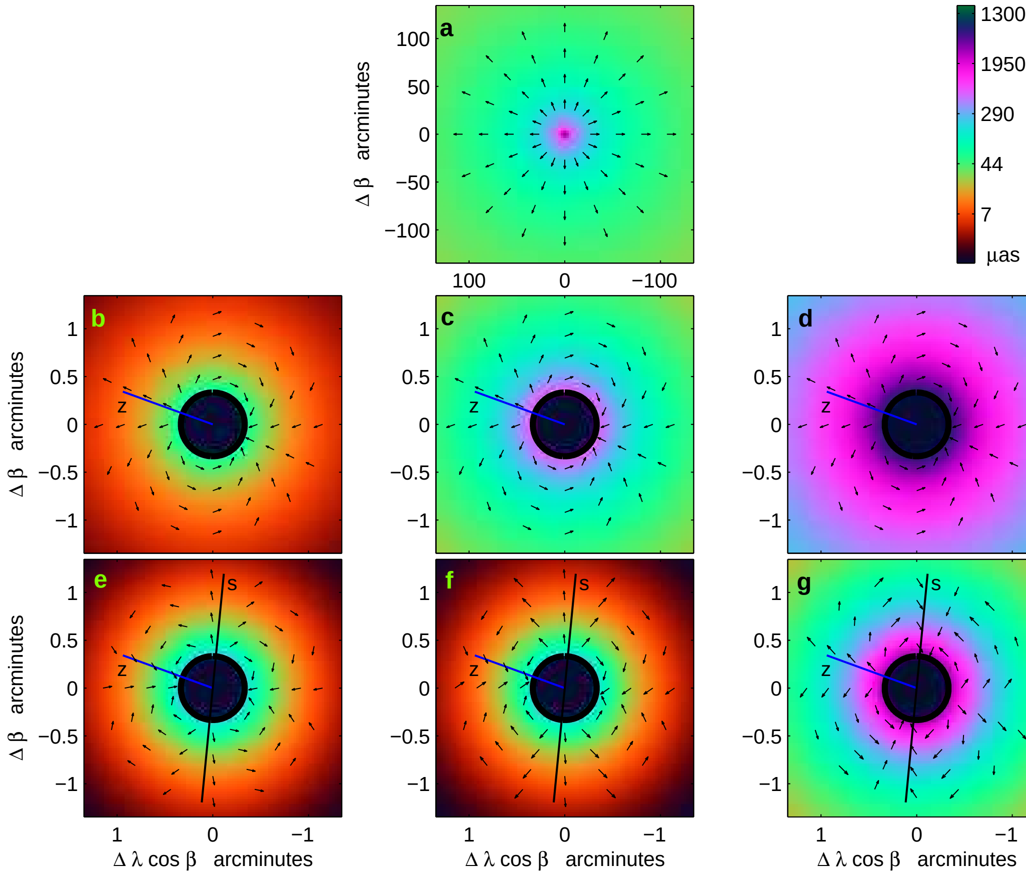

Figure 7.14 Gravitational deflection of light by the multipolar fields of Jupiter. The magnitude of deflection is given by color in a logarithmic scale shown at the top, and the small arrows in each plot indicate the direction of the deflection. The monopole deflection pattern is shown in plot (a). The dipolar deflection patterns caused by the shifting $L \sin\sigma$ planet's center of mass from the origin of the coordinate frame as projected on the plane of the sky, are depicted in the central row of the plots, for $(L \sin\sigma)_1 = 350$ (b), $(L \sin\sigma)_2 = 3500$ (c), and $(L \sin\sigma)_3 = 35000$ km (d). The bottom row of the plots (e-g) shows the combined quadrupolar deflection patterns generated by the intrinsic oblateness of the planet and by the translation. The direction of rotation axis s projected on the plane of the sky, is indicated with a thin black line. Note that the quadrupolar deflection at $(L \sin\sigma)_1 = 350$ km is dominated by that from the intrinsic quadrupole J_2, whereas for $(L \sin\sigma)_3 = 35000$ km, on the contrary, the translational quadrupolar deflection exceeds the light deflection caused by J_2.

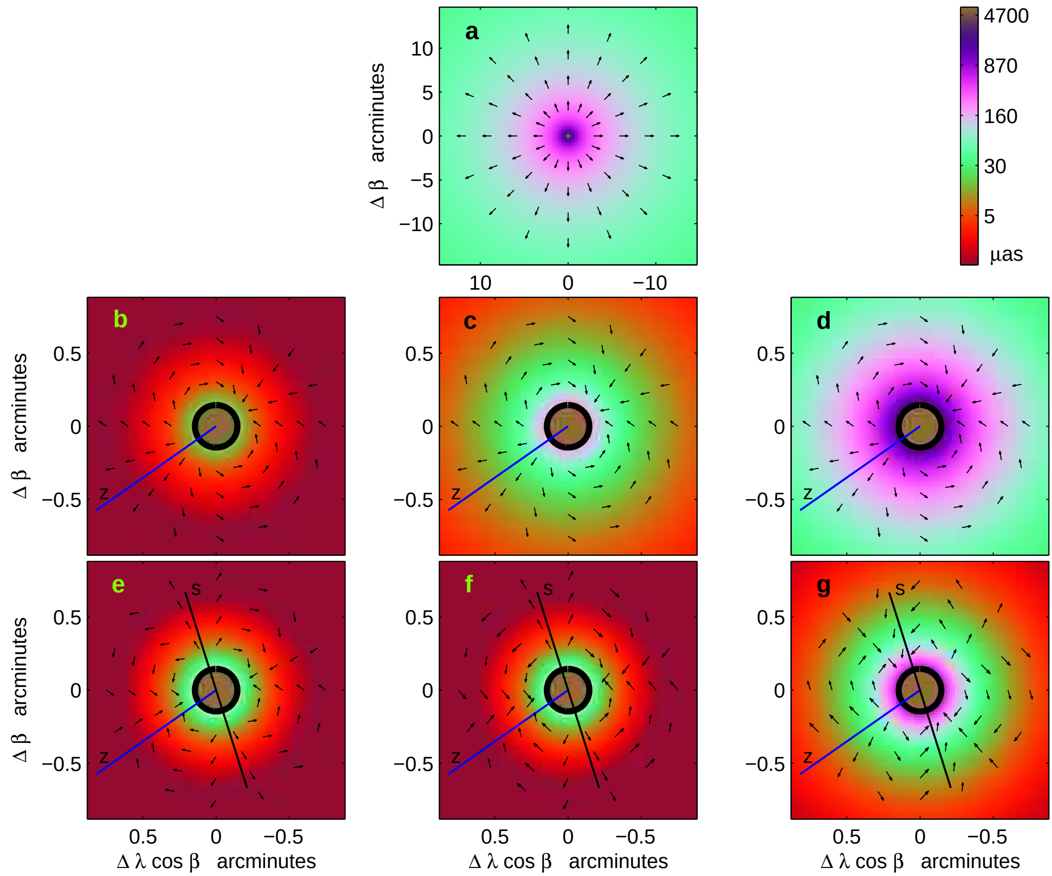

Figure 7.15 Deflection of light by Saturn caused by its mass (a), dipole (middle raw), and quadrupole (bottom raw) moments. Magnitudes of the deflection are expressed by color in the logarithmic grade shown in the upper right corner, and the direction of the deflection is shown by small arrows. Projected component of the vector of displacement of the center of mass of Saturn from the origin of the coordinate frame takes values (from left to right) of $(L\sin\sigma)_1 = 250$ (b), $(L\sin\sigma)_2 = 2500$ (c), and $(L\sin\sigma)_3 = 25000$ km (d). The lines associated with the unit vectors z and s indicate the directions of the displacement and rotational axis of the planet as projected on the plane of the sky respectively.

dipole translation x_P^i are nonphysical in general relativity. Hence, if general relativity is valid one can remove the dipolar deflection by choosing the origin of the coordinate system, used for calculation of the deflection angles, at the center of mass of the planet taken at the retarded instant of time s_1 with respect to observer. This situation is shown in Figs. 7.12, 7.13b where the past light cone is made of the null characteristics of the light rays coming from all stars to observer, and the future gravity null cone is made of the null characteristics of the gravity-field equation (7.344) that is the region of the causal interaction of the gravity force of the moving planet with any other particles including photons.

The deflection angle α^i is an observable and gauge-invariant quantity measured with respect to the unperturbed direction k^i of incoming photon defined in the asymptotically-flat spacetime at past null infinity [22)]. For this reason, the relativistic effects associated with the gravitational deflection of light have direct physical interpretation that is discussed in the rest of the present section.

7.9.5.2 ***Snapshot patterns***

Snapshot pattern of the gravitational deflection of light by a massive body is produced as a result of instantaneous measurement of the deflections of the apparent positions of all stars surrounding the light-ray deflecting body. It is a kind of instantaneous photographic picture of the stellar field with the light-ray deflecting body at its foreground. In order to get the snapshot deflection pattern produced by a massive planet one substitutes equations (7.349)-(7.352) to equations (7.372)-(7.374). It yields

$$\boldsymbol{\alpha}_M = \alpha\frac{1+\cos\chi}{2}\boldsymbol{n}\,, \tag{7.375}$$

$$\boldsymbol{\alpha}_D = \alpha\frac{L}{d}\Big[(\boldsymbol{z}\cdot\boldsymbol{n})\boldsymbol{n}-(\boldsymbol{z}\cdot\boldsymbol{m})\boldsymbol{m}\Big]\,, \tag{7.376}$$

$$\begin{aligned}\boldsymbol{\alpha}_Q = {} & \alpha J_2\frac{R^2}{d^2}\Big\{\Big[(\boldsymbol{s}\cdot\boldsymbol{n})^2-(\boldsymbol{s}\cdot\boldsymbol{m})^2\Big]\boldsymbol{n}-2(\boldsymbol{s}\cdot\boldsymbol{n})(\boldsymbol{s}\cdot\boldsymbol{m})\boldsymbol{m}\Big\} \\ & +\alpha\frac{L^2}{d^2}\Big\{\Big[(\boldsymbol{z}\cdot\boldsymbol{n})^2-(\boldsymbol{z}\cdot\boldsymbol{m})^2\Big]\boldsymbol{n}-2(\boldsymbol{z}\cdot\boldsymbol{n})(\boldsymbol{z}\cdot\boldsymbol{m})\boldsymbol{m}\Big\}\,,\end{aligned} \tag{7.377}$$

where

$$\alpha = \alpha_{\rm limb}\frac{R}{d}\,, \tag{7.378}$$

is the light-ray deflection angle caused by the planetary mass M,

$$\alpha_{\rm limb} = (1-\boldsymbol{k}\cdot\boldsymbol{\beta}_P)\,\frac{4GM}{c^2R} \tag{7.379}$$

is the deflection angle of the light ray grazing the planetary limb (R is the mean radius of the planet), $L=|\boldsymbol{x}_P(s_1)|$ is the displacement of the planetary center of mass from the origin of the coordinate system, the unit vector $\boldsymbol{z}=\boldsymbol{x}_P(s_1)/L$ points from the

22) This is a precise mathematical definition of the stellar positions in any astrometric catalog [Soffel, 1989]

origin of the coordinate system toward the center of mass of the planet, $\boldsymbol{\beta}_P = \boldsymbol{v}_P/c$, $\boldsymbol{v}_P = d\boldsymbol{x}_P/dt$ is the planet's velocity, and all time-dependent quantities like $\boldsymbol{x}_P$, $\boldsymbol{v}_P$, $\boldsymbol{z}$, and $\boldsymbol{s}$, are computed at the retarded time $s_1 = t_1 - r_1/c$ corresponding to the time of observation t_1.

Making use of conformal transformation technique [Damour and Esposito-Farèse, 1998b] one can prove that equations (7.375)-(7.377) are still applicable in the class of a generic scalar-tensor theory of gravity if a formal replacement: $\alpha \to \alpha_\gamma \equiv (1+\gamma)\alpha/2$ is made, where γ is the PPN parameter characterizing the impact of the scalar field on the curvature of space [Will, 1993]. Notice that one has moved the term depending on the first time derivative of the dipole moment from equation (7.376) to equation (7.375). This velocity-dependent term is purely radial (that is, directed along vector $\boldsymbol{n}$) and corresponds to the relativistic correction to the planetary mass M caused by the radial Doppler shift of light frequency. This gravitomagnetic correction correlates with the PPN parameter γ setting a natural limit on the precision of its measurement from a single-epoch experiment [Kopeikin, 2009]. In general relativity, $\gamma = 1$, and the maximal general-relativistic deflection of light on the planetary limbs are respectively: $\alpha_{\rm limb} \simeq 16280$ μas for Jupiter, and $\alpha_{\rm limb} \simeq 5772$ μas for Saturn [Brumberg *et al.*, 1990; Crosta and Mignard, 2006]. Taking into account in equation (7.379) the gravitomagnetic correction to the total mass, makes the radial deflection modulated up to a few μas, depending on the relative orbital velocity of the planet and observer. This radial-velocity correction to the mass may be a kind of interest in studying moving gravitational lenses [Wucknitz and Sperhake, 2004]. It is worthwhile to emphasize that some publications (see, for example, [Capozziello *et al.*, 1999; Gurvits and Mitrofanov, 1986; Sereno, 2002; Sereno and Cardone, 2002]) have an erroneous pre-factor $1 - 2\boldsymbol{k}\cdot\boldsymbol{\beta}_P$ in equation (7.379). Equation (7.379) with correct pre-factor $1 - \boldsymbol{k}\cdot\boldsymbol{\beta}_P$ was first derived by Birkinshaw & Gull [Birkinshaw and Gull, 1983], and re-confirmed later on by other authors [Frittelli, 2003b; Frittelli *et al.*, 2002; Kopeikin and Schäfer, 1999; Wucknitz and Sperhake, 2004].

It is worthwhile to emphasize that the entire theory of the microarcsecond light-ray deflection experiments with planets given in this section, is gauge-independent and Lorentz-invariant which means that all equations are valid in any arbitrary reference frame moving with respect to the original coordinate system $x^\alpha = (t, x^i)$ with a constant velocity V^i. Equations (7.375)-(7.377) of the light-ray deflection angles remain invariant with the velocity v_P^i replaced to $v_P^i - V^i$. Lorentz-invariance of equations (7.375)-(7.377) can be easily understood if one remembers that all quantities entering these equations are referred to the events connected by the null cones shown in Figure 7.11 and Figure 7.12 that represent the Minkowski diagram of the interaction of light with gravity that is invariant with respect to the Lorentz transformations by its own nature.

Equations (7.375)-(7.377) define the deflection patterns caused respectively by the monopole, dipole, and quadrupole components of the planetary gravitational field defined with respect to the coordinate system $x^\alpha = (t, x^i)$. Because velocity of motion of planets in the solar system are much smaller than the speed of light and gravity one can assume that during the time of propagation of photon across the solar system

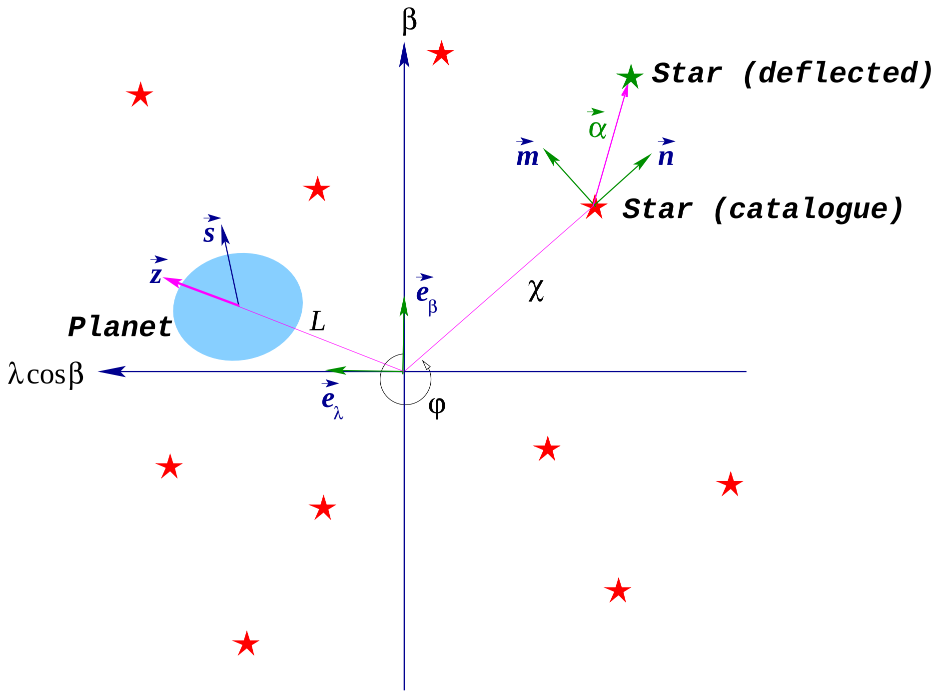

Figure 7.16 Local tangential coordinate system in the plane of the sky related to the global ecliptic coordinate system is shown. The unit vectors $\boldsymbol{e}_\lambda$ and $\boldsymbol{e}_\beta$ point in the directions of the increasing ecliptic longitude and latitude, respectively. The unit vector $\boldsymbol{k}$ lies along the line of propagation of light, is orthogonal to the plane of the sky and is directed toward observer (the reader). The angular distance of the unaffected position of a star from the origin of the coordinate frame is χ (c.f. Figure 7.10), which is at position angle φ from the north (direction $\boldsymbol{e}_\beta$). The total angle of the gravitational deflection of light of the star, $\boldsymbol{\alpha}$, has both radial, α_n, and ortho-radial, α_m, components: $\boldsymbol{\alpha} = \alpha_n \boldsymbol{n} + \alpha_m \boldsymbol{m}$.

the maximal displacement x^i_P of the center of mass of the planet that is allowed by the ephemeris systematic error or some other unexpected violation in the data processing computational approach, is smaller than the impact parameter, $L/d < 1$. Let us consider the light-ray deflection patterns for a single epoch of observation, t_1, mapping the deflections in the plane of the sky around the planet like it will appear in a photographic snapshot. To visualize the instantaneous two-dimensional deflection patterns, it is instructive to project equations (7.375)-(7.377) on the directions that are parallel and orthogonal to the ecliptic plane of the solar system as shown in Figure 7.16. The unit vectors n^i and m^i are expanded in such local tangential coordinates into

$$\boldsymbol{n} = \boldsymbol{e}_\beta \cos\varphi + \boldsymbol{e}_\lambda \sin\varphi \,, \tag{7.380}$$

$$\boldsymbol{m} = -\boldsymbol{e}_\beta \sin\varphi + \boldsymbol{e}_\lambda \cos\varphi \,, \tag{7.381}$$

where φ is called the position angle [Murray, 1983], the unit vectors $\boldsymbol{e}_\lambda$ and $\boldsymbol{e}_\beta$ are mutually orthogonal and directed along the increasing ecliptic longitude λ and latitude β respectively. The unit vectors $\boldsymbol{z}$ and $\boldsymbol{s}$ are expanded in this tangential frame as follows

$$\boldsymbol{z} = \boldsymbol{e}_\beta \sin\sigma\cos\psi + \boldsymbol{e}_\lambda \sin\sigma\sin\psi + \boldsymbol{k}\cos\sigma\,, \tag{7.382}$$

$$\boldsymbol{s} = \boldsymbol{e}_\beta \sin\Theta\cos\omega + \boldsymbol{e}_\lambda \sin\Theta\sin\omega + \boldsymbol{k}\cos\Theta\,, \tag{7.383}$$

where ψ and ω are respectively the position angles of the unit vectors $\boldsymbol{z}$ and $\boldsymbol{s}$ in the plane of the sky, and the angles σ and Θ characterize the degree of deviation of vectors $\boldsymbol{z}$ and $\boldsymbol{s}$ out of the plane of the sky.

Substituting equations (7.380)-(7.383) to equations (7.375)-(7.377) yields

$$\boldsymbol{\alpha}_M = \alpha\frac{1+\cos\chi}{2}\left(\boldsymbol{e}_\beta\cos\varphi + \boldsymbol{e}_\lambda\sin\varphi\right)\,, \tag{7.384}$$

$$\boldsymbol{\alpha}_D = \alpha\sin\sigma\frac{L}{d}\Big[\boldsymbol{e}_\beta\cos(\psi-2\varphi) - \boldsymbol{e}_\lambda\sin(\psi-2\varphi)\Big]\,, \tag{7.385}$$

$$\begin{aligned}\boldsymbol{\alpha}_Q = {} & \alpha J_2\sin^2\Theta\frac{R^2}{d^2}\Big[\boldsymbol{e}_\beta\cos(2\psi-3\varphi) - \boldsymbol{e}_\lambda\sin(2\psi-3\varphi)\Big] \\ & + \alpha\sin^2\sigma\frac{L^2}{d^2}\Big[\boldsymbol{e}_\beta\cos(2\omega-3\varphi) - \boldsymbol{e}_\lambda\sin(2\omega-3\varphi)\Big]\,.\end{aligned} \tag{7.386}$$

In this form, the equations clearly show that the directions of the deflections $\boldsymbol{\alpha}_M$, $\boldsymbol{\alpha}_D$ and $\boldsymbol{\alpha}_Q$ vary proportional to $\sin\varphi$, $\sin 2\varphi$ and $\sin 3\varphi$, respectively, as one goes from one star to another in the background stellar field around the planetary center. The amplitudes of the deflections α_M, α_D, and α_Q fall off as $1/\chi$, $1/\chi^2$, and $1/\chi^3$, respectively, where χ is the impact angle in the plane of the sky between the star and the coordinate system origin (see Figure 7.10). The reader should also notice the presence of the angles σ and Θ in the deflection angles $\boldsymbol{\alpha}_D$ and $\boldsymbol{\alpha}_Q$. The dipolar and quadrupolar deflections of light are maximal when the displacement $\boldsymbol{x}_P$ and the rotational axis of the planet are lying in the plane of the sky where $\sin\sigma = \sin\Theta = 1$. The deflection angles α_D and α_Q are reduced to zero if the rotational axis of the planet or its displacement from the origin of the coordinate system are along the line of sight, that is along vector $\boldsymbol{k}$. Let us point out that the term with J_2 was first derived by Klioner [1991b], re-discovered independently in papers [French *et al.*, 1993; Hubbard *et al.*, 1993] and thoroughly analyzed in works [Crosta and Mignard, 2006; Kopeikin and Makarov, 2007]. General formalism for deriving the quadrupole and all higher-order multipolar light-ray deflections in stationary gravitational field was published by Kopeikin [1997c]. It was extended later on to arbitrary time-dependent multipoles [Kopeikin and Korobkov, 2005; Kopeikin *et al.*, 2006].

Snapshot patterns of the deflections $\boldsymbol{\alpha}_M$, $\boldsymbol{\alpha}_D$, and $\boldsymbol{\alpha}_Q$ for a background field of stars surrounding Jupiter and Saturn are shown in Figure 7.14 and Figure 7.15 respectively, where one has used color grades to denote the magnitude of the deflection as a function of the impact angle χ of the light ray from the planet, and small arrows to specify the direction of the deflection, which depends on the position angle φ in accordance with equations (7.384)-(7.386). The patterns are shown for three different magnitudes $L\sin\sigma$ of the displacement vector $\boldsymbol{x}_P$: $(L\sin\sigma)_1 = 350$ km (left), $(L\sin\sigma)_2 = 3500$ km (middle), $(L\sin\sigma)_3 = 35000$ km (right) for Jupiter, and

$(L\sin\sigma)_1 = 250$ km (left), $(L\sin\sigma)_2 = 2500$ km (middle), and $(L\sin\sigma)_3 = 25000$ km (right) for Saturn.

The monopolar deflection of light $\boldsymbol{\alpha}_M$ is depicted in subplots (a) of Figure 7.14 and Figure 7.15. It is purely radial (that is , takes place along vector $\boldsymbol{n}$), and does not depend on the magnitude and direction of the displacement vector $\boldsymbol{x}_P$ but depends on the radial velocity $\boldsymbol{k}\cdot\boldsymbol{v}_P$ of the planet in correspondence with equation (7.379). The dipolar deflection of light, $\boldsymbol{\alpha}_D$, is directly proportional to the magnitude $L\sin\sigma$ of the displacement of the planet's center of mass from the origin of the coordinate frame as projected on the plane of the sky, and its orientation is determined by the component of vector $\boldsymbol{x}_P$ projected on the plane of the sky. The quadrupolar deflection of light, $\boldsymbol{\alpha}_Q$, is quadratically proportional to the magnitude of the projected displacement $L\sin\sigma$ between the origin of the coordinate frame and the planet's center of mass as well as on the magnitude of the intrinsic quadrupole moment of the planet projected on the plane of the sky $J_2\sin\Theta$. The orientation of the quadrupolar deflection pattern is determined by the projections of vectors $\boldsymbol{z}$ and $\boldsymbol{s}$ on the plane of the sky. When the displacement $L\sin\sigma \ll \sqrt{J_2}R\sin\Theta$, the effect of the translational quadrupole $Mx_P^{<i}x_P^{j>}$ on the light deflection is relatively small, and the orientation of the quadrupolar pattern is defined by the projection of $\boldsymbol{s}$ on the plane of the sky as shown in subplots (e) in Figure 7.14 and Figure 7.15. If $L\sin\sigma \simeq \sqrt{J_2}R\sin\Theta$, the quadrupolar pattern may be complicated and hard to interpret as the deflection of light by the translational quadrupole $Mx_P^{<i}x_P^{j>}$ is comparable to that caused by the planetary intrinsic quadrupole J_2. If, for whatever reason, $\sqrt{J_2}R\sin\Theta \ll L\sin\sigma$, the deflection of light by the intrinsic quadrupole moment of the planet is swamped by the much stronger deflection caused by the translational quadrupole as shown in subplots (g) in Figure 7.14 and Figure 7.15.

7.9.5.3 *Dynamic patterns of the light deflection*

Equations (7.384)-(7.386) can be also used to work out a dynamic visualization of the light deflection patterns as a function of time reckoned at the point of observation. In other words, one will be interested in finding apparent trajectory of a visible position of star as a function of the relative angular position of the star with respect to the planet. Let us stipulate that planet is moving on the celestial sphere below a star along a great circle with a minimal separation angle $\chi_{\rm m} \equiv X_0$ between the star and the planet. As the light-ray deflection angle is rather small compared with $\chi_{\rm m}$, one can project this motion onto the plane of the sky so that it represents a straight line in the local coordinates of Figure 7.16. One further assumes for simplicity that this straight line is parallel and sufficiently close to the ecliptic, which is a rather realistic approximation since the orbital inclinations of Jupiter and Saturn with respect to ecliptic are small and do not exceed 2.5°. Nevertheless, in real astrometric experiments the inclination of the planetary orbits to the ecliptic must be taken into account. Equations (7.384)-(7.386) can be viewed as a one-to-one mapping of the planetary position specified by coordinates (X, Y) in the plane of the sky, to the locus of the star image at coordinates (x, y) continuously shifted by the planetary gravitational field. The geometry of the gravitational lens mapping is explained in more detail in Figure 7.17 and the book by Lawrence [1972] provides further insight to the mathematical structure of the central mapping. The gravity-unaffected position of the star, when the planet is at "infinity", is the center of the mapping which is also the origin of both coordinate frames: (X, Y) and (x, y) (notice that one has shifted Y axis along X axis for making better graphical visualization). In other words, the gravity-unaffected position of the star is at the point $x = 0$, $y = 0$ where also $X = 0$, $Y = 0$. The planet moves along a straight line parallel to Y axis at a constant separation $X = X_0$ from the gravity-unaffected position of the star. When the planet is at the position angle ϕ in the coordinates (X, Y), the deflected position of the star is at the position angle θ in the coordinates (x, y). Notice that in general, $\theta \neq \phi$ because the dipolar and quadrupolar light-ray deflections are not purely radial. Precise correspondence between the two angles is established by the central mapping transformation equations. The angle ϕ runs from $-\pi/2$ to $+\pi/2$ while the angle θ runs in the most common case from 0 to 2π (see Figure 7.17 and the text below).

Two components of the deflection angle $\boldsymbol{\alpha}_M$ are identified with the coordinates x and y: $\boldsymbol{\alpha}_M \equiv (x, y)$, and substitutions $\sin\phi = Y/d$, $\cos\phi = X/d$, $d = \sqrt{X^2 + Y^2} = \sqrt{X_0^2 + Y^2}$ are made in equation (7.384) where the approximation $\cos\chi = 1$ is used for simplicity. It yields the monopolar light-deflection mapping equations in the following form

$$x = 2r\,\frac{X^2}{X^2 + Y^2}\,, \qquad\qquad y = 2r\,\frac{XY}{X^2 + Y^2}\,, \tag{7.387}$$

where $r \equiv 2GM/(c^2 X)$, so that $2r$ is the maximal deflection angle reached at the time when the angular distance between the planet and the star is minimal ($Y = 0$). The mapping given by equations (7.387) is purely radial, which means that the gravity-deflected visible position of the star and the planet lie on a straight line passing through the center of the mapping (origin of the coordinates), and $\theta = \phi$ (see Figure 7.17). As the planet moves from $\phi = -\pi/2$ to $\phi = +\pi/2$ along Y axis, the deflected position of the star outlines the circle

$$(x - r)^2 + y^2 = r^2\,, \tag{7.388}$$

with the radius r and the origin located at the point $(x = r, y = 0)$. In polar coordinates $x = \rho\cos\theta$, $y = \rho\sin\theta$, the correspondence between θ and ϕ is simple: $\theta = \phi$, and pure radial as one has pointed out earlier. The monopolar deflection curve in the polar coordinates is

$$\rho = 2r\cos\phi\,, \tag{7.389}$$

which yields the circle shown in the left graph of Figure 7.18, where $x = \Delta\beta$, $y = \Delta\lambda$.

The dynamic curve of the dipolar deflection can be easier represented in coordinates (x', y') that are rotated clockwise with respect to the (x, y) coordinates through angle ψ, where ψ defines direction of vector $\boldsymbol{z}$ in the plane of the sky by equation (7.382),

$$x' = x\cos\psi - y\sin\psi\,, \qquad\qquad y' = x\sin\psi + y\cos\psi\,. \tag{7.390}$$

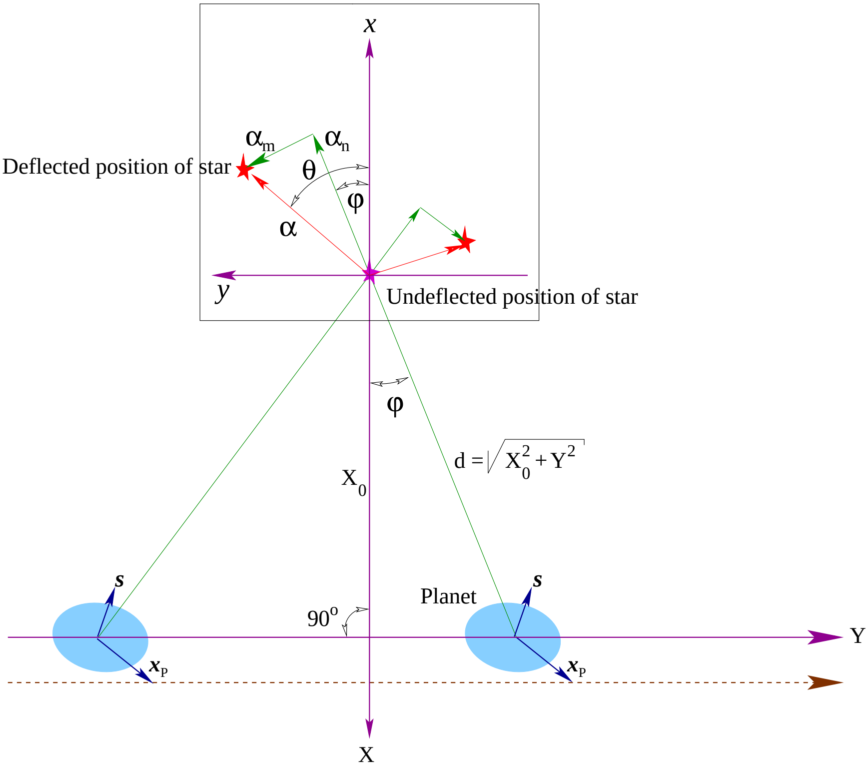

Figure 7.17 Plane geometry of the central mapping used for computation of the dynamic pattern (curve) of the gravitational deflection of the apparent position of star by planet as a function of time. The origins of the two coordinate charts (x, y) and X, Y coincide and are placed at the gravity-unaffected (catalog) position of the star. For the sake of convenience, Y axis is shown as displaced along the positive direction of X axis at the constant distance X_0. The available planetary ephemeris predicts that the planet moves continuously from left to right along Y axis with its X coordinate being fixed. Two positions of the planet are shown to explain the mapping geometry. The dashed line is parallel to Y axis and shows trajectory of the planet's center of mass which differ from the planetary ephemeris,thus, giving rise to the dipolar component of the light deflection. The minimal distance between the planet and the unaffected position of the star is $X = X_0$. Apparent position of the star is deflected from its unperturbed (catalog) position in both radial, α_n, and ortho-radial, α_m, directions respectively, with the total deflection $\boldsymbol{\alpha} = \alpha_n \boldsymbol{n} + \alpha_m \boldsymbol{m}$. The mapping establishes a one-to-one correspondence F between the coordinates (X, Y) of the planetary position and the coordinates (x, y) of the deflected position of the star, that is $F : (d \to \alpha, \phi \to \theta)$ where $d = \sqrt{X_0^2 + Y^2}$.

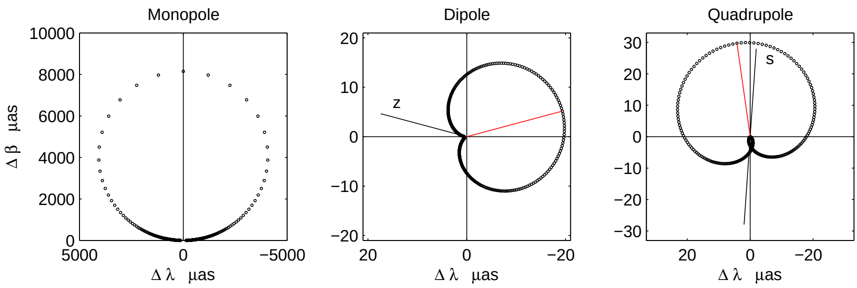

Figure 7.18 Left, middle, and right graphs show respectively the monopolar (circle), dipolar (cardioid), and quadrupolar (the Caley's sextic) light-ray deflection curves. These curves are obtained by the central mapping transform explained in Figure 7.17 and section 7.9.5.3. The planet moves with a constant velocity and passes directly below the catalog position of the star associated with the origin of the coordinates. One has chosen the impact parameter $X_0 = 40$ arcsecond, and the orbital position's error projected on the plane of the sky, is $L \sin\sigma = 250$ km. The time step between the individual points is 0.02 days for the monopolar deflection curve, and 0.002 days for the other two curves. Dynamic light-ray deflection curves for Saturn have smaller amplitude for the same value of the impact parameter but similar shape. Red lines on the middle and right graphs show direction of the coordinate axis x' and $\hat{x}$ rotated clockwise with respect to x axis at the angles ψ and 2ω respectively (see text).

In the new coordinates, the dipolar mapping transformation is deduced from equation (7.385) and is given by

$$x' = 2p\,\frac{X^2(X^2 - Y^2)}{(X^2 + Y^2)^2}\,, \qquad y' = 4p\,\frac{X^3 Y}{(X^2 + Y^2)^2}\,, \tag{7.391}$$

where $p \equiv rL\sin\sigma/X$. Besides the radial deflection component, the dipolar mapping generally contains an ortho-radial deflection component as follows from equation (7.391) and demonstrated in Figure 7.17. As the planet moves along Y axis and the angle ϕ changes from $\phi = -\pi/2$ to $\phi = +\pi/2$, the deflected position of the star outlines the curve

$$(x'^2 + y'^2 - px')^2 = p^2(x'^2 + y'^2)\,, \tag{7.392}$$

that should be superimposed on the monopolar deflection curve (the circle). The dipolar deflection curve alone is shown in the middle graph of Figure 7.18. This curve, called cardioid [Lawrence, 1972], is symmetric around the axis x' with a cusp at the origin. In polar coordinates, $x' = \rho\cos\theta'$, $y = \rho\sin\theta'$, $\theta' = \theta + \psi$, the correspondence between θ' and ϕ is $\theta' = 2\phi$, and in these coordinates the cardioid curve takes the form

$$\rho = p\,(1 + \cos 2\phi)\ . \tag{7.393}$$

Its orientation depends on the angle ψ, and the magnitude on the ratio $L\sin\sigma/X$. Notice that the vector $\boldsymbol{x}_P$ was kept constant in one's calculation. In reality, the planet can have a residual transverse velocity with respect to its pre-calculated ephemeris trajectory because of the statistical errors in construction of the planetary epehemris. In this case, the dipolar deflection curve will be slightly different from cardioid.

The dynamic trajectory of the stellar apparent position due to the quadrupolar deflection of light becomes simple in the plane-of-the-sky coordinates $(\hat{x}, \hat{y})$ rotated clockwise with respect to the (x, y) coordinates through an angle 2ψ,

$$\hat{x} = x \cos 2\psi - y \sin 2\psi \, , \qquad \hat{y} = x \sin 2\psi + y \cos 2\psi \, . \tag{7.394}$$

In these new coordinates, the quadrupolar mapping transformation may now be inferred from equation (7.386)

$$\hat{x} = 4q \frac{X^4(X^2 - 3Y^2)}{(X^2 + Y^2)^3} \, , \qquad \hat{y} = 4q \frac{X^3 Y(3X^2 - Y^2)}{(X^2 + Y^2)^3} \, , \tag{7.395}$$

where the parameter $q \equiv q_L = (r/2)(L \sin \sigma / X)^2$ for the deflection component caused by the translational quadrupole $M x_P^{<i} x_P^{j>}$, and $q \equiv q_J = (r/2) J_2 (R \sin \Theta / X)^2$ for the term generated by the projected value of the intrinsic quadrupolar oblateness J_2 of the planet. The quadrupolar mapping contains an ortho-radial deflection besides the radial component. As the planet moves from $\phi = -\pi/2$ to $\phi = +\pi/2$ along Y axis, the deflected position of the star describes the curve

$$4(\hat{x}^2 + \hat{y}^2 - q\hat{x})^3 = 27q^2(\hat{x}^2 + \hat{y}^2)^2 \, , \tag{7.396}$$

which must be superimposed on the monopolar and dipolar deflection curves. The quadrupolar deflection alone is shown in the right graph of Figure 7.18. This plane algebraic curve is known as the Caley sextic [Lawrence, 1972]. It is symmetric around axis $\hat{x}$ and has a small secondary loop near the origin of the coordinates. The polar equation of the Caley sextic is obtained after substitution of the polar coordinates $\hat{x} = \rho \cos \hat{\theta}$, $\hat{y} = \rho \sin \hat{\theta}$, $\hat{\theta} = \theta + 2\psi$, into equation (7.396), and solving the cubic equation with respect to ρ. The quadrupolar mapping correspondence between the polar angles $\hat{\theta}$ and ϕ is $\hat{\theta} = 3\phi$, and the polar equation of the Caley sextic caused by the translational quadrupole

$$\rho = q_L (\cos 3\phi + 3 \cos \phi) \; . \tag{7.397}$$

The sextic curve for the deflection caused by the oblateness J_2, takes the same form after replacing parameter $q_L \to q_J$ and the angle $\psi \to \omega$. The quadrupolar deflection curve has a secondary loop passing through the origin of coordinate axes as shown in Figure 7.18. The cross-over point of the secondary loop is 1/8 of the maximum deflection $\rho = 4q_L$ in the mid-point of the primary loop. This means that the magnitude of the secondary loop is about few microarcsecond for Jupiter that is within the range of near-future experimental measurement by SIM, Gaia, and SKA.

Figure 7.19 shows the combined deflection trajectory of the star's position from the dipole cardioid and the quadrupole sextic presented in Figure 7.18. The resulting curve looks rather similar to cardioid of a somewhat larger amplitude, and is inclined at some intermediate angle with respect to the cardioid and sextic symmetry

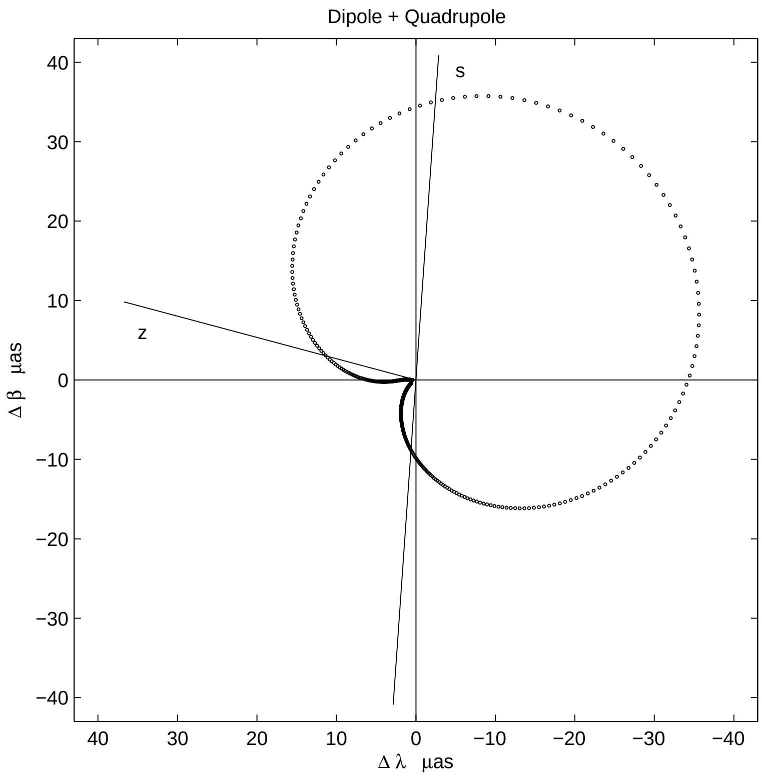

Figure 7.19 Combined dipolar and quadrupolar deflection of light by Jupiter for the same parameters and configuration as in Figure 7.18. Details of the quadrupolar deflection (the sextic curve) are hidden by the dipolar deflection (the cardioid). Separation of the two deflection terms is a non-trivial experimental problem, and generally will require multiple observations of a number of sources around the planet.

axes. Obviously, this curve swamps the quadrupolar deflection and can be confused with a pure dipolar one, because vector $\boldsymbol{x}_P$ is not known precisely before observations. This demonstrates the crucial importance of accurate determination of the spurious dipole component in any experiment designed to measure the intrinsic gravitational quadrupole J_2 by making use of the gravitational deflection of light. Since the ephemeris error is expected to be a smooth function of time, microarcsecond observations of multiple stars at different position angles to the direction of the dipole axis should help to separate these two contributions to the total deflection pattern.

7.9.6
Testing Relativity and Reference Frames

Most papers calculating light deflection in the solar system stipulate that the massive body deflecting the light, does not move. However, making use of this assumption in microarcsecond astrometric experiments with planets is inadequate as the plant moves as light propagates through the solar system, and it is important to know precisely how light interacts with the changing gravitational field of the planet. Our formalism is based on rigorous general-relativistic paradigm and takes into account the effect of planet's motion on the deflection angle through the retarded solution (7.345)-(7.347) of the Einstein equations where the retardation is due to the finite speed of gravity showing that gravity field is causal and operates on the null cone as shown in Figure 7.12. The planetary motion affects the data of light-ray deflection measurements in two ways [Kopeikin, 2009]:

- explicitly - in the form of terms depending on velocity, $\boldsymbol{v}_P$, acceleration, $\boldsymbol{a}_P$, etc., of the planet with respect to the coordinate frame;
- implicitly - in the form of the retarded time, $s_1 = t_1 - r_1/c$, in the time argument of coordinates, $\boldsymbol{x}_P = \boldsymbol{x}_P(s_1)$ of the planet as well as in its velocity and acceleration.

General relativity can be violated in higher-order approximations, which may become detectable in the near future by astrometric facilities operating at the 1 μas level of precision. This presumable violation of general relativity can be parameterized and tested in light-ray deflection experiments. Verification of general relativity and detection of beyond-Einstein phenomena is the main goal of experimental fundamental physics in space that includes the search for scalar fields, observation of gravitomagnetic and retardation-of-gravity effects associated with the time-dependence of planet's gravitational field, measurement light-deflection quadrupolar anisotropy, and more [Turyshev, 2009a]. This kind of astrometric experiments will also lead to essential improvement in accuracy of fundamental reference frames and planetary ephemerides in the solar system [Malkin *et al.*, 2009; Sovers *et al.*, 1998; Walter and Sovers, 2000]. In what follows, these experimental opportunities are discussed in more detail by analyzing different modes of the gravitational light-ray deflection.

7.9.6.1 ***The monopolar deflection***
The monopolar deflection of light is given by equations (7.375), (7.378), (7.379). It is standard experimental practice [Fomalont *et al.*, 2009a; Lambert and Le Poncin-

Lafitte, 2009] to introduce the PPN parameter γ leading to the replacement [Will, 1993]

$$\alpha \rightarrow \alpha_\gamma \equiv \frac{1+\gamma}{2}\alpha \tag{7.398}$$

in equation (7.375) as well as equations (7.376), (7.377). The parameter γ characterizes deviation of gravity from a pure geometry and is associated with the presence of hypothetical scalar fields remaining from the epoch of the Big Bang [Damour and Nordtvedt, 1993a]. This parameter has been measured in the solar light-ray deflection experiments with a precision approaching to the level of $10^{-4} \div 10^{-5}$ [Bertotti *et al.*, 2003; Fomalont *et al.*, 2009a; Kopeikin *et al.*, 2007; Lambert and Le Poncin-Lafitte, 2009].

Orbital motion of a planet generates h_{0i} component of the gravitomagnetic field which depends on the time derivative of the dipole moment as appears in equation (7.346). This gravitomagnetic component leads to the *explicit* dependence of the monopolar light-ray deflection on the line-of-sight component of the orbital velocity $\boldsymbol{v}_P$ of the planet, effectively modulating the planetary mass as shown in equation (7.379). The gravitomagnetic term was derived in [Kopeikin and Schäfer, 1999] and analytically confirmed in [Frittelli, 2003a,b; Klioner, 2003b; Wucknitz and Sperhake, 2004]. It can reach about 1.6 μas for a star observed at the limb of Jupiter.

The gravitomagnetic term is the result of the asymmetric exchange of energy between the photon and the gravitational field of the moving body. It is well-known [Brumberg, 1991] that in case of a static gravitational field the overall angle of light deflection does not depend on the energy (frequency) of the photon. The photon gains energy as it approaches the gravitating body, and loses it as it moves away. If the body is at rest with respect to the chosen coordinate system, the gain and the loss of energy cancel out, but this balance is violated if the body is moving radially. Appearance of the Doppler modulation of the deflection of light can be also understood if one remembers that the deflection angle is associated with spatial components of the null vector of photon. Lorentz transformation from a static frame of the planet to the moving frame of observer transforms and mixes all four components of the wave vector and its time component affects the spatial components in the moving frame. This is clearly shown in the paper by Klioner [Klioner, 2003b].

The gravitomagnetic deflection of light can be measured in a single-epoch observation only if (1) mass of the planet is known sufficiently well; (2) the parameter γ is excluded since it correlates linearly with the gravitomagnetic term [Kopeikin, 2009]. On the other hand, if one can conduct observations in different epoches, the gravitomagnetic term in equation (7.379) will periodically modulate the radial deflection because of the relative motion of planet with respect to observer so that the line-of-sight velocity changes. Hence, in this case it can be measured independently of the parameter γ. In case of Jupiter the amplitude of this gravitomagnetic modulation reaches 3.2 μas and has a main period of one year due to the orbital motion of the Earth [Kopeikin and Makarov, 2007].

7.9.6.2 *The dipolar deflection*

The dipolar deflection of light is given by equations (7.376), (7.378) where the coordinates of the planetary center of mass $\boldsymbol{x}_P = \boldsymbol{x}_P(s_1)$ depend on the retarded instant of time $s_1 = t_1 - r_1/c$ [23] which is a solution of the gravity null-cone equation (7.348). These retarded coordinates of the planet define the dipole moment $\boldsymbol{I}(s_1) = M\boldsymbol{x}_P(s_1)$ in the multipolar expansion of the gravitational field of the planet. In general relativity, the law of conservation of linear momentum allows us, at least theoretically, to eliminate the dipole moment by placing the center of mass of the planet at the origin of the coordinate frame.

If $\boldsymbol{x}_P(s_1) = 0$, all coordinate-dependent effects in the gravitational deflection of light vanish. It means that the planet deflects light from its retarded position defined by the retarded solution of the gravity wave equation (7.344) as shown in Figure 7.12. This retardation is due to the fact that light and gravity interact on the hypersurface of intersection of two null cones - the past light cone of observer and the future gravity cone of the moving planet. This important consequence of general relativity was pointed out in [Kopeikin, 2001] and was experimentally confirmed in 2002 light-ray deflection experiment with Jupiter [Fomalont and Kopeikin, 2003] with precision $\sim 20\%$ [24]. SIM, Gaia, and SKA can improve this measurement by, at least, a factor of 10.

In a practical experimental setup, the retardation of gravity effect can be parameterized by introducing a parameter ϵ to the retarded time, $s \to s_\epsilon = t - \epsilon\, r/c$ [Kopeikin, 2004] in the dipolar light-ray deflection angle (**??**). If $\epsilon = 0$ the gravity null cone degenerates and the gravitational field of the planet deflects photons instantaneously. General relativity postulates that $\epsilon = 1$, which corresponds to the case of gravity propagating with the same speed as light. This consideration makes it evident that the position of Jupiter's center of mass in data processing algorithm must be taken at the retarded instant of time $s_1 = t_1 - r_1/c$ with respect to observer.

For arbitrary $\epsilon \neq 0$ (that is when the speed of gravity and light differ) light is deflected by the planet from its orbital position taken at the time s_ϵ. It leads to a non-vanishing displacement vector $\boldsymbol{L}$

$$\boldsymbol{L} \equiv \boldsymbol{x}_P(s_1) - \boldsymbol{x}_P(s_\epsilon) = \boldsymbol{v}_P(\epsilon - 1)\frac{r_1}{c} - \frac{1}{2}\boldsymbol{a}_P(\epsilon - 1)^2\frac{r_1^2}{c^2} + O\left[(\epsilon - 1)^3\right] , \tag{7.399}$$

where $\boldsymbol{v}_P$ and $\boldsymbol{a}_P$ are velocity and acceleration of the planet respectively. If gravity null cone is different from the light cone, it brings about the dipolar anisotropy

$$\begin{aligned} \boldsymbol{\alpha}_D &= \alpha(\epsilon - 1)\frac{r_1}{cd}\Big[(\boldsymbol{v}_P\cdot\boldsymbol{n})\boldsymbol{n} - (\boldsymbol{v}_P\cdot\boldsymbol{m})\boldsymbol{m}\Big] \\ &- \alpha(\epsilon - 1)^2\frac{r_1^2}{2cd}\Big[(\boldsymbol{a}_P\cdot\boldsymbol{n})\boldsymbol{n} - (\boldsymbol{a}_P\cdot\boldsymbol{m})\boldsymbol{m}\Big] , \end{aligned} \tag{7.400}$$

23) One reminds that $t = t_1$ is the time of observation and $r = r_1$ is the distance between the observer and the planet, and c is the speed of gravity.

24) Further discussion of the gravitational physics of the Jupiter's experiment is given in review article by Kopeikin and Fomalont [2006] that also analyzes a number of misconceptions in physical interpretation of the experiment presented in [Will, 2006].

which can be measured to evaluate the retardation of gravity parameter ϵ. For a star observed on the limb of Jupiter $d = R$, where R is the radius of the planet, the velocity-dependent term in equation (7.400) is $\sim \alpha_{\rm limb}(v/c)(r_1/R) \simeq 8140$ μas, and the acceleration-dependent term is $\sim \alpha_{\rm limb}(v^2/c^2)(r_1/2R) \simeq 0.2$ μas. The reader should notice that the dipolar deflection (7.400) depends on the transversal component of the planetary velocity only. Since the velocity of the planet is almost lying in the ecliptic plane, it makes the position angle $\psi = -\pi/2$ of the unit vector $\boldsymbol{z} = \boldsymbol{L}/L$ in equation (7.382) so that the symmetry axis of the cardioid in Figure 7.18 gets aligned with Y axis. This also allows us to physically interpret the dipolar deflection associated with the retardation of gravity as caused by the dragging of light ray by gravitomagnetic field of Jupiter originating from the first time derivatives of its gravitational field as it travels along its orbit [Arakida, 2004; Kopeikin, 2003b, 2006; Schäfer and Bartelmann, 2006; Sereno, 2005, 2008].

A VLBI experiment was conducted in 2002 [Fomalont and Kopeikin, 2003] to measure the speed of gravity through the search for the residual dipolar deflection of light of quasar J0842+1835 caused by the moving gravitational field of Jupiter that could arise due to the presumable difference between the speed of gravity and light. The experiment did not show any deviation from general relativistic model of light deflection taking into account the retarded nature of gravitational field of moving planet.

Assuming that general relativity is fully valid, the measurement of the dipolar anisotropy of light deflection can be used to determine position of the center of mass of the planet on its orbit more accurately than any other currently available astrometric techniques [Kopeikin and Makarov, 2007; Malkin *et al.*, 2009].

7.9.6.3 *The quadrupolar deflection*

It is evident from equation (7.377) that in order to measure the quadrupolar deflection of light caused by the planetary oblateness J_2, the following condition on the displacement vector $\boldsymbol{x}_P$ must be satisfied

$$L \sin\sigma \le \Delta_L < \sqrt{J_2} R \sin\Theta\,, \tag{7.401}$$

where Δ_L is the measurement error of $L\sin\sigma$. If condition (7.401) is fulfilled, all terms in the second line of equation (7.377) depending on the ratio $(L/d)^2$ can be abolished. In fact, $\Delta_L \neq 0$ because the planetary ephemeris is known with a limited accuracy. Hence, one should consider vector $\boldsymbol{x}_P$ as a fitting parameter that is to be determined from the light-ray deflection observation by successive iterations. JPL ephemerides provides the initial value of $\boldsymbol{x}_P$ for such iterations, which can be improved if the measurement accuracy on the deflected positions is high enough [Kopeikin and Makarov, 2007]. For Jupiter $\sqrt{J_2}R \simeq 8500km$, so that in order to measure the intrinsic quadrupole moment J_2 of the planet its orbital position must be known to $\Delta_L < 8500$ km. Currently available JPL ephemerides for major solar system bodies, determined from direct optical/radio observations and spacecraft tracking, are believed to be more accurate [Pireaux *et al.*, 2007; Standish, 2004]. However, one should be careful in taking correct value of time in the position of the planet. A non-zero value of the dipole moment of gravitational field emerges if the time of observation t_1 instead of the retarded time $s_1 = t_1 - r_1/c$ is used in the light-ray deflection equations whenever position of the planet is computed, which corresponds to $\epsilon = 0$ in equation (7.399). In this case $L \sin\sigma \approx 35000$ km for Jupiter, and condition (7.401) is violated, thus, making the intrinsic quadrupolar deflection of light unmeasurable.

After the displacement $\boldsymbol{x}_P$ is determined from the observed dipole anisotropy in the light-ray deflection pattern, one can shift the origin of the coordinate system in the sky to the center-of-mass of the light-ray deflecting planet to suppress the spurious dipolar term as much as possible. If general relativity is a correct theory of gravity, this will reduce both the dipolar and quadrupolar deflections caused by the translational displacement $\boldsymbol{L}$ of the planetary center of mass. On the other hand, if general relativity were violated one would not be able to remove the dipole anisotropy because the structure of the light-ray deflection equations in alternative theory of gravity is different from general-relativistic equations (7.375)-(7.377). This would make more difficult to measure J_2 in the models of the light deflection based on an alternative theory of gravity.

8
Relativistic Geodesy

8.1 Introduction

Geodesy is concerned with the precise determination of the area and physical surface of the Earth, and in particular the definition of a reference Earth's shape known as the geoid [Hofmann-Wellenhof and Moritz, 2005; Moritz, 1980]. The geoid is an equipotential surface of gravity, fairly close to ellipsoidal (oblate spheroid) in shape because of the counter-gravity centrifugal forces generated by the rotation of the Earth about its axis, and fairly irregular because of the variability in composition (matter's density) of the Earth beneath each point on the geoid [1)] (see figure-geodesy1). The reference-ellipsoid model of the Earth attempts to define its shape in terms of a smooth ellipsoid, and the use of satellite measurements have led to the development of the World Geodetic System (WGS-84) ellipsoid as the best ellipsoidal representation of the geoid. The maximum difference between the geoid and the WGS-84 ellipsoid is 50-100 meters ($\delta r/r_\oplus \sim 10^{-5}$). Ellipsoids have also been constructed for individual continents and countries because different ellipsoids give better fits to the geoid at different locations, for example, the Clarke 1866 ellipsoid for the United States or the Krasovsky 1940 ellipsoid for the USSR.

The goal of this chapter is to introduce the reader to the problems related to manifestation of relativistic effects in gravimetric and geodetic measurements, which have recently reached a necessary level of precision. Relativistic geodesy on the ground is a very promising method to bridge the gap between geometrical navigation and gravity field determination in establishing homogeneous world height system [Müller *et al.*, 2008a; Schlüter and Behrend, 2007; Sovers *et al.*, 1998]. Relativistic corrections to the Newtonian formulas of classical gravimetry and geodesy are characterized by two small parameters: $\varepsilon \sim \omega^2 R_E^2/c^2 \simeq 2.3\times10^{-12}$ and $\eta \sim GM_E/c^2R_E \simeq 0.7\times10^{-9}$, where G is the Newtonian gravitational constant, c - the fundamental speed, ω_E - the angular velocity of the Earth's rotation, R_E and M_E are the mean radius and mass of the Earth respectively. Registration of relativistic effects can be possible, if and only if, a relative precision of measuring devices is comparable with the magnitude of parameters ε and η.

Modern technology allows to build high-precision superconducting gravimeters that are a spring type gravimeters in which the mechanical spring is replaced by a magnetic levitation of a superconducting sphere in the field of superconducting, persistent current coils. The superconducting gravimeters are the most sensitive to various periodic variations of the Earth's gravity acceleration $g_\oplus$ [2)] with a magnitude of the order of $\leq 10^{-3}$ μgal from periods of a few thousand seconds to the monthly and annual solid earth tides and the Chandler wobble [Goodkind, 1999]. It exceeds precision of the best ballistic gravimeters that is about 1 μgal [Goodkind, 1999; Torge, 1990].

1) The geoid's irregularities are described by differentiable functions and are not necessarily coincident with those exhibited by the real Earth's surface which is not differentiable in many places like mountains.

2) $g_\oplus = 981$ gal, where 1 gal = 1 cm/s$^2 \simeq 10^{-3}\, g_\oplus$.

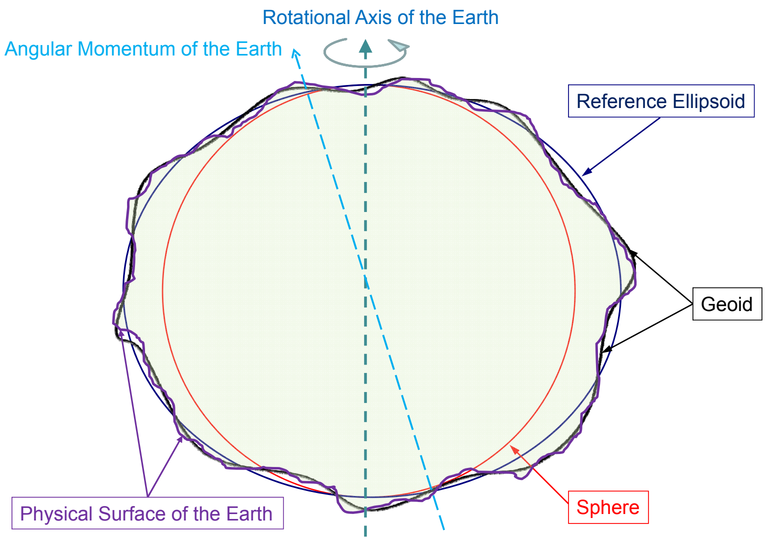

Figure 8.1 Physical surface of the Earth and its relationship to various reference surfaces of the Earth: sphere, ellipsoid, and geoid - adopted in geodesy. The height differences between the surfaces are exaggerated. A spherical model of the Earth offers a simple surface, which is mathematically easy to deal with. While the sphere is a close approximation of the true figure of the Earth and satisfactory for many purposes, to the geodesists interested in the measurement of over-continental distances and oceans, a more exact Figures are necessary [Moritz, 1989]. Directions of the rotational axis of the Earth and its angular momentum do not coincide leading to the movement of Earth's rotation axis across its surface - the polar motion - with amplitude only a few meters.

Another device used in gravimetry and geodesy to study gravitational field and its variations, is gradiometer that measures gradients of the gravity acceleration. One type of gradiometer consists of two gravimeters, which experience the same accelerations, Earth tides, latitude effects, etc. The gradient is simply the subtraction of one gravimeter response from the other, divided by the baseline distance between them. The most advanced gradiometer of this type is realized in the GRACE space mission - a joint project between the National Aeronautics and Space Administration (NASA) and the Deutsches Zentrum für Luft-und Raumfahrt (DLR) which represents twin satellites connected over two-way microwave-ranging link instrument that monitors the satellite-satellite separation and can detect changes as small as 10 microns. Accelerometers aboard each satellite help sort out the effects of non-gravitational forces, and the Global Positioning System (GPS) tracks the position of the GRACE spacecrafts above Earth. The GRACE satellites were launched on March 17, 2002 into near polar orbits with an initial altitude of about 500 km and flying nearly 220 km apart [Levi, 2003].

The European Space Agency launched another satellite to measure Earth's mass distribution: the Gravity Field and Steady-State Ocean Circulation Explorer (GOCE). Whereas GRACE gives high resolution for large-scale features, GOCE provides high-resolution measurements of smaller-scale features. GOCE's goal is to measure the height of the geoid with 1 cm resolution at a spatial resolution of 100 km. This is done with the three pairs of accelerometers aboard the single GOCE satellite that measure simultaneously the gravity gradients in three different directions.

Measuring capability of gradiometers has dimension of acceleration/distance that is expressed in terms of a special unit called Eötvös [3]. For example, GOCE's gradiometer has precision of 10^{-3} E/$\sqrt{\mathrm{H}}$ in the measurement bandwidth between 5×10^{-3} and 0.1 Hz [Bernard and Touboul, 1987; Sneeuw, 2002]. Promising work have been carried out on the design of electrostatic gradiometers for testing the equivalence principle in space [Touboul and Bernard, 2003; Touboul *et al.*, 1996] and on the development of superconducting gradiometer [Chan *et al.*, 1987; Chan and Paik, 1987; Moody *et al.*, 1986; Paik, 1981] that can be capable to detect gradients of the Earth's gravitational field with precision 10^{-7} E. Other promising efforts have beem aimed to build an atom interferometer-based gravity gradiometer [Fixler *et al.*, 2007; McGuirk *et al.*, 2002; Snadden *et al.*, 1998]. Meanwhile, the geodetic coordinates of the observer can be routinely determined now by VLBI, Doppler or laser ranging techniques with an accuracy of 1 cm. In some exceptional cases the accuracy of geodetic measurements approaches 1 mm [Kang *et al.*, 2009]. It is not difficult to see that the relative accuracy of gravimetric and geodetic measurements is comparable with the magnitude of the weak-field parameter η, and in some cases with ε. It motivates theorists to push forward a corresponding development of theoretical methods for adequate interpretation of geodetic and gravimetric measurements with the post-Newtonian corrections taken into account.

From the point of view of practical application of such advanced methods, the most interesting is related to improving definition, establishment and realization of

3) 1 Eötvös ≡ 1 E = 10^{-9} s^{-2}.

the international terrestrial reference frame (ITRF) [Altamimi and Collilieux, 2008; Lambert and Bizouard, 2002; Mueller and Rapp, 1989; Schlüter and Behrend, 2007] and relativistic time scales in the vicinity of the Earth [Brumberg and Kopeikin, 1990; Guinot and Seidelmann, 1988; Huang *et al.*, 1989; Irwin and Fukushima, 1999; Klioner *et al.*, 2009a]. Currently, the instability of the International Atomic Time (TAI), is estimated as 0.5×10^{-15} for averaging times of 20-40 days [Guinot and Felicitas Arias, 2005]. However, in the near future it will be possible to improve performance of atomic frequency standards to the stability $10^{-17} \div 10^{-18}$ [Gibble, 2007; Maleki *et al.*, 2008; Mattison, 1989]. Employment of such standards for realization of TAI will require an independent knowledge of the geoid heights with an accuracy better than 1 cm. Definition of geoid's surface with such accuracy has to inevitably include relativistic corrections.

On the other hand, a world-wide network of the extremely precise atomic clocks synchronized with respect to the Terrestrial Time (TT) [4] by means of a high-precision communication channel can be used in order to solve the inverse problem that is determination of geoid heights from a reference ellipsoid by means of comparison of the clock' rates [Bjerhammar, 1985,?; Kopejkin, 1991a]. The basic principle of this method is drastically different from the classical procedures of determination of geoid heights [Brovar *et al.*, 2001; Hofmann-Wellenhof and Moritz, 2005; Pellinen *et al.*, 1980]. Moreover, it does not require a preliminary knowledge of the distribution of the Earth's density and, therefore, can be applied directly for measuring the geoid' heights with respect to a chosen equipotential surface.

Of course, development of relativistic gravimetry and geodesy is impossible without reliable high-precision classical theory of gravimetric and geodetic measurements. The basic elements of such theory were worked out in 1849 by Stokes who derived the formula for computing the geoid [5] from surface gravity measurements. This formula became immortalized as "Stokes's integral" [Heiskanen and Moritz, 1967]. Stokes's theory faced a serious difficulty in practical applications in the middle of 20-th century as it required the gravimetric data measured on the geoid itself which is unrealistic. Crucial progress in resolving this difficulty was achieved in the theory of physical geodesy by M. S. Molodensky in his Doctor of Science thesis defended in Moscow in 1945 [Molodensky, 1945; Moritz and Yurkina, 2000]. M. S. Molodensky published most of his works in Russian journals, which were translated into English much later, so that Molodensky's theory propagated to western scientists with enormous delay. Clear and elegant introduction to the theory of Molodensky and results, which were obtained on its base up to 1980, is contained in books [Moritz, 1980; Moritz and Molodensky, 1971]. Subsequent developments can be tracked through publications [Brovar *et al.*, 2001; Hofmann-Wellenhof and Moritz, 2005; Pellinen, 1987; Sanso', 1981] and references therein. Original Molodensky's theory made an emphasis on the basic importance of the physical Earth's

4) The definition of TT is given by IAU 2000 Resolution B1.9 [Rickman, 2001] and IAU 1991 Resolution A4 [Bergeron, 1992].

5) Stokes used another word for the equipotential reference surface. The term "geoid" was introduced later by J.B. Listing in 1873 [Torge, 1991].

surface. Today, the gravimetric data are continuously determined by the geodetic satellites and the physical surface of the Earth is measured using GPS and Satellite Laser Ranging (SLR) techniques. This drastically extends various applications of the Molodensky's theory to physical geodesy and to the theory of the Earth's figure. Molodensky's theory has an elegant solution in terms of a perturbation approach given in the form of the Molodensky's series where the first term in the series takes us back to the Stokes theory.

Will [1971] was, perhaps, the first, who began theoretical investigation of relativistic effects in gravimetry. In the framework of the parameterized post-Newtonian (PPN) formalism he considered possible dependence of the Newtonian constant G on spatial variations of the Earth's gravity field that are allowed by some alternative theories of gravity. Results of the study are summarized in textbook [Will, 1993]. Bjerhammar [1985] pointed out to the possibility of practical applications of principles of general relativity for solving the problem of determination of geoid heights. The first attempt to consider the influence of the post-Newtonian gravitational field on the results of measurement of the Earth's gravity has been undertaken by Soffel *et al.* [1988a,b] who also suggested definition of relativistic geoid as well [Soffel, 1989]. However, it is necessary to note that the works of Bjerhammar [1985] and Soffel *et al.* [1988a,b] did not exhaust the theoretical part of the problem. In particular, these authors have imposed rather restricting limits on accuracy of the theory of geodetic measurements, which resulted in omitting many post-Newtonian terms from their final formulas. For example, results of work [Bjerhammar, 1985] are applicable, in fact, only in the Newtonian limit of general relativity. Our paper [Kopejkin, 1991a] extended the applicability of the approach to relativistic geodesy proposed by Soffel [1989].

The first calculation of relativistic effects in gradiometry was done by Shirokov [Shirokov, 1973] who solved the equation of deviation of geodesics and calculated the period of free oscillations of the relative distance between two test particles orbiting around a spherically-symmetric body (a Schwarzschild black hole) in close trajectories. It was found that the frequency of oscillations in the direction lying in the orbital plane and that in the direction perpendicular to the orbit differ from their Newtonian counterparts. Paper [Shirokov, 1973] also proposed the experiment for measurement of this effect in the approximation of spherically-symmetric Earth. Melkumova & Khlebnikov [Melkumova and Khlebnikov, 1990]) extended the theory of Shirokov's effect to the case of oblate and axially symmetric Earth. The next step in development of relativistic gradiometry has been done in the work by Braginsky & Polnarev [Braginskiĭ and Polnarev, 1980] and by Mashhoon & Theiss [Mashhoon and Theiss, 1982a,b] who discussed how the gravitomagnetic tidal field caused by rotation of a central massive body affects the geodesic motion of two test particles coupled by a string. Distinctive feature of Mashhoon-Theiss's approach is that they described the gravitational field of the central body by means of the exact Schwarzschild solution perturbed by the intrinsic angular momentum (spin) of the body in a linearized approximation. This allowed them to solve the problem without restricting time duration of the experiment inherently imposed by the post-Newtonian approximation. Later on, Mashhoon & Theiss extended their study to the lunar motion and discovered a new, long-periodic relativistic effect in the motion of the Moon caused by rotation of the Sun [Mashhoon and Theiss, 1986, 1991]. Soffel *et al.* [1988a,b]; Soffel [1989] and Gill *et al.* [1989] published a series of papers on the post-Newtonian gradiometry partially motivated by the line of research proposed by Mashhoon & Theiss. They were not able to reproduce the Mashhoon-Theiss results [Mashhoon and Theiss, 1986, 1991] in the post-Newtonian approximation of the gravitational field of a rotating and oblate body. Scrutiny analysis of the underlying assumptions of these works shows that Gill et al. [Gill *et al.*, 1989] discarded the second and high-order post-Newtonian terms in the solution of the equation of the geodesic deviation of two test masses. However, it is these, high-order post-Newtonian terms are critical in the derivation of the Mashhoon-Theiss effect [Mashhoon, 1985] which goes essentially beyond the limitations imposed on the dynamic evolution of the Earth-Moon system by the method of the post-Newtonian approximations.

The basic theoretical equation of relativistic gradiometry is that of the deviation of geodesics. The detailed analysis of this equation from the point of view of the measuring process and feasibility of detection of relativistic effects in satellite gradiometry has been carried out by Paik [1989, 2008] and Mashhoon *et al.* [1989] in the approximation of an isolated, spherically-symmetric and uniformly rotating Earth. Some useful theoretical results are contained also in publications [Khlebnikov, 1989; Melkumova and Khlebnikov, 1990]. Our paper [Kopejkin, 1991a] extends the analysis of Mashhoon *et al.* [1989] to the case of an N-body system.

The present chapter develops the post-Newtonian theory of gravimetric and geodetic measurements in application to the case of the real Earth with external bodies taken into account. To make discussion less independent on previous chapters the basic relativistic equations and descriptions of the barycentric, geocentric and topocentric reference frames will be recalled. One will find the post-Newtonian corrections to the Earth's gravity potential and its gradient, discuss relativistic equations of geoid and a level surface and derive relativistic equivalent of Clairaut's equation defining the flattening of the Earth as a function of the angular speed of rotation and the mass density distribution.

8.2
Basic Equations

This chapter will imply general relativity everywhere and will not consider other alternative theories of gravitation. This is because the urgent goal of relativistic geodesy is to extend the definitions of reference frames and data processing algorithms of gravimetric and gradiometric measurements from the Newtonian gravity to the most reliable relativistic extension of the theory of gravity - general relativity. Geodetic testing of other theories of gravity is a secondary goal for more distant future which requires much more theoretical and experimental work in order to clearly disentangle the effects of the Newtonian geodetic parameters (like Earth's mass, multipole moments, center of mass, Love's elasticity numbers, etc.) from the effects of the alternative theories.

The spacetime in general relativity represents a curved Riemannian manifold, endowed with a tensor field of the second rank - the metric tensor $g_{\alpha\beta}$. The metric tensor is obtained as solution of the Einstein field equations with the imposed boundary and initial conditions. The Einstein equations in arbitrary coordinate system have the following form, which should be compare with equation (4.2),

$$R_{\alpha\beta} = \frac{8\pi G}{c^2}\left(T_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}T\right) , \tag{8.1}$$

where $R_{\alpha\beta}$ is the Ricci tensor defined in (3.201); $T_{\alpha\beta}$ is the energy-momentum tensor of matter, and $T = g^{\alpha\beta}T_{\alpha\beta}$. It is useful to point out here that the Einstein equations generalize the Poisson equation of the Newtonian theory of gravity and the metric tensor extends the notion of the Newtonian gravitational potential.

In what follows, a solution of the Einstein equations in the vicinity of the Earth's center of mass will be especially important. It can be determined only after one defines the specific form of the energy-momentum tensor of the Earth's matter. Numerous gravimetric and geological measurements revealed that the Earth is not in the state of exact hydrostatic equilibrium [Hofmann-Wellenhof and Moritz, 2005; Moritz, 1989]. Therefore, the Earth's matter cannot be considered as a perfect (non-viscous) fluid and should be described by means of the energy-momentum tensor, $T_{\alpha\beta}$, which incorporates anisotropic stresses. In accordance with equation (4.8) one defines

$$c^2T^{\alpha\beta} = \left(\rho c^2 + \rho\Pi\right)u^\alpha u^\beta + \pi^{\alpha\beta} , \tag{8.2}$$

where ρ is the rest mass density of the Earth's matter, Π - the specific internal energy density of the matter, $u^\alpha = dx^\alpha/cd\tau$ - the four-velocity of a microscopic element of the matter, $\pi^{\alpha\beta}$ is the covariant form of the stress tensor of matter satisfying the following conditions $\pi^{\alpha\beta} = \pi^{\beta\alpha}$ and $\pi_{\alpha\beta}u^\alpha = \pi_{\alpha\beta}u^\alpha u^\beta = 0$. The former condition means that the stress tensor is symmetric, and the latter condition means, that in the local frame of reference comoving with the element of the Earth's matter, tensor $\pi^{\alpha\beta}$ has components which are solely spatial. Moreover, in the local frame π^{ij} is the ordinary tensor of stresses of the classical theory of elasticity [Landau and Lifshit's, 1959]. A particular form of the stress tensor for a perfect fluid is

$$\pi^{\alpha\beta} = p\left(u^\alpha u^\beta + g^{\alpha\beta}\right) , \tag{8.3}$$

where p is the isotropic pressure. The trace $\pi^\alpha_\alpha = \pi^{\alpha\beta} g_{\alpha\beta}$ of the stress tensor of the perfect fluid is equal to $3p$, since the four-velocity is subject to the condition, $u^\alpha u_\alpha = -1$, and $g^{\alpha\beta} g_{\alpha\beta} = 4$. More realistic equations for the stress tensor include the local values of other parameters like viscosity and/or elastic parameters [Landau and Lifshit's, 1959] which can be determined from geodetic measurements.

Data processing of geodetic (and any other kind of practical) observations requires from us a construction of a reference frame associated with a four-dimensional coordinate chart, x^α ($\alpha = 0, 1, 2, 3$). Usually, x^0 is taken as time coordinate, and x^i ($i = 1, 2, 3$) are considered as spatial coordinates. The distance ds in spacetime between any two infinitesimally close events with coordinates x^α and $x^\alpha + dx^\alpha$, can be expressed through the increments of the coordinates dx^α and the metric tensor $g_{\alpha\beta} = g_{\alpha\beta}(t, \boldsymbol{x})$ as

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta = g_{00}\left(dx^0\right)^2 + 2g_{0i} dx^0 dx^i + g_{ij} dx^i dx^j . \tag{8.4}$$

The metric in general relativity characterizes not only gravitational field, but the geometric properties of spacetime as well. The distance ds (or in the relativistic language - interval) is scalar, and can be spacelike ($ds^2 > 0$), timelike ($ds^2 < 0$), or null ($ds^2 = 0$) [6]. When ds is timelike, it is possible to introduce parameter τ, such that $ds^2 = -c^2 d\tau^2$, and $d\tau^2 > 0$. This parameter has physical meaning of the proper time of observer, moving along the timelike worldline and measuring time with the help of, for example, atomic clock. This is a conventional point of view on the proper time adopted, for example, in the IERS Conventions [McCarthy and Petit, 2004]. One should emphasize, however, that this is just a hypothesis adopted in general relativity for facilitating practical measurements. Strictly speaking, atomic clock operates on the basis not a geometric, but quantum mechanical principles. This means that the atomic time may differ from the proper time which can be counted only with the help of an ideal clock defined on the basis of a simplest physical phenomena reflecting geometric structure of spacetime: motion of free test particles and light signals, as stated by R. Marzke and John A. Wheeler in 1964 [Marzke and Wheeler, 1964] [7]. The ticks of the *photon clock* arise from the permanent exchange of light signals between two mirrors separated by a constant distance. It is possible to check and keep this constancy with the help of the motion of other free particles and light signals, even allowing for curved spacetime [Carrier, 1994]. The overall practical procedure of the time measurement with the help of the Marzke-Wheeler clock works properly only in the free fall and have not yet been practically implemented though the International Space Station can be an ideal place for building this clock.

6) The null interval describes propagation of wave fronts of any massless field including both gravity and light

7) See also [Carrier, 1994; Castagnino, 1968; Desloge, 1989; Feyerabend, 1994; MacDonald, 1992] for further discussion of the geometry-based clock.

The Einstein equations will be solved by means of the post-Newtonian iterations as explained in section 4.2.3. Physical quantities u^α, $g_{\alpha\beta}$ and $\sigma^{\alpha\beta}$ are expanded in powers of $1/c$ (in fact, ε and η):

$$u^0 = -1 + c^{-2} \overset{(2)}{u^0} + O\left(c^{-4}\right) , \tag{8.5}$$

$$u^i = c^{-1} u^0 v^i , \tag{8.6}$$

$$\sigma^{00} = O\left(c^{-2}\right) , \tag{8.7}$$

$$\sigma^{0i} = O\left(c^{-1}\right) , \tag{8.8}$$

$$\sigma^{ij} = t^{ij} + O\left(c^{-2}\right) , \tag{8.9}$$

$$g_{00} = -1 + c^{-2} \overset{(2)}{h}_{00} + c^{-4} \overset{(4)}{h}_{00} + O\left(c^{-5}\right) , \tag{8.10}$$

$$g_{0i} = c^{-3} \overset{(3)}{h}_{0i} + O\left(c^{-5}\right) , \tag{8.11}$$

$$g_{ij} = \delta_{ij} + c^{-2} \overset{(2)}{h}_{ij} + O\left(c^{-4}\right) , \tag{8.12}$$

These expansions are inserted into the Einstein equations (8.1), which in the first post-Newtonian approximation and in the harmonic coordinates [Fock, 1964] acquire the form

$$\Delta \overset{(2)}{h}_{00} = -8\pi G\rho^* , \tag{8.13}$$

$$\Delta \overset{(2)}{h}_{ij} = -8\pi G\rho^* \delta_{ij} , \tag{8.14}$$

$$\Delta \overset{(3)}{h}_{0i} = -16\pi G\rho^* v^i , \tag{8.15}$$

$$\Delta \overset{(4)}{h}_{00} = -8\pi G\rho^* \left(\frac{3}{2} v^2 + \Pi + \frac{t^{kk}}{\rho^*} - \frac{1}{2} \overset{(2)}{h}_{00} \right) + \frac{\partial^2 \overset{(2)}{h}_{00}}{\partial t^2} - \frac{1}{2} \Delta \left(\overset{(2)}{h}_{00} \right)^2 , \tag{8.16}$$

where $\rho^* = \rho\sqrt{-g} u^0$ is the invariant density, introduced by Fock [Fock, 1964] and $\Delta \equiv \delta^{ij}\partial_i\partial_j$ is the Laplace operator in the Euclidean space.

For development of relativistic theory of gravimetric and geodetic measurements it is necessary to construct three basic coordinate charts associated with the solar system barycentric (BRF), geocentric (GRF) and topocentric (TRF) or satellite (SRF) reference frames. From the mathematical point of view, the construction of these coordinates is done with the help of solution of the Einstein equations (8.13)-(8.16) with specific boundary conditions imposed on the components of metric tensor as explained in section 4.2.3. In BRF, the boundary conditions reflect the fact that the solar system is isolated and gravitational field is regular everywhere. Construction of GRF requires different boundary conditions. It is chosen from the reasons, that in the vicinity of the geocenter mainly the Earth's gravitational field dominates, as the gravitational field of external bodies appears only in the form of tidal terms (see section 4.5. When constructing TRF (SRF), one considers the observer as a massless point particle. Therefore, in the neighborhood of the observer, the gravitational potential contains only inertial and tidal terms and can be represented in the form of a harmonic polynomial in powers of spatial topocentric (satellite) coordinates. That is the boundary condition imposed on the metric tensor in the TRF (SRF).

Solution of the post-Newtonian field equations (8.13)-(8.16) in BRF contains the overall information about the functional structure of the metric tensor. However, in GRF and TRF (SRF) solution for the metric tensors have, at first, some unknown functions of time, which describe the gravitational field of external (with respect to Earth or to observer) bodies. For example, in TRF these functions have the physical meaning of the Earth's gravity and its gradient [Brumberg and Kopejkin, 1989b; Kopeikin, 1990; Kopejkin, 1991b]. Determination of the unknown functions in GRF and TRF (SRF) metrics as well as relativistic relationships between the coordinates is achieved by means of matching of the asymptotic expansions of the components of the metric tensor in the domain of overlapping of the different coordinate charts as explained in section 5.1.

Let us briefly repeat the idea of the asymptotic matching technique. For this sake the BRF coordinates are denoted by $x^\alpha = (x^0, x^i) = (ct, x^i)$ and the GRF coordinates by $w^\alpha = (w^0, w^i) = (cu, w^i)$, where t and u are the barycentric and geocentric coordinate times. The BRF and GRF coordinates are overlapping in the vicinity of Earth's geocenter restricted by the lunar orbit. In this domain functions w^α depend on x^α and vice versa, that is $w^\alpha = w^\alpha(t, x^i)$ and $x^\alpha = x^\alpha(u, w^i)$. It is possible to expand w^α in the vicinity of the Earth's center of mass worldline $x^i_{\scriptscriptstyle E}(t)$ in powers of the relative distance $R^i_{\scriptscriptstyle E} = x^i - x_{\scriptscriptstyle E}(t)$ with so far unknown coefficients depending on time t only. Then, one takes the BRF and GRF metrics and expand all terms describing gravitational field of external masses in powers of $R^i_{\scriptscriptstyle E}$. Matching is done with the help of a tensor law of transformation of the metric tensor

$$g_{\alpha\beta}\left(t, x^i\right) = \hat{g}_{\mu\nu}\left(u, w^i\right) \frac{\partial w^\mu}{\partial x^\alpha} \frac{\partial w^\nu}{\partial x^\beta} \ , \tag{8.17}$$

which should be considered as an equation for determination of the coordinate transformations. Indeed, the left-hand-side (l.h.s.) of equation (8.17) is known, but its right-hand-side (r.h.s.) contains in the metric $\hat{g}_{\mu\nu}$ and $\partial w^\mu / \partial x^\alpha$ the above-mentioned unknown functions, which are determined by means of equating similar terms in the l.h.s. and r.h.s of (8.17). This accomplishes matching of the BRF and GRF metric tensors. Matching of the GRF and TRF (SRF) metric tensors is done in the same manner [Brumberg and Kopejkin, 1989b; Kopeikin, 1989a; Kopejkin, 1991b] except that equation (8.17) should be replaced in this case by the following one

$$\hat{g}_{\alpha\beta}\left(u, w^i\right) = G_{\mu\nu}\left(\tau, \xi^i\right) \frac{\partial \xi^\mu}{\partial w^\alpha} \frac{\partial \xi^\nu}{\partial w^\beta} \ , \tag{8.18}$$

where $\xi^\alpha = (\xi^0, \xi^i) = (c\tau, \xi^i)$ are the TRF (SRF) coordinates and $G_{\mu\nu}\left(\tau, \xi^i\right)$ represents the TRF (SRF) metric tensor.

The outlined method is sufficiently simple, mathematically rigorous and does not contain ambiguous assumptions. It allows us to construct the self-consistent relativistic theory of astronomical reference frames and time scales in the Solar system, to derive relativistic equation of motion of the Earth's satellite in the GRF [Brumberg and Kopeikin, 1989; Brumberg and Kopejkin, 1989a] and that of the planetary center of mass in BRF with all subtle effects taken into account [Kopeikin and Vlasov, 2004], to describe the relativistic algorithm of data reduction of optical and radio astronomical observations that is directly acceptable for practice.

Alternative approach to solving the similar problems was suggested by Ashby & Bertotti [Ashby and Bertotti, 1984, 1986] and Fukushima et al. [Fukushima, 1988; Fukushima *et al.*, 1986]. Their method relies upon construction of the Fermi normal coordinates (3.214) in the vicinity of a self-gravitating body (Earth, planet, etc.). However, the given method requires introduction of some special assumptions - the absence of rotation and spherically-symmetric shape of the gravitating bodies, a special choice of the external background metric, restricting the motion of the Earth's center of mass along a geodesic worldline of the background metric, etc., which is not consistent with practice. Moreover, construction of the Fermi normal coordinates is more difficult as compared to the construction of harmonic coordinates which is achieved by straightforward solution of the Einstein equations (8.13)-(8.16).

Let us now go to the concrete description of the basic equations of the relativistic coordinate frames in gravimetry and geodesy. Description of the global BRF was given in section 4.3 to which the reader should refer for more detail. The local GRF was outlined in section 4.5 but its mathematical structure is briefly repeated below. The local topocentric frame of observer (TRF) plays a crucial role in geodetic measurements and will be introduced in section 8.4.

8.3 Geocentric Reference Frame

A metric tensor in the dynamically non-rotating GRF $w^\alpha = (cu, w^i) = (cu, \boldsymbol{w})$ is obtained by means of solution of equations (8.13)-(8.16), and has the following form

$$\begin{aligned}
\hat{g}_{00}(u,\boldsymbol{w}) &= -1 + \frac{2}{c^2}\left[U_E(u,\boldsymbol{w}) + Q_i w^i + \frac{1}{2}Q_{ij}w^i w^j + \frac{1}{6}Q_{ijk}w^i w^j w^k\right] \qquad (8.19)\\
&\quad + \frac{2}{c^4}\left[\Phi_E(u,\boldsymbol{w}) - U_E^2(u,\boldsymbol{w}) - \frac{1}{2}\frac{\partial^2\chi_E(u,\boldsymbol{w})}{\partial u^2}\right] + O\left(\frac{r^4}{c^2}\right) + O\left(\frac{r^2}{c^4}\right) + O\left(\frac{1}{c^5}\right) ,\\
\hat{g}_{0i}(u,\boldsymbol{w}) &= -\frac{4}{c^3}U_E^i(u,\boldsymbol{w}) + O\left(\frac{r^2}{c^3}\right) + O\left(\frac{1}{c^5}\right) , \qquad (8.20)\\
\hat{g}_{ij}(u,\boldsymbol{w}) &= \delta_{ij}\left\{1 + \frac{2}{c^2}\left[U_E(u,\boldsymbol{w}) + Q_k w^k + \frac{1}{2}Q_{kp}w^k w^p + \frac{1}{6}Q_{kpq}w^k w^p w^q\right]\right\} \qquad (8.21)\\
&\quad + O\left(\frac{r^4}{c^2}\right) + O\left(\frac{1}{c^4}\right) .
\end{aligned}$$

Here the internal gravitational potentials of the Earth are defined by equations

$$U_E(u,\boldsymbol{w}) = G\int_{V_E}\frac{\rho^*(u,\boldsymbol{w}')}{|\boldsymbol{w}-\boldsymbol{w}'|}d^3w' , \qquad (8.22)$$

$$U_E^i(u,\boldsymbol{w}) = G\int_{V_E}\frac{\rho^*(u,\boldsymbol{w}')v^i(u,\boldsymbol{w}')}{|\boldsymbol{w}-\boldsymbol{w}'|}d^3w' , \qquad (8.23)$$

$$\chi_E(u,\boldsymbol{w}) = -G\int_{V_E}\rho^*(u,\boldsymbol{w}')|\boldsymbol{w}-\boldsymbol{w}'|d^3w' , \qquad (8.24)$$

$$\Phi_E(u,\boldsymbol{w}) = G\int_{V_E}\frac{\rho^*(u,\boldsymbol{w}')\left[\frac{3}{2}v^2(u,\boldsymbol{w}') - U_E(u,\boldsymbol{w}') + \Pi(u,\boldsymbol{w}')\right] + t^{kk}(u,\boldsymbol{w}')}{|\boldsymbol{w}-\boldsymbol{w}'|}d^3w' \qquad (8.25)$$

The inertial gravity force due to the coupling between the Earth's oblateness and the tidal gravitational octupole Q_{ijk} is given in equation (6.20) which is approximated by

$$Q_i = -\frac{1}{2}M_E^{-1}Q_{ijk}I_E^{jk} + O\left(J_3\frac{L_E^3}{R_E^3}\right) + O\left(\frac{1}{c^2}\right) , \qquad (8.26)$$

where the baryon's rest mass of the Earth

$$M_E = \int_{V_E}\rho^*(u,\boldsymbol{w})d^3w , \qquad (8.27)$$

and the Earth's second moment of inertia

$$I_E^{ij} = \int_{V_E}\rho^*(u,\boldsymbol{w})w^i w^j d^3w . \qquad (8.28)$$

Equation (8.26) for the inertial force should be compared with more complete expression for this force given in equation (6.20). The inertial force given by equation

(8.26) is known in geodynamics as an indirect J_2 effect [Gruber *et al.*, 2009; Newhall *et al.*, 1983].

The tidal gravitational quadrupole and octupole moments are given by equation (5.59)

$$Q_{ij} = \bar{U}_{,ij}(t, \boldsymbol{x}_E) + O\left(\frac{1}{c^2}\right), \tag{8.29}$$

$$Q_{ijk} = \bar{U}_{,ijk}(t, \boldsymbol{x}_E) + O\left(\frac{1}{c^2}\right), \tag{8.30}$$

where the external gravitational potential

$$\bar{U}(t, \boldsymbol{x}_E) = \sum_{A \neq E} U_A(t, \boldsymbol{x}), \tag{8.31}$$

$$U_A(t, \boldsymbol{x}) = G \int_{V_A} \frac{\rho^*(t, \boldsymbol{x}')}{|\boldsymbol{x} - \boldsymbol{x}'|} d^3x', \tag{8.32}$$

includes all bodies of the solar system other than Earth. One has omitted relativistic terms in formulas (8.29), (8.30) since these terms have magnitudes, which are negligibly small for the ground-based gravimetric and geodetic measurements. However, the relativistic influence of tidal fields may be important for high-resolution determination of the fine-multipole structure of the gravitational field of the Earth in future satellite missions and for testing general relativity with lunar laser ranging [Kopeikin *et al.*, 2008; Müller *et al.*, 2008b].

Other notations that were used in above equations, are $r = |\boldsymbol{w}| = \sqrt{\delta_{ij} w^i w^j}$, $v^i = dw^i/du$ - velocity of the Earth's matter with respect to GRF; L_E - the mean radius of the Earth; $R_E = |\boldsymbol{x} - \boldsymbol{x}_E|$; $x^\alpha = (ct, x^i) = (ct, \boldsymbol{x})$ - the barycentric (BRF) coordinates of the Solar system, connected with the GRF coordinates by means of the post-Newtonian transformations given in section (5.2.6); $\boldsymbol{x}_E = x^i_E(t)$ - the BRF coordinates of the Earth's center of mass; V_E and V_A denotes the matter volumes of the Earth and another gravitating body A respectively; summation in the formula (8.31) is carried out over all gravitating bodies of the Solar system other than Earth; J_2 and J_3 - the oblateness factors of the Earth of the second and third order respectively.

Metric potentials U_E, U^i_E, Φ_E, χ_E characterize gravitational field of the Earth only. They can be expanded outside the Earth in the internal multipole harmonics with respect to the powers of L_E/r:

$$U_E = \frac{GM_E}{r} + G \sum_{l=2} \frac{(-1)^l}{l!} J_E^{i_1 i_2 \dots i_l} \frac{\partial^l}{\partial w^{i_1} \partial w^{i_2} \dots \partial w^{i_l}} \left(\frac{1}{r}\right), \tag{8.33}$$

$$U^i_E = \frac{G(\boldsymbol{S}_E \times \boldsymbol{w})^i}{2r^3} + \frac{G\dot{I}^{ij}_E w^j}{2r^3} + O\left(\frac{L^3_E}{r^3}\omega_E\right), \tag{8.34}$$

$$\Phi_E = \frac{G}{r} \int_{V_E} \left(\frac{1}{2}\rho^* v^2 - \frac{1}{2}\rho^* U_E + \rho^* \Pi\right) d^3w + \frac{G\ddot{I}^{kk}_E}{2r} + O\left(\frac{L^4_E}{r^2}\omega^2_E\right), \tag{8.35}$$

$$\chi_E = -GM_E r - \frac{GI^{kk}}{3r} - \frac{3}{2}\frac{J^{ij}_E w^i w^j}{r^3} + O\left(\frac{L^3_E}{r^2}\right). \tag{8.36}$$

Here the internal multipoles of the Earth are

$$J_E^{i_1 i_2 \dots i_l} = \int_{V_E} \rho^*(u, \boldsymbol{w}) w^{<i_1} w^{i_2} \dots w^{i_l>} d^3 w \,, \tag{8.37}$$

$$S_E^{\,i} = \int_{V_E} \rho^*(u, \boldsymbol{w}) (\boldsymbol{w} \times \boldsymbol{v})^i d^3 w \,, \tag{8.38}$$

where $i_1, i_2, \dots i_l$ are the spatial indices of the GRF coordinates; the angular brackets around the indices denote STF tensor; ω_E is the module of the angular velocity of the Earth's rotation; $J_E^{i_1 i_2 \dots i_l}$ is the Newtonian mass multipole moments of the Earth's gravitational field of the order l; $S_E^{\,i}$ is the intrinsic angular momentum (spin) of the Earth; each dot over any function means the total derivative with respect to the coordinate time u of the GRF. In the expansion of potential U_E the infinite number of the multipole moments is taken into account. This is because U_E appears in the Newtonian terms and, therefore, has to be calculated with exhaustive completeness. In expansion of the other potentials, which appear always only in the post-Newtonian terms, it is quite enough to take into account several first terms.

The mass multipole moments $J_E^{i_1 i_2 \dots i_l}$ describe the dependence on the Earth's internal mass distribution. They are inseparably linked with the geopotential coefficients C_{lm} and S_{lm} $(m = 0, 1, 2, \dots, l)$ [Soffel *et al.*, 2003] in the sense that there exists one-to-one mapping between the geopotential coefficients and the STF multipoles $J_E^{i_1 i_2 \dots i_l}$. In a simplified case of a mass distribution that is symmetric with respect to the axis of rotation one can express the mass STF multipoles as follows

$$J_E^{i_1 i_2 \dots i_l} = M_E R_E^l J_l s^{<i_1} s^{i_2} \dots s^{i_l>} \,, \tag{8.39}$$

where s^i is the unit vector directed along the axis of rotation, $J_l = -C_{l0}$ are the zonal coefficients characterizing the amount of oblateness of each zonal harmonic in the expansion of geopotential U_E, and the angular brackets indicate that the symmetric and trace-free combination of the unit vectors must be taken. In conclusion of this section one notices that the external potentials U_A $(A \neq E)$ can be decomposed to the multipolar series in the same way as it was done for geopotential [Kopejkin, 1988a].

8.4
Topocentric Reference Frame

Topocentric reference frame (TRF) is more familiar to relativists under the name of the proper reference frame of observer [Misner *et al.*, 1973]. From practical point of view TRF is the frame that is used for doing local measurements in spacetime [Brumberg and Kopejkin, 1989b; Klioner, 2004; Kopejkin, 1991a,b; Misner *et al.*, 1973]. TRF can be introduced in many different ways. For example, the Fermi normal coordinates [Misner *et al.*, 1973; Schouten, 1954] are often used in order to describe the functional structure of the metric tensor in the vicinity of the worldline of accelerated and rotating observer as demonstrated in papers [Li and Ni, 1979; Marzlin, 1994; Ni and Zimmermann, 1978b]. However, according to the principle of general covariance of general relativity any coordinates can be used in order to describe the local measurements. One has used the harmonic coordinates as their construction is more straightforward and requires solution of vacuum Einstein equations in a close vicinity of the worldline of observer. The solution consists of the set of harmonic polynomials with the coefficients which describe the inertial and tidal forces existing in the vicinity of observer. Detailed mathematical procedure is outlined in papers [Suen, 1986; Zhang, 1986].

Let us introduce the TRF coordinates $\xi^\alpha = (c\tau, \xi^i) = (c\tau, \boldsymbol{\xi})$ that cover the immediate vicinity of a massless observer, which is always located at the origin of the TRF. The TRF metric tensor $G_{\alpha\beta}(\tau, \boldsymbol{\xi})$ is obtained from equations (8.13)-(8.16), where functions ρ^*, Π, and t^{ij} are now set equal to zero identically. One stipulates that the TRF is dynamically non-rotating which means that G_{0i} component of the metric tensor does not have terms being proportional to the angular velocity of rotation. Then, with sufficient accuracy one has for the TRF metric tensor the following expressions [Kopejkin, 1991a,b]

$$G_{00}(\tau, \boldsymbol{\xi}) = -1 + \frac{2}{c^2}\left(E_i\xi^i + \frac{1}{2}E_{ij}\xi^i\xi^j + \frac{1}{6}E_{ijk}\xi^i\xi^j\xi^k\right) + O\left(\frac{\xi^4}{c^2}\right) + O\left(\frac{\xi^2}{c^4}\right) + O\left(\frac{1}{c^5}\right), \tag{8.40}$$

$$G_{0i}(\tau, \boldsymbol{\xi}) = O\left(\frac{\xi^2}{c^3}\right) + O\left(\frac{1}{c^5}\right), \tag{8.41}$$

$$G_{ij}(\tau, \boldsymbol{\xi}) = \delta_{ij}\left[1 + \frac{2}{c^2}\left(E_k\xi^k + \frac{1}{2}E_{kp}\xi^k\xi^p + \frac{1}{6}E_{kpq}\xi^k\xi^p\xi^q\right)\right] + O\left(\frac{\xi^4}{c^3}\right) + O\left(\frac{1}{c^4}\right) \tag{8.42}$$

where the quantities E_i , E_{ij}, E_{ijk} depend on the coordinate time τ only, and are STF tensors taken at the origin of TRF as follows from the solution of the homogeneous Laplace equations. Their precise form will be given later, when discussing applications for gravimetry and gradiometry.

The topocentric coordinate time τ at the origin of TRF represents a proper time of the observer that can be measured with observer's clock. The spatial coordinates of the TRF, ξ^i, in the infinitesimally close neighborhood of the observer measure the proper distance and directions and can be interpreted as the standard Euclidean coordinates of the Newtonian theory. Rotating TRF can be easily introduced in a standard way by making a coordinate transformation to a rigidly rotating frame as shown below in equation (8.58). The rotating TRF coordinates obtained in this way are not harmonic as the harmonic gauge condition does not admit rigidly rotating frames of reference [Brumberg and Kopejkin, 1989b; Kopejkin, 1991b].

Physical meaning of quantities E_i, E_{ij}, E_{ijk} is established by means of consideration of the topocentric equations of motion of a test particle with negligibly small mass m moving with respect to observer. Let F^α be the external non-gravitational force exerted on the particle. Then, the particle moves along an accelerated worldline described by the perturbed equation of geodesic [Landau and Lifshitz, 1975; Misner *et al.*, 1973]

$$\frac{d^2\xi^\alpha}{d\lambda^2} + c^2\Gamma^\alpha_{00}\left(\frac{d\tau}{d\lambda}\right)^2 + 2c\Gamma^\alpha_{0i}\frac{d\tau}{d\lambda}\frac{d\xi^i}{d\lambda} + \Gamma^\alpha_{ij}\frac{d\xi^i}{d\lambda}\frac{d\xi^j}{d\lambda} = \frac{F^\alpha}{m}, \tag{8.43}$$

where λ is the proper time of the particle (affine parameter), $\Gamma^{\alpha}_{\beta\gamma}$ - the Christoffel symbols, and the four-force satisfies the condition of orthogonality $g_{\alpha\beta}F^{\alpha}(d\xi^{\beta}/d\lambda) = 0$, which means that in the proper reference frame of the test particle the four-force is a measurable observable quantity having only spatial components. In TRF the four-force has all four components different from zero because TRF is not the proper reference frame of the test particle under consideration. However, because the four-force is orthogonal to four-velocity of the test particle, its time component is not independent and can be always expressed in terms of its spatial components.

Calculating the Christoffel symbols and transforming the proper time λ to topocentric coordinate time τ with the help of the definition of proper time,

$$-d\lambda^2 = G_{00}d\tau^2 + \frac{2}{c}G_{0i}d\tau d\xi^i + \frac{1}{c^2}G_{ij}d\xi^i d\xi^j \,, \tag{8.44}$$

and the metric (8.40)-(8.42), one obtains the spatial components of the equations of motion of the test particle

$$\frac{d^2\xi^i}{d\tau^2} = \frac{F^i}{m} + E_i + E_{ij}\xi^j + \frac{1}{2}E_{ijk}\xi^j\xi^k - \frac{4}{c^2}E_iE_k\xi^k + O\left(\xi^3\right) + O\left(\frac{\xi^2}{c^2}\right) . \tag{8.45}$$

This equation is nothing else but the second Newton's law written down in a non-inertial coordinate system because of the presence of the inertial acceleration E_i in the right side of it. One can see that the Coriolis and centrifugal accelerations are absent in (8.45). It means that the TRF is dynamically non-rotating. It follows from equation (8.45) that the acceleration E_i taken with the sign minus (that is $-E_i$) represents acceleration of the origin of the TRF caused by its non-inertial motion. Indeed, let's assume that the test particle is passing through the origin of TRF and is in the state of a free fall, which means $F^i = 0$ and $\xi^i = 0$. Equation (8.45) tells us that in this case $d^2\xi^i/d\tau^2 = E_i$. Since the freely-falling test particle moves along the geodesic worldline and can not experience any physical acceleration by definition, one concludes that the TRF is a non-inertial frame that moves with acceleration $-E_i$, which creates a "fictitious" inertial acceleration E_i of the particle with an opposite sign.

One can measure E_i by keeping the test particle at the origin of TRF. For this sake, the experimentalist must apply a certain force $F^i \neq 0$ to the particle to ensure fulfillment of the following conditions: $\xi^i = d\xi^i/d\tau = d^2\xi^i/d\tau^2 = 0$, which mean that the particle is always at the origin of the TRF. Then, equation (8.45) tells us that $E^i = F^i/m$. Thus, measuring the force F^i with the help of an accelerometer (gravimeter), one obtains the exhaustive quantitative information about E_i. In section 8.6 it will be shown that for the ground-based observer, E_i represents the acceleration of Earth's gravity with the post-Newtonian corrections taken into account. Topocentric frame attached to a spacecraft orbiting the Earth is called satellite reference frame (SRF). Mathematical description of SRF is the same as TRF but the inertial acceleration $E_i = 0$ if the satellite is equipped with a drag-free control system. This system comprises a proof mass which is disconnected from the satellite, a sensor to measure any displacements between the proof mass and the satellite, as actuators a number of low-level (ideally proportional) thrusters, and the associated electronics and software (drag-free control algorithms) [Debra, 2003]. Protected inside the satellite, the proof mass follows a purely gravitational orbit. The position of the satellite with respect to the proof mass is measured and fed back to thrusters, which force the satellite to follow the proof mass.

The E_{ij} and E_{ijk} describe the tidal forces which can be expressed in terms of the Riemannian curvature of spacetime and its derivatives. These forces can be measured with the help of a gradiometer, located at the origin of the TRF (SRF). Functional structure of E_i, E_{ij}, E_{ijk} as well as the relationship between the GRF and TRF coordinates are determined by means of matching of the asymptotic expansions of the metric tensors given by equations (8.19)-(8.21) and (8.40)-(8.42) as it was described in section 8.3 and 8.4. These formulas constitute the basis of the relativistic gravimetry, gradiometry, and geodesy. Results of the matching procedure are described in the following sections.

8.5
Relationship between the Geocentric and Topocentric Frames

Let us stipulate that observer is permanently located at the origin of TRF and moves in spacetime along a worldline described in GRF by spatial coordinates $w^i_T = w^i_T(u)$ depending on time u. One is looking for a post-Newtonian transformation between GRF and TRF coordinates in the following form [Brumberg and Kopejkin, 1989b; Kopejkin, 1991a,b]

$$\tau = u - \frac{1}{c^2}\left[V(u) + v^k_T r^k_T\right] - \frac{S(u)}{c^4} + O\left(\frac{r_T}{c^4}\right) + O\left(\frac{1}{c^5}\right), \tag{8.46}$$

$$\xi^i = r^i_T + \frac{1}{c^2}\left[\frac{1}{2}v^i_T v^k_T + R^{ik}(u) + Z^{ik}(u)\right] r^k_T + O\left(\frac{r^2_T}{c^2}\right) + O\left(\frac{1}{c^4}\right), \tag{8.47}$$

where $r^i_T = w^i - w^i_T(u)$, and $v^i_T = dw^i_T/du$ is velocity of the observer with respect to the dynamically non-rotating GRF [8].

Functions $V(u)$, $S(u)$, and matrices $R^{ik}(u) = -R^{ki}(u)$ and $Z^{ik}(u) = Z^{ki}(u)$ depend only on the geocentric time u and are determined from the matching procedure. Inserting formulas (8.19)-(8.21), (8.40)-(8.42), (8.46), (8.47) to equation (8.18) and matching the GRF and TRF coordinates (that is equating similar terms in both sides of equation (8.18)), one concludes that these functions satisfy the following equa-

8) This velocity is associated with either the diurnal motion of the ground-based observer or with the velocity of a satellite orbiting the Earth.

tions

$$\frac{dV}{du} = \frac{1}{2}v_T^2 + U_E(\boldsymbol{w}_T) + Q_i w_T^i + \frac{1}{2}Q_{ij}w_T^i w_T^j + \frac{1}{6}Q_{ijk}w_T^i w_T^j w_T^k + O\left(w_T^4\right) , \tag{8.48}$$

$$\frac{dS}{du} = \frac{1}{8}v_T^4 + \frac{3}{2}v_T^2 U_E(\boldsymbol{w}_T) - 4v_T^i U_E^i(\boldsymbol{w}_T) - \frac{1}{2}U_E^2(\boldsymbol{w}_T) + \Phi_E(\boldsymbol{w}_T) - \frac{1}{2}\frac{\partial^2\chi_E(\boldsymbol{w}_T)}{\partial u^2} + O\left(w_T\right) , \tag{8.49}$$

$$\frac{dR^{ik}}{du} = 3v_T^{[i}U_E^{,k]}(\boldsymbol{w}_T) - 4U_E^{[i,k]}(\boldsymbol{w}_T) + v_T^{[i}E^{k]} + O(w_T) , \tag{8.50}$$

$$Z^{ik} = \delta^{ik}\left[U_E(\boldsymbol{w}_T) + Q_p w_T^p + \frac{1}{2}Q_{pq}w_T^p w_T^q + \frac{1}{6}Q_{jpq}w_T^j w_T^p w_T^q\right] + O\left(w_T^4\right) . \tag{8.51}$$

Formulas (8.46), (8.48) and (8.49) give the relationship between the geocentric and topocentric coordinate time scales. In case of $r_T^i = 0$, they give rise to relationship between the proper time of observer and the geocentric time u. The symmetric matrix Z^{ik} describes the isotropic gravitational contraction of the grid of the TRR spatial coordinates with respect to that of the GRF spatial coordinates.

The anti-symmetric matrix R^{ik} in equations (8.47), (8.50) describes a post-Newtonian precession of the spatial axes of the TRF coordinates with respect to the spatial GRF coordinates. Spin of a torque-free gyroscope that is placed to the origin of TRF, will undergo this precession in accordance with equation

$$\frac{dS^i}{d\tau} = \left(\boldsymbol{\Omega}_{pN} \times \boldsymbol{S}\right)^i , \tag{8.52}$$

where $\boldsymbol{\Omega}_{pN} = (\Omega_{pN}^i) = \frac{1}{2}\epsilon^{ijk}dR^{jk}/d\tau$ is the angular velocity of the post-Newtonian precession, and one has replaced the time derivative with respect to time u by the time derivative with respect to the proper time of observer τ as the difference between them is of the post-Newtonian order of magnitude and can be neglected in equation (8.52). Using the ordinary notation of vector analysis for gradient, $(\partial_i) = \boldsymbol{\nabla}$, equation for the angular velocity of the relativistic precession can be recast to the form

$$\Omega_{pN}^i = \frac{3}{2}\left(\boldsymbol{v}_T \times \boldsymbol{\nabla}U_E\right)^i + 2\left(\boldsymbol{\nabla} \times \boldsymbol{U}_E\right)^i + \frac{1}{2}\left(\boldsymbol{v}_T \times \boldsymbol{E}\right)^i , \tag{8.53}$$

where vector-potential $\boldsymbol{U}_E = (U_E^i)$ which is explicitly given in equation (8.34). Standard textbooks on general relativity [Misner *et al.*, 1973; Weinberg, 1972] derive the relativistic precession (8.53) of gyroscope's spin by making use of the Fermi-Walker transport along accelerated worldline of the gyroscope's center of mass. Our derivation of equation (8.53) originates from the post-Newtonian precession of the local TRF derived by the asymptotic matching technique [9].

The first term in the right side of equation (8.50) is the de Sitter (geodetic) precession caused by the motion of observer with respect to GRF. The second term in the right side of equation (8.50) is the Lense-Thirring (gravitomagnetic) precession caused by Earth's rotation which drags the local inertial frames [Ciufolini and

9) Notice that the notation for the local acceleration measured by the observer differs by sign from that used in the textbooks [Misner *et al.*, 1973; Will, 1993]. Textbook [Weinberg, 1972] ignores the acceleration E^i and uses definitions, $\phi =: -U_E$ and $\zeta^i =: -4U_E^i$.

Wheeler, 1995; Misner *et al.*, 1973]. The measuring of the geodetic and gravitomagnetic precessions has been achieved in the Gravity Probe-B (GPB) Relativity Gyroscope Experiment [Everitt *et al.*, 2008, 2009b; Turneaure *et al.*, 2003]. An independent measurement of the gravitomagnetic precession of the orbit of LAGEOS satellite was achieved by Ignazio Ciufolini with collaborators [Ciufolini, 2007; Ciufolini and Pavlis, 2004; Ciufolini *et al.*, 2006] and independently confirmed by John Ries [Ries, 2009]. The third term in the left side of equation (8.50) represents the Thomas precession [Misner *et al.*, 1973], which is not equal to zero only for the observers which experience a local acceleration $E_i \neq 0$. It means that the Thomas precession is absent, for example, for a GPB gyroscope onboard of satellite which is in a free-fall flight. For the ground-based observer, the Thomas precession couples with the geodetic precession, since in the TRF $E_i \simeq \partial_i U_{\scriptscriptstyle E}(\boldsymbol{w}_{\scriptscriptstyle T})$ (see equations (8.54), (8.55). There exists also tidal contributions to the geodetic and gravitomagnetic precession caused by the presence of external with respect to the Earth bodies [Kopeikin, 1989a; Kopejkin, 1991b]. The magnitude of the tidal terms to the post-Newtonian precession is very small amounting only to $\leq 10^{-3}$ arcsecond per 100 years. Therefore, measuring of the tidal corrections to equation (8.50) is problematic.

8.6 Post-Newtonian Gravimetry

Gravimetry is the science of measuring the strength and direction of Earth's gravitational field at different locations. Therefore, the basic formula of gravimetry is an expression for the acceleration of Earth's gravity. According to the physical nature of the metric tensor the force of gravity is given in the topocentric reference frame by a (co)vector E_i entering G_{00} component of the metric tensor as shown in equation (8.40). Exact analytic expression for this vector is derived from the matching of asymptotic expansions of GRF and TRF metric tensors according to equation (8.18). It yields

$$E_i = g_i^N + \frac{1}{c^2} g_i^{PN} + O\left(\frac{1}{c^4}\right) , \tag{8.54}$$

where

$$g_i^N = V_{E,i}(\boldsymbol{w}_T) - a_T^i + Q_i + Q_{ij} w_T^j + \frac{1}{2} Q_{ijk} w_T^j w_T^k + O\left(w_T^3\right) , \tag{8.55}$$

$$\begin{aligned} g_i^{PN} = {} & 4U^i_{E,u}(\boldsymbol{w}_T) + 4v_T^k U^i_{E,k}(\boldsymbol{w}_T) - 4v_T^k U^k_{E,i}(\boldsymbol{w}_T) \\ & -3v_T^i U_{E,u}(\boldsymbol{w}_T) - 3a_T^i U_E(\boldsymbol{w}_T) - \frac{7}{2} v_T^i v_T^k U_{E,k}(\boldsymbol{w}_T) \\ & +2v_T^2 U_{E,i}(\boldsymbol{w}_T) - \frac{1}{2} v_T^i \left(a_T^k v_T^k\right) - v_T^2 a_T^i + g_j^N R^{ij} + O(w_T) , \end{aligned} \tag{8.56}$$

and one has introduced a generalized gravitational potential

$$V_E(\boldsymbol{w}) = U_E(\boldsymbol{w}) + \frac{1}{c^2}\left[\Phi_E(\boldsymbol{w}) - \frac{1}{2} U_E^2(\boldsymbol{w}) - \frac{1}{2}\chi_{E,uu}(\boldsymbol{w})\right] , \tag{8.57}$$

consisting of the Newtonian potential U_E and a post-Newtonian contribution to it (terms proportional to $1/c^2$). Quantities $\boldsymbol{w}_T$, $\boldsymbol{v}_T = d\boldsymbol{w}_T/du$, $\boldsymbol{a}_T = d\boldsymbol{v}_T/du$ in the above equations are respectively the GRF coordinates, velocity, and acceleration of the observer measuring the gravity field with a gravimeter that is placed at the origin of TRF; R^{ij} is a matrix of the relativistic precession defined by formula (8.50).

The coordinates w_T^i are connected with the geodetic coordinates y_T^i of the International Terrestrial Reference Frame (ITRF) by the orthogonal matrix of a continuous rotational transformation [McCarthy and Petit, 2004]:

$$y_T^i = S^{ij} N^{jk} P^{kq} w_T^q , \tag{8.58}$$

where $S^{ij} = S^{ij}(u)$, $N^{ij} = N^{ij}(u)$, $P^{ij} = P^{ij}(u)$ are matrices of the diurnal rotation, nutation and precession of the Earth, which are functions of time u and determined by the IERS Standards [McCarthy and Petit, 2004]. One notices that the current convention is to include the post-Newtonian precession to the matrix of precession P^{ij} taking into account that the spatial axes of the geocentric frame are fixed with respect to the International Celestial Reference Frame based on radio quasars [Fey *et al.*, 2004; Souchay and R., 2010]. Differentiation of both sides of equation (8.58) with respect to the geocentric coordinate time u, one obtains the explicit expressions for the geocentric velocity and acceleration of the observer:

$$v^i_T = (\boldsymbol{\Omega} \times \boldsymbol{w}_T)^i + v^i_{TT} \,, \tag{8.59}$$

$$a^i_T = (\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \boldsymbol{w}_T))^i + 2(\boldsymbol{\Omega} \times \boldsymbol{v}_{TT})^i + (\dot{\boldsymbol{\Omega}} \times \boldsymbol{w}_T)^i + a^i_{TT} \,, \tag{8.60}$$

where the first term in the right side of equation (8.59) represents velocity of observer due to the diurnal rotation of the Earth, $\boldsymbol{\Omega} = (\Omega^i)$ is the instantaneous angular velocity of the Earth's rotation with respect to the kinematically non-rotating geocentric frame (GRF); v^i_{TT} and $a^i_{TT} = dv^i_{TT}/du$ are velocity and acceleration of the observer with respect to the terrestrial reference frame GRF$^+$, which rotates with respect to GRF with the angular velocity Ω^i. When observer is at rest with respect to the Earth's crust, velocity v^i_{TT} may be different from zero only because of geophysical factors (crustal motion).

The first term in the right side of equation (8.60) is the centrifugal acceleration of observer due to the diurnal rotation of the Earth. It decreases the force of gravity for observers located at any point on Earth but poles. The second term in the right side of equation (8.60) is the Coriolis acceleration which is different from zero only for observers moving with respect to the GRF$^+$ with velocity v^i_{TT}. Third term in right side of equation (8.60) is the acceleration due to the angular acceleration of Earth's rotation (time variations of the rotational velocity of the Earth). Acceleration a^i_{TT} due to the regular crustal motion is negligibly small [10)]. If the observer is doing gravimetric measurement from a moving platform (ship, aircraft), the contribution of v^i_{TT} and a^i_{TT} in the readings of measuring gravimetric devices is rather large and should be taken into account. This correction is known as the Eötvös effect [Dehlinger, 1978].

One has pointed out already that the quantity g^N_i extends the Newtonian definition of the Earth's gravity by taking into account the gradient of the other post-Newtonian potentials, Φ_E and χ_E, which are included to the first term in the right side of equation (8.55) by means of definition (8.57). In particular, the presence of the term Φ_E in this definition indicates that besides the matter's density ρ^* the source of the gravity force is the density of the internal self-gravitational energy of the Earth as well as the kinetic and thermodynamics energy, and stresses (pressure) of matter. Their contribution has a maximal amplitude of the order of a few μgals but they can not be separated from the Newtonian gravity gradient in the gravimetric measurement.

The term a^i_T in the right side of equation (8.55) is the centrifugal acceleration (8.60) decreasing the gravity force depending of the latitude of observer. The third term Q_i in the right side of equation (8.55) describes the acceleration of the geoid's center-of-mass worldline from that of a spherically-symmetric Earth. The main contribution to Q_i is due to the interaction between the quadrupole moment J^{ik} of the Earth and the octupole tidal gravitational field Q_{ijk} of external bodies (mostly the Moon) as shown in equation (8.26). It is not difficult to evaluate that Q_i changes periodically

10) The case of earthquakes, when the acceleration a^i_{TT} can be rather large for a short interval of time, is an important seismological parameter to measure but it has a loose connection to the present book and is not analyzed.

and has an amplitude of the order of 3 nanogals. The present-time super-conducting gravimeters are not yet sufficiently precise for a direct measurement of the inertial acceleration Q_i on the ground. However, it must be taken into account in analysis of observations of geodetic satellites like LAGEOS, GRACE [Levi, 2003], GOCI [Albertella *et al.*, 2002]. The term Q_i must be taken into account in the barycentric equations of motion of the Earth and Moon for constructing high-precision ephemeris of these bodies as shown in paper by Newhall *et al.* [1983]). This is because the acceleration Q_i is comparable in magnitude with the Earth's figure effects [Newhall *et al.*, 1983] and leads to noticeable periodic variation of the Earth-Moon distance. This variation is approximately equal to $|Q_i|\times$ (the mean motion of the Moon)$^{-2} \simeq 10$ meters - much larger than the accuracy of the lunar laser ranging (LLR) measurements to retro-reflectors on the Moon [Müller *et al.*, 2008b].

The forth and fifth terms in the right side of equation (8.55) describe the direct Newtonian influence of the tidal gravitational field of external bodies on the magnitude of the Earth's gravity acceleration. The indirect influence of this field appears in the form of the tidal deformation of the Earth's figure leading to tidal variation of the Earth's gravitational potential U_E [Melchior, 1983; Montenbruck and Gill, 2000].

The direct post-Newtonian correction g_i^{PN} to the Newtonian gravity depends on the gravitomagnetic potential U_E^i, and velocity and acceleration of observer coupled with the Newtonian potential U_E. These terms reach maximal amplitude of the order of a few nanogals on the Earth's equator and reduce to zero at the poles.

The relativistic term $g_k^N R^{ik}$ in the right side of equation (8.56) appears because of the relativistic precession of TRF spatial axes with respect to GRF, and manifests itself as a slowly varying-in-direction force in the equation of motion (8.45) of the test particle. This force will change, for example, orientation of the swinging plane of Foucault's pendulum with respect to the dynamically non-rotating spatial axes of GRF. The indicated effect can be used, in principle, for measuring the gravitomagnetic potential U_E^i of the Earth as it was suggested by Braginsky et al. [Braginskii *et al.*, 1984].

The post-Newtonian corrections to the Newtonian tidal variations of gravity were not included to formula (8.56) since they are negligibly small - approximately by a factor of 10^{-9} smaller as compared to the classic tidal variation of gravity $\sim Q_{ij}w_T^j$. Explicit form of the tidal post-Newtonian corrections to E_i have been derived in [Kopejkin, 1991b].

Finally, one considers the case when observer is located at the satellite orbiting the Earth. Let us denote coordinates of the satellite $(w^i) = \boldsymbol{w}$, and its velocity $(v^i) = \boldsymbol{v} = d\boldsymbol{w}/du$. In this case the orbital acceleration of the satellite is $a^i = dv^i/du$ and the acceleration E_i is due to the non-gravitational forces (air drag, Yarkovsky's effect [Beekman, 2006], solar pressure, etc.) preventing satellite to move along a geodesic worldline [11]. In this case, equations (8.54)-(8.57) become the equation of motion of

11) If the satellite is equipped with the drag-free system control the non-gravitational acceleration $E_i = 0$.

the satellite

$$
\begin{aligned}
a^i &= V_{E,i}(\boldsymbol{w}) - E_i + Q_i + Q_{ij}w^j + \frac{1}{2}Q_{ijk}w^j w^k \\
&+ \frac{1}{c^2}\Big[4U^i_{E,u}(\boldsymbol{w}) + 4v^k U^i_{E,k}(\boldsymbol{w}) - 4v^k U^k_{E,i}(\boldsymbol{w}) \\
&\quad -3v^i U_{E,u}(\boldsymbol{w}) - 3U_E(\boldsymbol{w})U_{E,i}(\boldsymbol{w}) - 4v^i v^k U_{E,k}(\boldsymbol{w}) \\
&\quad +v^2 U_{E,i}(\boldsymbol{w})\Big] + O\left(w^3\right) + O\left(\frac{w}{c^2}\right) ,
\end{aligned} \tag{8.61}
$$

where the inertial acceleration Q_i of the Earth's geocenter causes the indirect J_2 effect in the motion of satellites [Gruber *et al.*, 2009], and terms with Q_{ij}, Q_{ijk}, etc., are direct tidal accelerations from external bodies (Moon, Sun, planets) [Montenbruck and Gill, 2000]. The post-Newtonian perturbations of the satellite's orbit have been analyzed in [Brumberg, 1991; Brumberg and Kopeikin, 1989; Brumberg and Kopejkin, 1989a; Martin *et al.*, 1985; Soffel, 1989]. Among the post-Newtonian perturbations, the most remarkable is the gravitomagnetic acceleration that is formed by terms with vector-potential U^i_E in the second line of equation (8.61)

$$
a^i_{\rm GM} = \frac{8}{c^2}v^k U_E^{[i,k]}(\boldsymbol{w}) , \tag{8.62}
$$

where one has taken into account that in the approximation of the rigidly-rotating Earth the partial time derivative from vector-potential U^i_E is zero. The gravitomagnetic acceleration can be recast to a more usual, double vector product form

$$
a^i_{\rm GM} = -\frac{4}{c^2}\left(\boldsymbol{v}\times(\boldsymbol{\nabla}\times\boldsymbol{U}_E)\right)^i , \tag{8.63}
$$

where again notation $\boldsymbol{U}_E = (U^i_E)$ has been used. Assuming that the Earth is rotating stationary, one can neglect the term with the time derivative in expression (8.34). Substituting the spin-dependent term in equation (8.34) to equation (8.63) one obtains

$$
a^i_{\rm GM} = (\boldsymbol{v}\times\boldsymbol{\Omega}_{LT})^i , \tag{8.64}
$$

where $\boldsymbol{\Omega}_{LT} = (\Omega^i_{LT})$ is the angular velocity of a slow gravitomagnetic rotation of the inertial frame caused by Earth's rotation. It is given by equation

$$
\Omega^i_{LT} = \frac{2GS_E}{c^2 r^3}\left[s^i - \frac{3w^i(\boldsymbol{s}\cdot\boldsymbol{w})}{r^2}\right] , \tag{8.65}
$$

where $\boldsymbol{s} = (s^i)$ is the unit vector in the direction of the Earth's spin $\boldsymbol{S}_E = S_E\boldsymbol{s}$. The gravitomagnetic acceleration of the satellite is similar to the effect of the Coriolis acceleration acting on a moving body in a rotational frame of reference. It leads to the precession of satellite's perigee and node (the intersection of the Earth's equatorial plane with the satellite's orbit) - the effect, envisaged by Lense and Thirring in 1918 [Lense and Thirring, 1918]. The angular velocity of the Lense-Thirring precession of satellite's node and perigee is calculated on the basis of the equations for osculating elements which yield (after averaging over one period of orbital revolution)

$$\Omega^{\rm node}_{LT} = \frac{2GS_{\scriptscriptstyle E}}{c^2 a^3 (1-e^2)^{3/2}} \,, \tag{8.66}$$

$$\Omega^{\rm perigee}_{LT} = -3\cos i \,\Omega_{\rm node} \,, \tag{8.67}$$

where i is the inclination of the orbital plane of the satellite to the Earth's equator, a and e are the semi-major axis and eccentricity of the satellite's orbit.

The Lense-Thirring effect plays a major dynamical role in the accretion disks of super-massive spinning black hole and in the alignment of jets in active galactic nuclei and quasars [Melia, 2003; Thorne *et al.*, 1986]. However, Ω^i_{LT} is extremely small for a satellite orbiting the Earth. For example, for a LAGEOS satellite having an orbital semi-major axis 12,238 km the Lense-Thirring angular shift of its node is only about 33 mas yr^{-1}, that is, nearly 1.4 m yr^{-1} in space. Van Patten and Everitt [Van Patten and Everitt, 1976] discussed feasibility of the detection of the Lense-Thirring effect with two counter-orbiting satellites placed on the polar orbit. Ciufolini [Ciufolini, 1986] proposed a more general technique for doing measurement of the Lense-Thirring precession with two satellites having supplementary orbital inclinations. It allows to suppress the Newtonian "noise" signal in the precessional motion of the node and perigee of the satellite's orbit produced by the low-order harmonics $J^{i_1 i_2 \ldots i_l}_{\scriptscriptstyle E}$ in expansion of the gravitational field of the Earth. In any case, Ciufolini's method still requires precise knowledge of the multipole moments of the Earth's gravitational field. Recent advances in mapping gravitational field of the Earth with GRACE satellite allowed Ciufolini & Pavlis [Ciufolini and Pavlis, 2004] and Ries [Ries, 2009] to measure the gravitomagnetic Lense-Thirring effect by comparing a specifically-chosen, "differential" combination of the orbital elements of LAGEOS and LAGEOS II geodynamic satellites.

8.7
Post-Newtonian Gradiometry

Gravity gradiometry is a measure of differential acceleration of the Earth's gravity field over a unit distance. Gradiometry is widely used in practical explorations by oil, gas and mining companies to directly measure the density of the subsurface and, effectively, the rate of change of rock properties. From this information they can build up a picture of sub-surface anomalies and more accurately target oil, gas and mineral deposits. However, this section is mostly concerned with the application of gradiometry to the study of the fundamental properties of the Earth's gravity field.

Basic equation of gradiometry in both the Newtonian and relativistic gravity is the equation of deviation of time-like worldlines of two proof masses. In most cases these worldlines are assumed to be geodesic (see section 3.7.5) but this assumption is not required here. Measurable quantity in this equation is an invariant "vector" distance $\mathcal{X}^i = \mathcal{X}^i_2 - \mathcal{X}^i_1$ connecting the two proof masses. In what follows, one will derive a differential equation for $\mathcal{X}^i$ that will be defined as a "vector" distance taken along a light geodesic connecting both proof masses with the invariant measure of the distance between them being equal to the proper time τ measured by clocks attached to the first mass. This is equivalent to introduction of Synge's optical coordinates [Synge, 1964] of the second mass with respect to the first one. Let us stipulate that the first proof mass is always located at the origin of the topocentric reference frame with the spatial coordinates $\xi^i_1 = 0$ that means $\mathcal{X}^i_1 = 0$. Let's a light signal (photon) is emitted at time $\tau_1 = 0$ from the first proof mass. It will reach the second proof mass at time of reception $\tau_2 = \tau$ when topocentric spatial coordinates of the second proof mass are $\xi^i_2 = \xi^i$. Because the topocentric time τ coincides with the proper time of the first mass, the optical coordinates of the second mass are

$$\mathcal{X}^i = cK^i\tau\ , \tag{8.68}$$

where one has taken into account that $\mathcal{X}^i = \mathcal{X}^i_2$ in case of $\mathcal{X}^i_1 = 0$, and K^i is the unit vector in the direction of the second particle: $\delta_{ij}K^iK^j = 1$. Optical (or null) coordinates are directly measurable quantities that are invariants. Full concept of these coordinates has been worked out in the book of Synge [Synge, 1964] on the basis of the, so-called, world function. Current gradiometry indeed uses electromagnetic signals in order to track precise positions of the proof masses (satellites in case of the CRACE mission). Therefore, the concept of the optical coordinates is highly relevant to the discussion of the basic gradiometric equation.

In order to connect the optical coordinates $\mathcal{X}^i$ with the topocentric coordinates ξ^i one will solve equation of the light geodesic in the topocentric coordinates

$$\frac{d^2\xi^i}{d\tau^2} = \frac{c^2}{2}\frac{\partial G_{00}}{\partial \xi^i} - \frac{\partial G_{00}}{\partial \xi^k}\frac{d\xi^i}{d\tau}\frac{d\xi^k}{d\tau} - \left(\frac{\partial G_{ik}}{\partial \xi^p} - \frac{1}{2}\frac{\partial G_{pk}}{\partial \xi^i}\right)\frac{d\xi^p}{d\tau}\frac{d\xi^k}{d\tau}\ , \tag{8.69}$$

where all terms with the time derivatives from the topocentric metric $G_{\alpha\beta}$ have been neglected due to their smallness, and the parameter τ along the light geodesic is the topocentric time that coincides with the proper time of the first proof mass always located at the origin of the topocentric frame. Substituting the metric coefficients (8.40)-(8.42) to equation (8.69) and reducing similar terms yield

$$\frac{d^2\xi^i}{d\tau^2} = \left[\left(1 + \frac{1}{c^2}\frac{d\xi^p}{d\tau}\frac{d\xi^p}{d\tau}\right)\delta^{ip} - \frac{4}{c^2}\frac{d\xi^i}{d\tau}\frac{d\xi^p}{d\tau}\right]\left(E_p + E_{pq}\xi^q\right) , \tag{8.70}$$

where one has omitted quadratic and higher-order terms with respect to the spatial coordinates ξ^i. Equation (8.70) is solved by iterations. Unperturbed solution is $\xi^i = ck^i\tau$, where k^i is a unit vector along the unperturbed light ray: $\delta_{pq}k^pk^q = 1$. Substituting the unperturbed solution to the right side of equation (8.70) and integrating with respect to τ give us

$$\xi^i = ck^i\tau + \tau^2\left(\delta^{ip} - 2k^ik^p\right)\left(E_p + \frac{c}{3}E_{pq}k^q\tau\right) . \tag{8.71}$$

The unperturbed unit vector k^i is yet unknown and must be expressed in terms of the unit vector K^i entering the definition of optical coordinates. This definition comes about naturally from the requirement that the light signal indeed reaches the second proof mass at the time τ and at the point ξ^i. It means that $K^i = \xi^i/|\boldsymbol{\xi}|$, where $|\boldsymbol{\xi}|$ is the coordinate distance $|\boldsymbol{\xi}| = \sqrt{\delta_{pq}\xi^p\xi^q}$. Simple calculation reveals that

$$k^i = K^i - \frac{\tau}{c}\left(\delta^{ip} - k^ik^p\right)\left(E_p + \frac{c}{3}E_{pq}k^q\tau\right) . \tag{8.72}$$

After substituting this expression to equation (8.71) and taking into account definition (8.68) one obtains for the optical coordinate of the second proof mass

$$\mathcal{X}^i = \xi^i + \frac{1}{c^2}\left(E_p\xi^p + \frac{1}{3}E_{pq}\xi^p\xi^q\right)\xi^i + O\left(\frac{\xi^4}{c^2}\right) , \tag{8.73}$$

which yields the coordinate transform between the topocentric and optical coordinates.

Let us take now the second time derivative along the worldlines of the masses from both sides of equation (8.73), and substitute for the second time derivative of coordinates ξ^i its value form the equation of motion (8.45). It brings about

$$\frac{d^2\mathcal{X}^i}{d\tau^2} = \frac{F_2^i}{m_2} + E^i + \left(E_{ij} - \frac{3}{c^2}E_{<i}E_{j>}\right)\xi^j + \frac{1}{2}E_{ijp}\xi^j\xi^p + O\left(\xi^3\right) + O\left(\frac{\xi^2}{c^2}\right) , \tag{8.74}$$

where all velocity-dependent terms as well as terms with the time derivatives from E_i, E_{ij}, etc. have been neglected in accordance with previous assumptions. Equation of motion of the first proof mass is an exact relationship

$$\frac{d^2\xi^i}{d\tau^2} = \frac{F_1^i}{m_1} + E^i = 0 , \tag{8.75}$$

which follows from the condition that the first mass is always located at the origin of the topocentric frame, so that the physical force F_1^i compensates the inertial force

$m_1 E^i$. Making use of equation (8.75) in equation (8.74) gives us the basic equation of gradiometry

$$\frac{d^2\mathcal{X}^i}{d\tau^2} = F^i + K_{ij}\mathcal{X}^j + \frac{1}{2}K_{ijp}\mathcal{X}^j\mathcal{X}^p + O\left(\mathcal{X}^3\right) + O\left(\frac{\mathcal{X}^2}{c^2}\right) , \tag{8.76}$$

where

$$K_{ij} = E_{ij} - \frac{3}{c^2}E_{<i}E_{j>} , \tag{8.77}$$

$$K_{ijk} = E_{ijk} , \tag{8.78}$$

and the force per unit mass

$$f^i \equiv \frac{F_2^i}{m_2} - \frac{F_1^i}{m_1} , \tag{8.79}$$

represents all empirical non-gravitational forces that constrain the relative motion of the proof masses. In particular, if the both masses are moving along geodesics $f^i = 0$. One can easily confirm that the linear term in equation (8.76) is the component of the Riemann tensor

$$K_{ij} = -R_{0i0j} , \tag{8.80}$$

while the quadratic term is its spatial derivative

$$K_{ijp} = -R_{0i0j,p} . \tag{8.81}$$

The reader can notice that equation (8.77) apparently includes the quadratic dependence on the inertial acceleration E_i. In fact, this dependence is illusory as the component E_{ij} has not been defined yet, and, as shown later, has the same term which cancels the quadratic dependence on the acceleration. The physical reason for this is that the Riemann tensor R_{0i0j} characterizes the curvature of space-time and can not depend on the acceleration of the particles that is solely a property of its worldline alone.

The quantities K_{ij}, K_{ijk} represent the quadrupole and octupole tidal fields as measured by a local observer at the origin of the topocentric reference frame. Measurement of these fields is the main practical task of gradiometry. The explicit functional form of K_{ij}, K_{ijk} with the post-Newtonian corrections taken into account is obtained after making the asymptotic matching of the GRF and TRF metric tensors. One has

$$K_{ij} = K_{ij}^N + c^{-2}K_{ij}^{PN} + O\left(\frac{1}{c^4}\right) , \tag{8.82}$$

$$K_{ijp} = K_{ijp}^N + O\left(\frac{1}{c^2}\right) , \tag{8.83}$$

where the Newtonian tidal matrices are

$$K_{ij}^N = U_{E,<ij>}(\boldsymbol{w}_T) + 3Q_{ij} + 15Q_{ijp}w_T^p + O(w_T^2) , \tag{8.84}$$

$$K_{ijp}^N = U_{E,<ijp>}(\boldsymbol{w}_T) + 15Q_{ijp} + O(w_T) , \tag{8.85}$$

and the post-Newtonian contribution to the tidal matrix is

$$K_{ij}^{PN} = K_{ij}^{EM} + K_{ij}^{GM} + 2K_{p<i}^{N} R_{j>p} \,, \tag{8.86}$$

$$\begin{aligned} K_{ij}^{EM} = {} & \Phi_{E,<ij>}(\boldsymbol{w}_T) - 2U_E(\boldsymbol{w}_T)U_{E,<ij>}(\boldsymbol{w}_T) - \frac{1}{2}\chi_{E,uu<ij>}(\boldsymbol{w}_T) \\ & -2v_T^{<i}U_E^{,j>}(\boldsymbol{w}_T) + 2v_T^2 U_{E,<ij>}(\boldsymbol{w}_T) - 3v_T^k v_T^{<i} U_E^{,j>k}(\boldsymbol{w}_T) \\ & +3a_T^{<i}a_T^{j>} - 6a_T^{<i}U_E^{,j>}(\boldsymbol{w}_T) - 3E_{<i}E_{j>} + O\left(v_T^2 Q_{ij}\right) \,, \end{aligned} \tag{8.87}$$

$$K_{ij}^{GM} = 4\left[U_E^{<i,j>u}(\boldsymbol{w}_T) + v_T^k U_E^{<i,j>k}(\boldsymbol{w}_T) - v_T^k U_E^{k,<ij>}(\boldsymbol{w}_T)\right] . \tag{8.88}$$

These formulas are more general as compared to the results of work [Mashhoon *et al.*, 1989] that investigated the case of an isolated, spherically-symmetric and uniformly rotating Earth without explicit formulation of relativistic tidal effects of other bodies of the solar system. However, the external bodies also produce tidal effects, which appear in equations (8.84), (8.85) in the form of the external tidal multipoles Q_{ij}, Q_{ijk}, and they should be taken into account in precise gradiometry.

The term K_{ij}^N represents the Newtonian gradient of the Earth's gravity field. The main contribution to K_{ij}^N is from the second derivatives of the Newtonian potential, $U_{E,<ij>}$, describing the quadrupole tidal matrix of the Earth's gravitational field. Replacing in this term the expansion of U_E according to formula (8.33), one obtains:

$$\begin{aligned} U_{E,<ij>}(\boldsymbol{w}) = {} & \frac{3GM}{r^5}\left(w^i w^j - \frac{1}{3}r^2\delta^{ij}\right) \\ & +G\sum_{l=2}\frac{(2l+3)!!}{l!}\frac{J_E^{k_1\ldots k_l}w^{k_1}\ldots w^{k_l}}{r^{2l+5}}\left(w^i w^j - \frac{1}{2l+3}r^2\delta^{ij}\right) \\ & -2G\sum_{l=2}\frac{(2l+1)!!}{(l-1)!}\frac{J_E^{k_1k_2\ldots k_l(i}w^{j)}w^{k_1}\ldots w^{k_l}}{r^{2l+3}} \,. \end{aligned} \tag{8.89}$$

The first term in the right side of equation (8.89) is of the order of GM_E/r^3. It can reach magnitude of the order of 3000 E ($\sim 3\times10^{-6}$ s^{-2}) on Earth's surface. The next order of magnitude term in the right side of equation (8.89) is associated with the influence of the Earth's quadrupole moment J_E^{ij}. It is proportional to Earth's dynamical oblateness factor $J_2 = 0.00108263$ and reaches magnitude of the order of 70 E ($\sim 7\times10^{-8}$ s^{-2}) on Earth's surface. Earth's multipole moments of higher order $J^{i_1i_2\ldots i_l}$ ($l \geq 3$) contribute to the Newtonian gradient of the gravity force no more than 10^{-1} E. Satellite gradiometry is currently a flourishing field of the Earth's gravity field research. Two space missions - GRACE [Levi, 2003] and GOCE [Albertella *et al.*, 2002] - represent a new generation of space gradiometers consisting of two satellites (several proof masses in case of GOCE) flying closely to each other and connecting via microwave ranging link that is used for measuring the relative distance $\mathcal{X}^i$ between the satellites. Spatial resolution of GRACE is 320 km which means that it can measure all multipole harmonics in the expansion (8.89) to the order $l = 120$. GOCE main measuring device is an Electrostatic Gravity Gradiometer (EGG) that consists of three pairs of identical accelerometers, which are about 50 cm apart and form three gradiometer arms mounted orthogonal to one another on the ultra-stable structure. One arm is aligned with the satellite's trajectory, one perpendicular to the trajectory, and one pointing approximately towards the center of the Earth. By combining these different acceleration measurements, it is possible to derive the gravity-gradient components in the expansion (8.89). Spatial resolution of GOCE is 100 km which corresponds to the multipole harmonics with $l = 400$.

The second and third terms in expression (8.84) for the K^N_{ij} represent the quadrupole and octupole tidal gravity's gradient caused by the external bodies - Moon, Sun, and planets. It is sufficient in gradiometric analysis to consider the external bodies as the point masses with gravitational potential $U_A = GM_A/R_A$, where $R_A = |\boldsymbol{R}_A|$ $\boldsymbol{R}_A = R^i_A = x^i - x^i_A(t)$, $x^i_A(t)$ are the BRF coordinates of an external body A, which are connected to the geocentric coordinates by the post-Newtonian transformations given in section 5.2.6. Taking into account the Newtonian tidal terms only, one obtains (compare with equations (5.59), (5.89))

$$Q_{ij} = 3\sum_{A\neq E}\frac{GM_A}{R^3_{EA}}\left(N^i_{EA}N^j_{EA} - \frac{1}{3}\delta^{ij}\right) + O\left(\frac{1}{c^2}\right) , \tag{8.90}$$

$$Q_{ijk} = 15\sum_{A\neq E}\frac{GM_A}{R^4_{EA}}\left(N^i_{EA}N^j_{EA}N^k_{EA} - \frac{1}{5}\delta^{ij}N^k_{EA} - \frac{1}{5}\delta^{ik}N^j_{EA} - \frac{1}{5}\delta^{jk}N^i_{EA}\right) + O\left(\frac{1}{c^2}\right) \tag{8.91}$$

where $N^i_{EA} = R^i_{EA}/R_{EA}$, $R_{EA} = |\boldsymbol{R}_{EA}|$, $\boldsymbol{R}_{EA} = R^i_{EA} = x^i_E(t) - x^i_A(t)$ represents a difference between the BRF spatial coordinates of the Earth's center of mass (geocenter) and the external body A. The most important external bodies in gradiometric measurements are Moon and Sun. Magnitude of the external quadrupole Q_{ij} for the Moon is of the order of 1.7×10^{-4} E and for the Sun it amounts to 8×10^{-5} E, approximately two times smaller than the tidal quadrupole of the Moon. Magnitude of the contribution of the octupole field of the external masses to the quadrupole tidal effects is smaller by a factor of R_E/R_{EA} as compared with the direct contribution from the quadrupole components Q_{ij} seen in equation (8.84). For example, in case of the Moon $Q_{ijk}w^k_T \simeq 3\times10^{-6}$ E on the Earth's surface.

The term K^N_{ijk} describes the Newtonian octupole tidal gravitational field of the Earth. With sufficient accuracy the first term in the right side of equation (8.85) has the following form:

$$U_{E,<ijk>} = -\frac{15GM_E}{r^7}\left(w^iw^jw^k - \frac{r^2}{5}\delta^{ij}w^k - \frac{r^2}{5}\delta^{jk}w^i - \frac{r^2}{5}\delta^{ik}w^j\right) + O(J_2) . \tag{8.92}$$

Contribution of this term to the differential acceleration between two, freely-falling particles, as described by equation (8.45), is by a factor ξ/R_E smaller than the effect of the local quadrupole field E_{ij}, where ξ is a separation between the two particles. For the GOCE gradiometer, this is equivalent to the magnitude of $K^N_{ijk}\xi^k \simeq 2.3\times10^{-4}$ E, which is comparable with the tidal gravitational effect of the external quadrupole Q_{ij} from the Moon.

The term K^{PN}_{ij} represents the post-Newtonian tidal matrix of the gravity's gradient of the Earth's field. It consists of three parts: the "gravitoelectric" matrix - K^{GE}_{ij}, the "gravitomagnetic" matrix - K^{GM}_{ij}, and the term $2K^N_{p<i}R_{j>p}$, which arises owing to the relativistic precession of the dynamically non-rotating TRF with respect to the spatial axis of GRF. In principle, there are also post-Newtonian contributions from the external masses - Moon, Sun, planets - but they are so small that can be neglected without reservation (for discussion of these terms see [Kopejkin, 1991b]).

The matrix K^{GE}_{ij} is defined by the formula (8.87). It is not difficult to understand that potential Φ_E in the $g_{00}(u, \boldsymbol{w})$ component of the GRF metric tensor re-defines the Earth mass and the multiple moments because it has the same structure as the Newtonian potential U_E. Therefore, the term $\Phi_{E,<ij>}$ linearly combines with $U_{E,<ij>}$ and can not be observed separately. One stipulates in the post-newtonian approximation that the Earth is in a stationary rotation and does not change its multipole moments as times goes. It allows us to set all terms depending on partial time derivatives equal to zero. Remaining terms in equation (8.87) yield

$$K^{GE}_{ij} = \frac{3GM_E}{r^3}\left[\left(2v^2 - \frac{3GM_E}{r}\right)n^{<i}n^{j>} + v^{<i}v^{j>} - 3(\boldsymbol{n}\cdot\boldsymbol{v})v^{<i}n^{j>}\right] + O(J_2)\ , \qquad (8.93)$$

where $n^i = w^i/r$, $v^i = dw^i/du$. A magnitude of elements of the matrix (8.93) is approximately equal to several μE in the immediate vicinity of the Earth's surface. As for the matrix K^{GM}_{ij}, one can obtain the explicit expression of its main terms with the help of formulas (8.34) and (8.88):

$$K^{GM}_{ij} = \frac{6GS_E}{r^4}\left[5(\boldsymbol{n}\cdot\boldsymbol{v})(\boldsymbol{s}\times\boldsymbol{n})^{<i}n^{j>} - 5((\boldsymbol{s}\times\boldsymbol{n})\cdot\boldsymbol{v})n^{<i}n^{j>} - 3(\boldsymbol{s}\times\boldsymbol{v})^{<i}n^{j>} - (\boldsymbol{s}\times\boldsymbol{n})^{<i}v^{j>}\right]\ , \qquad (8.94)$$

where $S_E = (\boldsymbol{S}_E\cdot\boldsymbol{S}_E)^{1/2}$, $\boldsymbol{s} = s^i = S^i_E/S_E$. A magnitude of $c^{-2}K^{GM}_{ij}$ is of the order of several hundreds nanoEötvös at the low satellite orbits and approximately ten times smaller on the Earth s surface (because velocity $\boldsymbol{v}$ on the Earth's surface is that of the Earth's rotation).

Let us now specify the post-Newtonian tidal matrices K^{EM}_{ij} and K^{GM}_{ij} for a gradiometer deployed on-board of a satellite moving along a circular orbit. The circular orbit is characterized by instantaneous osculating orbital elements: eccentricity $e = 0$, semi-major axis a, inclination relative to the Earth's equator i, and angle of nodes Ω (see Figure 8.2). Because of post-Newtonian perturbations, these orbital elements, in particular i and Ω, will not be strictly constant, but instead they will vary with time. This time variations of the orbital elements are important in calculation of the Newtonian tidal matrix K^N_{ij} since they can correlate with the direct post-Newtonian terms. One emphasizes that the time variation of the orbital elements is not responsible for the precessional variation of the Newtonian tidal matrix shown as the last term in equation (8.86) which is solely due to the post-Newtonian precession of the spatial axis of the satellite reference frames described by the matrix R_{ij}. Position of the satellite on the orbit is given by the unit vector $\boldsymbol{n} = (n^i)$

$$\boldsymbol{n} = (\cos\Omega\cos\psi - \cos i\sin\Omega\sin\psi)\,\boldsymbol{e}_1 + (\sin\Omega\cos\psi + \cos i\cos\Omega\sin\psi)\,\boldsymbol{e}_2 + \sin i\sin\psi\boldsymbol{e}_3\ , \qquad (8.95)$$

where $\boldsymbol{e}_1 = (1, 0, 0)$, $\boldsymbol{e}_2 = (0, 1, 0)$, $\boldsymbol{e}_3 = (0, 0, 1)$ are the unit vectors along w^1, w^2, w^3 axis respectively, $\psi = (2\pi/P_b)u$ is the angular position of the satellites, and

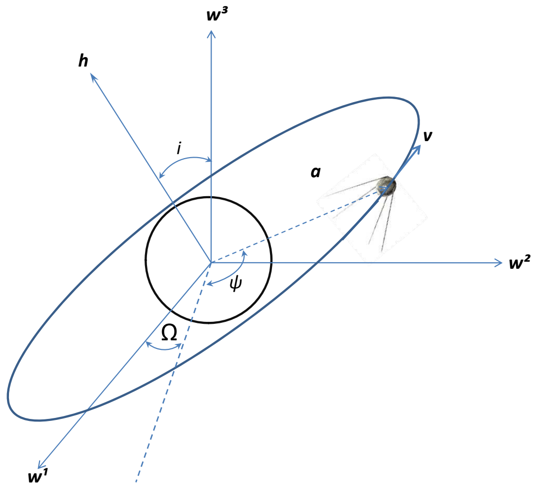

Figure 8.2 Satellite's circular orbit is shown in the geocentric coordinates $\boldsymbol{w} = (w^1, w^2, w^3)$. Earth's equatorial plane coincides with (w^1, w^2) plane. The satellite moves counterclockwise with velocity $\boldsymbol{v}$ and the angular momentum $\boldsymbol{h} = \boldsymbol{w} \times \boldsymbol{v}$ per unit mass.

$P_b = 2\pi\sqrt{a^3/GM_E}$ is its orbital period. Osculating velocity of the satellite $\boldsymbol{v} = d\boldsymbol{w}/du = -(GM_E/a)^{1/2}\,\hat{\boldsymbol{v}}$ where the unit vector

$$\hat{\boldsymbol{v}} = (\cos\Omega\sin\psi + \cos i\sin\Omega\cos\psi)\,\boldsymbol{e}_1 + (\sin\Omega\sin\psi - \cos i\cos\Omega\cos\psi)\,\boldsymbol{e}_2 - \sin i\cos\psi\,\boldsymbol{e}_3\ . \tag{8.96}$$

In what follows, one accepts that the node coincides with w^1 axis, that is the angle $\Omega = 0$. Because the orbit is circular the dot product $\boldsymbol{n}\cdot\boldsymbol{v} = 0$. Tedious but straightforward calculations give for the Newtonian (mass-induced) and post-Newtonian tidal matrices the following expressions [Mashhoon *et al.*, 1989]

$$K^N_{ij} = \frac{3GM_E}{2a^3}\begin{bmatrix} \frac{1}{3} + \cos 2\psi & \cos i\sin 2\psi & \sin i\sin 2\psi \\ \cos i\sin 2\psi & \frac{1}{3} - \sin^2 i - \cos^2 i\cos 2\psi & \sin i\cos i(1-\cos 2\psi) \\ \sin i\sin 2\psi & \sin i\cos i(1-\cos 2\psi) & \frac{1}{3} - \cos^2 i - \sin^2 i\cos 2\psi \end{bmatrix} \tag{8.97}$$

$$K^{GE}_{ij} = \frac{3G^2M_E^2}{c^2a^4}\begin{bmatrix} -\cos 2\psi & -\cos i\sin 2\psi & -\sin i\sin 2\psi \\ -\cos i\sin 2\psi & \cos^2 i\cos 2\psi & \sin i\cos i\cos 2\psi \\ -\sin i\sin 2\psi & \sin i\cos i\cos 2\psi & \sin^2 i\cos 2\psi \end{bmatrix} \tag{8.98}$$

$$K^{GM}_{ij} = \frac{6GS_E\nu}{c^3a^3}\begin{bmatrix} -\cos i\cos^2\psi & \left(1-\frac{3}{2}\cos^2 i\right)\sin 2\psi & -\frac{3}{2}\sin i\cos i\sin 2\psi \\ \left(1-\frac{3}{2}\cos^2 i\right)\sin 2\psi & -\cos i\left(1-5\sin^2 i\right)\sin^2\psi & \sin i\left(\frac{3}{2} - 2\cos^2\psi - 5\cos^2 i\sin^2\right. \\ -\frac{3}{2}\sin i\cos i\sin 2\psi & \sin i\left(\frac{3}{2} - 2\cos^2\psi - 5\cos^2 i\sin^2\psi\right) & \cos i\left(1 - 5\sin^2 i\sin^2\psi\right) \end{bmatrix} \tag{8.99}$$

Various technical aspects of the ultra-precision gradiometric measurements have been discussed during the past 20 years. Relativistic gravity's gradient K^{PN}_{ij} can be measured only in a space mission with a dedicated super-conducting gradiometer which was under development in the University of Maryland since the middle of 80-th [Chan *et al.*, 1987; Chan and Paik, 1987; Moody *et al.*, 1986; Paik, 1981]. Recent results in this development have been published by Paik in paper [Paik, 2008]. It is important to note that the measurement of the gravitoelectric matrix K^{GE}_{ij} is impossible as long as the inclination of the satellite's orbit is kept constant. The reason is that the functional structure of the tidal matrix K^{GE}_{ij} is almost identical to the Newtonian matrix K^N_{ij} [Kopejkin, 1991a; Mashhoon *et al.*, 1989]. Furthermore, the influence of the post-Newtonian tidal forces K^{GE}_{ij} on gradiometer will be masked by the errors in numerical values of the multipole moments $J_E^{i_1 i_2 \dots i_l}$ of the Earth. On the other hand, a vortex nature of a gravitomagnetic potential U_E^i of the Earth leads to the structure of the gravitomagnetic matrix K_{ij}^{GM}, which is quite different from both K_{ij}^N and K_{ij}^{GE}. Thus, the influence of the gravitomagnetic potential U_E^i on gradiometer can be separated from the Newtonian "noise". This idea was originally discussed by Braginsky & Polnarev [Braginskiĭ and Polnarev, 1980].

Another exciting application of space gradiometry for deeper understanding the laws of fundamental gravity is the satellite test of the equivalence principle (STEP) - the idea proposed by Worden and Everitt [Worden and Everitt, 1974]. The equivalence principle states the equality between gravitational, m_g, and inertial m_i mass, which makes the bodies of different composition falling with the same acceleration in a homogeneous and time-independent gravitational field. This postulate ($m_g = m_i$) cannot be proven, it can only be tested to higher and higher precision. There is no a priori reason why it should be strictly valid. Today, the experiments have reached a level of a few parts in 10^{-13} with a torsion balance in a ground-based laboratory [Schlamminger *et al.*, 2008] and lunar-laser ranging [Williams *et al.*, 2009]. A non-null result would constitute the discovery of a new fundamental interaction of nature [Damour, 2009b]. STEP will advance the testing of the equivalence principle from several parts in 10^{-13} to 1 part in 10^{-18}.

STEP will compare the accelerations of four pairs of proof masses in orbit [Worden *et al.*, 2000]. The drag-free control system on board of the satellite will make the masses free-floating and isolated from disturbing non-gravitational force f^i inside a cryogenic dewar with super-conducting shielding and ultra-high vacuum. Differential accelerations will be measured by a super-conducting circuit using a quantum interference device (SQUID) for the best sensitivity. Gravity gradient disturbances caused by the tidal matrix K_{ij}, are eliminated by precise placement of the mass centers on each other. Under these circumstances the equation (8.76) can be written-down as follows

$$\frac{d^2 X^i}{d\tau^2} = \eta \frac{\partial U_E}{\partial w^i} \,, \tag{8.100}$$

where the gradient of the geopotential U_E is multiplied with the Eötvös parameter

$$\eta = 2\frac{(m_g/m_i)_2 - (m_g/m_i)_1}{(m_g/m_i)_2 + (m_g/m_i)_1} \,, \tag{8.101}$$

characterizes the empirical violation of the equivalence principle due to the possible difference between the ratios m_g/m_i of the gravitational and inertial mass in each arm of the gradiometer. The STEP mission will be flown in a near-circular sun-synchronous orbit, to minimize temperature variations, for a period of six months. The best altitude is approximately 550 km.

Two other, less ambitious space missions have been proposed. MICROSCOPE is a French space mission for testing the equivalence principle [Touboul and Rodrigues, 2001]. The mission goal is the determination of the Eötvös parameter η with an accuracy of 10^{15}. GG (Galileo Galilei) mission is under development by the Italian

Space Agency [Nobili *et al.*, 2009]. The mission intends to measure the Eötvös parameter with accuracy up to 10^{-17}.

8.8 Relativistic Geoid

8.8.1 Definition of geoid in the post-Newtonian gravity

The geoid in the Newtonian gravity is defined as the equipotential surface which would coincide exactly with the mean ocean surface of the Earth, if the ocean's water were in equilibrium, at rest, and extended through the continents such as, for instance, with very narrow canals. The geoid is the physical figure of the Earth, a smooth but fairly irregular surface that corresponds not to the actual surface of the Earth's crust, but to a surface which can only be known through extensive gravitational measurements and calculations (see Figure 8.1). Despite being an important concept for almost two hundred years in the history of geodesy and geophysics, it has only been defined to high precision in recent decades. The geoid should be distinguished from the idealized geometrical figure of a reference ellipsoid of the Earth. The ellipsoid is used to approximate the bulk of the Earth's shape, and departures from the ellipsoid are represented by the geoid elevation above or below the ellipsoid. The geoid can be as low as 106 meters below the ellipsoid or as high as 85 meters above. The geoid models are periodically updated since the goal of a geoid model is to be consistent with actual physical characteristics as determined from current observations. Older geoid models, such as GEOID99 and even GEOID03, were built using observation data from the time of their formation, but are considered no longer consistent with the physical Earth. Satellite geodesy [Kaula, 1966; Milani *et al.*, 1987; Vaníček and Krakiwsky, 1986] brought important discoveries in theory of geoid computation. The Precise Geoid Solution (PGS) by P. Vaníček and co-workers [Martinec and Vaníček, 1994; Vaníček and Martinec, 1994] significantly improved on geoid computation. The solution is based on the approach formulated by Sir Gabriel Stokes in the 19th century and extended with F. R. Helmert's idea of condensing topography onto the geoid. It differs from M. S. Molodensky's approach [Moritz and Molodensky, 1971] where the geoid is replaced by a similar surface called quasi-geoid. The quasi-geoid is derived from gravity data procured at the surface of the Earth via an integral equation. The integration is carried out on a surface similar to the surface of the Earth, called the telluroid. The mathematical beauty of the Molodensky's approach has practical difficulty with integration over an irregular surface such as the telluroid. The approach by P. Vaníček and co-workers overcomes this difficulty and enables millimeter-to-centimeter accuracy in geoid computation - an order-of-magnitude improvement from previous classical solutions. However, further improvements face a serious problem as the post-Newtonian corrections to the Newtonian gravity field equations are of the same order of magnitude and must be taken into account properly to avoid undesirable bias in determination of the geoid.

In the framework of general relativity there exist two definitions of the relativistic geoid [Soffel *et al.*, 1988b; Soffel, 1989]:

- **Definition 1**. The relativistic u-geoid represents a two-dimensional surface, being close to the mean sea level, in any point of which the rate of the proper time τ of an ideal clock is constant with respect to the geocentric coordinate time u, i.e.

$$d\tau/du = \text{const.}\,, \tag{8.102}$$

under conditions $r^i_T = dr^i_T/du = 0$.

- **Definition 2.** The relativistic a-geoid represents a two-dimensional surface of constant geocentric time u, being close to the mean sea level and orthogonal everywhere to the topocentric direction of the Earth's gravity E_i (the plumb line), i.e.

$$E_i d\xi^i = 0\,, \tag{8.103}$$

under conditions $u = \text{const.}$

8.8.2
Post-Newtonian u-geoid

Let us begin with the derivation of the equation of the u-geoid. For this sake, one differentiates formula (8.46) with respect to the time u, take $r^i_T = dr^i_T/du = 0$, and, after that use formulas (8.48), (8.49). Then, condition (8.102) defines function W_0 that is constant on the surface of the relativistic u-geoid [Kopejkin, 1991a]

$$\begin{aligned} W_0 &= \frac{1}{2}v^2 + U_E + Q_i w^i + \frac{3}{2}Q_{ij}w^i w^j + \frac{5}{2}Q_{ijk}w^i w^j w^k \\ &+ \frac{1}{c^2}\left[\frac{1}{8}v^4 + \frac{3}{2}v^2 U_E - 4v^i U^i_E - \frac{1}{2}U^2_E + \Phi_E - \frac{1}{2}\chi_{E,uu}\right]\,, \end{aligned} \tag{8.104}$$

where one has suppressed the index T near the geocentric coordinates and velocity of observer because equation (8.104) is valid everywhere on the u-geoid. Thus, the surface of the u-geoid is the one of the constant potential W_0 which coincides in the Newtonian limit with the classical potential of the force of gravity where the term $1/2v^2$ represents the potential of the centrifugal force due to the Earth's rotation.

As follows from the derivation of equation (8.104), definition of the u-geoid is quite general, since it does not require any specific assumptions about distribution of density, velocity field and other characteristics of the Earth's matter or the requirement of the hydrostatic behavior of the Earth tides. Equation (8.104) makes more exact the equation of u-geoid derived by Soffel et al. [Soffel *et al.*, 1988b; Soffel, 1989].

8.8.3
Post-Newtonian a-geoid

Let us now derive equation of the a-geoid. First, one expresses the differential $d\xi^i$ of the topocentric coordinates on the geoid's surface in terms of the differential dw^i of the geocentric coordinates with the help of formula (8.47) and the conditions

$du = 0$ following from the definition of the a-geoid representing the hypersurface of a constant geocentric time. One has (index 'T' is suppressed again)

$$d\xi^i = dw^i + \frac{1}{c^2}\left(\frac{1}{2}v^i v^j + R^{ij} + Z^{ij}\right)dw^j + O\left(\frac{r_T}{c^2}\right) + O\left(\frac{1}{c^4}\right) . \tag{8.105}$$

Second, one calculates the scalar product of the acceleration of gravity E_i defined by equation (8.54), and $d\xi^i$ represented by equation (8.105). It yields

$$E_i d\xi^i = dW_0 + \frac{4}{c^2}\left(U^{i,u}_{E} + v^k U^{i,k}_{E} + U^{k}_{E} v^{k,i}\right) , \tag{8.106}$$

where the function W_0 is given by formula (8.104). Now, one supposes that the Earth's matter has a constant rigid-body rotation with velocity $v^i = (\boldsymbol{\omega}\times\boldsymbol{w})^i$, where ω^i is the angular velocity of the rotation. Then the-vector-potential U^i_{E} can be expressed as

$$U^i_{E} = v^i U_{E} + \epsilon^{ijk}\omega^j \chi_{E,k} . \tag{8.107}$$

This expression must be replaced in the corresponding terms in right side of equation (8.106), which are not incorporated to the total derivative dW_0. One obtains, rather remarkably that all such terms cancel out exactly. Thus, finally one gets a simple result

$$E_i d\xi^i = dW_0 , \tag{8.108}$$

which, along with the definition of the a-geoid, tell us that the equation of the a-geoid is the same as that for u-geoid and is given by formula (8.104). Moreover, one concludes from formula (8.108) that Earth's gravity force can be explicitly expressed on the relativistic geoid's surface in terms of the gradient of potential W_0

$$E_i = \left(\frac{\partial W_0}{\partial \xi^i}\right)_{u=\text{const.}} = \left(\frac{\partial W_0}{\partial w^k}\right)\left(\frac{\partial w^k}{\partial \xi^i}\right) . \tag{8.109}$$

It is instructive from methodological point of view to compare the equation of the relativistic geoid with that of the relativistic level surface of a self-gravitating and rotating fluid ball. It is well-known that in the Newtonian theory the level surface of the self-gravitating fluid coincides exactly with the equipotential surface. But is it hold true in the post-Newtonian approximation of general relativity? In other words, can the relativistic geoid be an equilibrium figure like in the Newtonian gravity?

8.8.4
Post-Newtonian level surface

In order to answer this question let us suppose that the Earth's matter consists of the perfect fluid with the energy-momentum tensor $T^{\alpha\beta}$ in the form of equations (8.2) and (8.3). The relativistic level surface is defined by means of condition of a constant pressure $p = 0$. The post-Newtonian hydrodynamic equation of the perfect fluid are derived from the law of conservation of the energy-momentum tensor that has the following form

$$\left\{\rho v^i\left[1+\frac{1}{c^2}\left(v^2+\Pi+2U_E\right)\right]+\frac{pv^i}{c^2}\right\}_{,u}+ \tag{8.110}$$
$$\left\{\rho v^i v^k\left[1+\frac{1}{c^2}\left(v^2+\Pi+2U_E\right)\right]+\frac{pv^iv^k}{c^2}\right\}_{,k}$$
$$+p_{,i}-\rho U_{E,i}-\rho\left(Q_i+3Q_{ik}w^k+\frac{15}{2}Q_{ijk}w^jw^k\right)$$
$$+\frac{\rho}{c^2}\left(4v^kU_E^{k,i}-4v^kU_E^{i,k}-4U_E^{i,u}=4v^iU_{E,u}+4v^iv^kU_{E,k}\right.$$
$$\left.-2v^2U_{E,i}+\frac{1}{2}\chi_{E,iuu}-2\Phi_{E,i}+2U_EU_{E,i}-\Pi U_{E,i}\right)$$
$$-\frac{1}{c^2}\left(2p_{,i}U_E+pU_{E,i}\right)=0\,.$$

Suppose that the Earth's matter has a constant rigid-body rotation about a fixed axis, and tides from external bodies (Sun and Moon) are in the hydrostatic equilibrium. These conditions mean that the Earth's density ρ, pressure p, specific internal energy Π and the Newtonian gravitational potential U_E do not depend on geocentric time u and their total time derivatives are equal to zero. Equation (8.110) can be drastically simplified and recast to a more compact form

$$\frac{\partial p}{\partial w^i}=\left(\rho c^2+\rho\Pi+p\right)\frac{\partial\log u^0}{\partial w^i} \tag{8.111}$$

where

$$u^0=\left[-\hat{g}_{00}(u,\boldsymbol{w})-2\hat{g}_{0i}(u,\boldsymbol{w})v^i-\hat{g}_{ik}(u,\boldsymbol{w})v^iv^k\right]^{-1/2} \tag{8.112}$$

represents the time component of the four-velocity of an element of the perfect fluid, $v^i=(\boldsymbol{\omega}\times\boldsymbol{w})^i$ is three-dimensional velocity of the fluid'd element, and the geocentric metric $\hat{g}_{\alpha\beta}$ is defined according to formulas (8.19)-(8.21). Let us note that equation (8.111) is the post-Newtonian analog of the Newtonian fundamental equation of hydrostatic equilibrium.

Take now the curl from the left and right sides of equation (8.111). Since the curl from a gradient $\partial f/\partial w^i$ of any scalar function f is equal to zero identically, one obtains the following relations

$$\boldsymbol{\nabla}\rho\times\boldsymbol{\nabla}p=0\,,\qquad\qquad \boldsymbol{\nabla}\Pi\times\boldsymbol{\nabla}p=0\,, \tag{8.113}$$

where the symbol $\boldsymbol{\nabla}$ denotes the operator of gradient $\boldsymbol{\nabla}\equiv(\partial/\partial w^i)$, and the symbol $\times$ denotes the Euclidean cross-product of two vectors. Relations (8.113) mean that any surface of a constant pressure p will coincide with the surfaces of constant density ρ and/or constant specific internal energy Π. Because the level surface is defined by the condition $p=0$, it can be also defined as the surface of constant density and/or the internal energy. Therefore, from equation (8.111) it follows that the relativistic level surface has the constant time component u^0 of four-velocity. It is not difficult to show that after substituting expressions (8.19)-(8.21) of the geocentric metric tensor into formula (8.112) one obtains:

$$u^0 = W_0 + \text{const.} \tag{8.114}$$

Thus, the equation of the relativistic level surface is the same as that of the relativistic geoid. From this derivation one can conclude that the relativistic geoid is the equilibrium figure, if any deviation of the Earth's matter from the perfect fluid and rigid rotation are neglected. This result facilitates incorporation of the relativistic concepts to geodesy because the primary definition of the geoid and the level surface is directly extrapolated from the Newtonian to the post-Newtonian approximation.

8.8.5
Post-Newtonian Clairaut's equation

The post-Newtonian Clairaut's equation is derived from the equation (8.104) where the terms of the order of $v^2 U_E$ were neglected due to their smallness. Then, the equipotential surface derived in equation (8.104) is simplified to

$$W_0 = \frac{1}{2}v^2 + U_E + \frac{1}{c^2}\left(\Phi_E - \frac{1}{2}U_E^2\right) = \text{const.} \tag{8.115}$$

Let us introduce a new potential $U_E^* \equiv U_E + c^{-2}\Phi_E$, which can be presented in the form of the integral

$$U_E^* = G\int_{V_E} \frac{\varepsilon^*(u, \boldsymbol{w}')d^3w'}{|\boldsymbol{w} - \boldsymbol{w}'|}\ , \tag{8.116}$$

with the effective density ε^* defined as

$$\varepsilon^* = \rho + \frac{1}{c^2}\left(\rho\Pi - \rho U_E + 3p\right)\ . \tag{8.117}$$

Equation (8.115) can be re-written now in the form:

$$W_0 = \frac{1}{2}v^2 + U_E^* - \frac{U_E^{*2}}{c^2}\ . \tag{8.118}$$

As it follows from equations (8.111) and (8.113), the potential W_0 is a function of density ρ only. Taking into account a formal analogy of the integral forms of U_E^* and U_E, one can use a standard technique of classical geodesy [Moritz, 1989] and split U_E^* in the two parts - the interior and exterior potentials at an internal point $\mathcal{P}$ lying on the surface of constant density $S_{\mathcal{P}}$ inside the body. Then, by making use of the well-known Legendre series for a centrifugal potential and function $|\boldsymbol{w} - \boldsymbol{w}'|^{-1}$, one obtains for W_0 the following main terms in the Legendre expansion with respect to the Legendre polynomials

$$W_0 = \frac{4\pi G}{3}\alpha^2\varepsilon^*\left[A_0(\alpha) + A_2(\alpha)P_2(\cos\theta)\right] - \frac{8\pi^2G^2}{9c^2}\alpha^4\varepsilon^{*2}\left[A_0^2(\alpha) + 2A_0(\alpha)A_2(\alpha)P_2(\cos\theta)\right]\ ,$$

(8.119)

where ε^* is the mean effective density of the Earth, α is the mean radius of the surface $S_{\mathcal{P}}$ divided by the mean radius R_E of the Earth and the functions $A_0(\alpha)$, $A_2(\alpha)$ are defined, for example, in [Moritz, 1989, equation 104] with replacement of density ρ by ε^*. Since $W_0 = W_0(\alpha)$ does not depend on the angle θ, it follows from (8.119) that $A_2(\alpha) = 0$. This condition yields the post-Newtonian Clairaut's equation, which retains in the approximation under consideration its classical form [12]

$$\frac{d^2 f}{dr^2} + \frac{6}{r}\frac{\varepsilon^*}{D(r)}\frac{df}{dr} + \frac{6}{r^2}\left[\frac{\varepsilon^*}{D(r)} - 1\right] f = 0\,, \tag{8.120}$$

where $f = f(r)$ represents the flattening of the Earth, and

$$D(r) = \frac{3}{r^3}\int_0^r \varepsilon^*(q) q^2 dq\,, \tag{8.121}$$

denotes the mean effective density inside the surface $S_{\mathcal{P}}$ calculated with taking into account the post-Newtonian contribution.

When deriving the Clairaut's equation (8.120) one neglected all terms of the order of f^2, f^3, and so on. In principle, development of a self-consistent relativistic theory of the equilibrium Figures of celestial bodies of the Solar system requires taking into account these terms.

12) Let us recall again that one has neglected in W_0 the terms of the order of $v^2 U_E$

9
Relativity in IAU Resolutions

9.1 Introduction

A series of resolutions passed by the International Astronomical Union (IAU) at its General Assemblies in 1997 and 2000 represents the most significant set of international agreements in positional astronomy in several decades. The approval of these resolutions culminated a process – not without controversy – that began with the formation of an inter-commission Working Group on Reference Systems at the 1985 IAU General Assembly in Delhi.

The resolutions came at the end of a remarkable decade for astrometry, geodesy, and dynamical astronomy. That decade witnessed the successes of the Hipparcos satellite and the Hubble Space Telescope (in both cases, after apparently fatal initial problems), the completion of the Global Positioning System, 25-year milestones in the use of very long baseline interferometry (VLBI) and lunar laser ranging (LLR) for astrometric and geodetic measurements, the discovery of Kuiper Belt objects and extra-solar planets, and the impact of comet Shoemaker-Levy 9 onto Jupiter. At the end of the decade, interest in near-Earth asteroids and advances in sensor design were motivating plans for rapid and deep all-sky surveys. Significant advances in theory also took place, facilitated by inexpensive computer power and the Internet. Positional and dynamical astronomy were enriched by a deeper understanding of chaos and resonances in the solar system, advances in the theory of the rotational dynamics of the Earth, and increasingly sophisticated models of how planetary and stellar systems form and evolve. It is not too much of an exaggeration to say that as a result of these and similar developments, the old idea that astrometry is an essential tool of astrophysics was rediscovered. The IAU resolutions thus came at a fortuitous time, providing a solid framework for interpreting the modern high-precision measurements that are revitalizing so many areas of astronomy.

This chapter is an attempt to explain these resolutions and provide guidance on their implementation. General relativity is at the heart of these resolutions; their ultimate purpose is to establish a modern physical foundation for astronomical measurements and models that can serve us well into the 21st century. This chapter also covers the IAU resolutions passed in 2006 and 2009, which refine the basic changes made in 1997 and 2000.

Of course, the IAU resolutions of 1997-2000 did not arise in a vacuum. Many people participated in various IAU working groups, colloquia, and symposia in the 1980s and 1990s on these topics, and some important resolutions were in fact passed by the IAU in the early 1990s. Furthermore, any set of international standards dealing with such fundamental matters as space and time must to some extent be based on, and provide continuity with, existing practice. Therefore, many of the new resolutions carry "baggage" from the past, and there is always the question of how much of this history (some of it quite convoluted) is important for those who simply wish to implement the latest recommendations. Material in this chapter generally avoids detailed history in an effort to present the most succinct and least confusing picture possible. However, many readers will be involved with modifying existing software systems, and some mention of previous practice is necessary simply to indicate what needs to be changed. A limited amount of background material also sometimes aids in understanding and provides a context for the new recommendations. The reader should be aware that the presentation of such material is selective and one has made no attempt at historical completeness.

It should be emphasized that the resolutions described here affect astronomical quantities only at the level of some tens of milliarcseconds or less at the present epoch. And, despite some initial misinformation to the contrary, familiar concepts such as the equinox and sidereal time have not been discarded. The largest systematic change is due to the new rate of precession, which is 0.3 arcsecond per century less than the previous (1976) rate; the change affects some types of astronomical coordinates and sidereal time. Astronomical software applications that work acceptably well now at the arcsecond or 0.1-arcsecond level (which would include most telescope control systems) will continue to work at that level, even when used with new sources of reference data, such as the Hipparcos, Tycho-2, or 2MASS star catalogs or one of the VCS radio source catalogs. Applications that are independent of the rotation of the Earth, such as those for differential (small-field) astrometry, are largely unaffected. For these kinds of systems, changes to computer code that implement the new resolutions are recommended as a long-term goal, to maintain standardization of algorithms throughout the astronomical community, but are not an immediate practical necessity.

9.1.1
Overview of the Resolutions

The IAU resolutions described in this chapter cover a range of fundamental topics in positional astronomy:

- **Relativity** Resolutions passed in 2000 provide the relativistic metric tensors for reference systems with origins at the solar system barycenter and the geocenter, and the transformation between the two systems. While these are mostly of use to theorists – for example, in the formulation of accurate models of observations – they provide a proper relativistic framework for current and future developments in precise astrometry, geodesy, and dynamical astronomy. (See section 9.2.)
- **Time Scales** Resolutions passed in 1991, 2000, and 2006 provide the definitions of various kinds of astronomical time and the relationships between them. Included are time scales based on the Système International (SI) second ("atomic" time scales) as well as those based on the rotation of the Earth. (See section 9.3.)
- **The Fundamental Astronomical Reference System** A resolution passed in 1997 established the International Celestial Reference System (ICRS), a high precision coordinate system with its origin at the solar system barycenter and "space fixed" (kinematically non-rotating) axes. The resolution included the specification of two sets of benchmark objects and their coordinates, one for radio observations (VLBI-measured positions of point-like extragalactic sources) and one for optical observations (Hipparcos-measured positions of stars). The IAU replaced the list of radio sources and their coordinates in 2009 while maintaining the overall ori-

entation of the system. These two sets of reference objects provide the practical implementation of the system and allow new observations to be related to it. (See section 9.4.)

- **Precession and Nutation** Resolutions passed in 2000 and 2006 provided a new precise definition of the celestial pole and endorsed specific theoretical developments for computing its instantaneous motion. The celestial pole to which these developments refer is called the Celestial Intermediate Pole (CIP); the instantaneous equatorial plane is orthogonal to the CIP. There are now precise algorithms for computing the pole's position on the celestial sphere at any time, in the form of new expressions for precession and nutation. (See section 9.6.)
- **Earth Rotation** A resolution passed in 2000 establishes new reference points, one on the celestial sphere and one on the surface of the Earth, for the measurement of the rotation of the Earth about its axis. The new points are called, respectively, the Celestial Intermediate Origin (CIO) and the Terrestrial Intermediate Origin (TIO). Both lie in the instantaneous equatorial plane. The rotation of the Earth is simply the geocentric angle, θ, between these two points, a linear function of Universal Time (UT1). The CIO is analogous to the equinox, the reference point on the celestial sphere for sidereal time. Unlike the equinox, however, the CIO has no motion along the instantaneous equator, and unlike sidereal time, θ is not "contaminated" by precession or nutation. The new CIO-TIO-based Earth rotation paradigm thus allows a clean separation of Earth rotation, precession, and nutation in the transformation between terrestrial and celestial reference systems. (See section 9.7.)

This chapter also includes a brief description (in section 9.5) of the de facto standard solar system model, produced and distributed by the Jet Propulsion Laboratory. This model, labeled DE405/LE405, provides the positions and velocities of the eight major planets, Pluto, and the Moon with respect to the solar system barycenter for any date and time between 1600 and 2200. The positions and velocities are given in rectangular coordinates, referred to the ICRS axes. This ephemeris is not the subject of any IAU resolutions but has become widely adopted internationally; for example, it is the basis for the tabulations in *The Astronomical Almanac* and it underlies some of the other algorithms presented in this chapter. However, DE405/LE405 has been in widespread use for about a decade and will likely be superseded within the next few years.

The IAU resolutions on positional astronomy adopted between 1997 and 2009 form an interrelated and coherent set of standards for positional astronomy. For instance, the definitions of the SI-based time scales rely on the relativity resolutions, and the position of the Celestial Intermediate Pole and the Celestial Intermediate Origin can only be properly computed using the new precession and nutation expressions. Many other links between the resolutions exist. In fact, attempting to apply the resolutions selectively can lead to quite incorrect (or impossible to interpret) results. This chapter is meant to provide an explanatory and computational framework for a holistic approach to implementing these resolutions in various astronomical applications. Let us hope that what is presented here does justice to the efforts of the many people who worked very hard over the last decade to take some important scientific ideas and work out their practical implications for positional astronomy, to the benefit of the entire scientific community.

9.1.2
About this Chapter

The six main subject areas described above are discussed in sections of this chapter. Each of the sections contains a list of the relevant IAU resolutions, a summary of the recommendations, an explanatory narrative, and, in most sections, a collection of formulas used in implementing the recommendations.

One assumes that readers have a basic knowledge of positional astronomy; that the terms right ascension, declination, sidereal time, precession, nutation, equinox, ecliptic, and ephemeris are familiar. Some experience in computing some type of positional astronomy data is useful, because the ultimate purpose of the chapter is to enable such computations to be carried out in accordance with the IAU resolutions that are discussed. The explanatory narratives deal primarily with new or unfamiliar concepts introduced by the resolutions – concepts that would not generally be described in most introductory textbooks on positional astronomy. This chapter is not a substitute for such textbooks.

IAU resolutions are referred to in the text in the form "resolution N of year", for example, "resolution B1.2 of 2000". The year refers to the year of the IAU General Assembly that passed the resolution. The proceedings of each General Assembly, including the text of the resolutions, are usually published the following year. An online reference for the text of IAU resolutions (beginning with those passed at the 1994 General Assembly) is the IAU Information Bulletin (IB) series distributed by IAU Secretariat [2010]. Resolutions are printed in the January IB following a General Assembly, i.e., IB numbers 74, 81, 88, 94, etc. An appendix C at the end of this book contains the complete text of the resolutions passed by the 1997, 2000, 2006, and 2009 General Assemblies, which are the focus of attention in this chapter.

9.1.3
Other Resources

An increasing number of publications, data, and software related to the recent IAU resolutions are becoming available.

A major online resource for implementing the IAU resolutions involving Earth rotation and time (see sections 9.3, 9.6, and 9.7) is the document of conventions used by the International Earth Rotation and Reference Systems Service (IERS): IERS Technical Note No. 32, edited by McCarthy and Petit [2004]. It is available in printed form from the IERS and also on the web [McCarthy and Petit, 2003b]. The online document contains links to Fortran subroutines that implement the recommended models. The document also contains algorithms specific to geodetic applications, such as tidal and geopotential models, that have not been the subject of IAU action and are not discussed in this chapter. The IERS also maintains an online list of FAQs on the IAU resolutions [IERS, 2010]. A new version of the IERS Conventions is in preparation and will be released at the end of 2010.

The IAU Working Group on Nomenclature for Fundamental Astronomy (2003-2006) has a website [Capitaine *et al.*, 2007] with many helpful documents, including a list of definitions (some of which are used in this chapter) and other educational material.

In addition to the IERS software, two other packages of computer subroutines are available for implementing the IAU resolutions: the Standards of Fundamental Astronomy (SOFA) [IAU SOFA Center, 2009], and the Naval Observatory Vector Astrometry Software (NOVAS) [Kaplan *et al.*, 2009]. Both packages are available in both Fortran and C. SOFA is a collection of routines managed by an international panel, the SOFA Reviewing Board, that works under the auspices of IAU Division 1. The board has adopted a set of coding standards for algorithm implementations and is soliciting code from the astrometric and geodetic communities that implements IAU models. Procedures (Fortran subroutines or C functions) are adapted to the coding standards and validated for accuracy before being added to the SOFA collection. NOVAS is an integrated package of procedures for the computation of a wide variety of common astrometric quantities and transformations. NOVAS dates back to the 1970s but has been continually updated to adhere to subsequent IAU resolutions.

The Astronomical Almanac [2010], beginning with the 2006 edition, is also a resource for implementing the IAU resolutions. Not only does it list various algorithms arising from or consistent with the resolutions, but its tabular data serve as numerical checks for independent developments. Both SOFA and NOVAS subroutines are used in preparing the tabulations in *The Astronomical Almanac*, and various checks have been made to ensure the consistency of the output of the two software packages.

9.2 Relativity

Relevant IAU resolutions:

- A4.I, A4.II, A4.III, A4.IV of 1991;
- B1.3, B1.4, B1.5 of 2000;
- B2 (rec. 2) of 2006.

In 2000, the IAU defined a system of spacetime coordinates for (1) the solar system, and (2) the Earth, within the framework of general relativity, by specifying the form of the metric tensors for each and the four-dimensional spacetime transformation between them. The former is called the Barycentric Celestial Reference System (BCRS) and the latter is called the Geocentric Celestial Reference System (GCRS). The BCRS is the system appropriate for the production of the basic ephemerides of solar system objects and the storage of astrometric reference data on galactic and extragalactic objects. The GCRS is the system appropriate for describing the rotation of the Earth, the orbits of Earth satellites, and geodetic quantities such as instrument locations and baselines. The analysis of precise observations inevitably involves quantities expressed in both systems and the transformations between them.

9.2.1 Background

Although the theory of relativity has been with us for over a century (Einstein's first papers on special relativity were published in 1905), it has only been within the last few decades that it has become a routine consideration in positional astronomy. The reason is simply that the observational effects of both special and general relativity are small. In the solar system, deviations from Newtonian physics did not need to be taken into account – except for the advance of the perihelion of Mercury – until the advent of highly precise space techniques in the 1960s and 1970s: radar ranging, spacecraft Doppler tracking, very long baseline interferometry (VLBI), pulsar timing, and satellite/lunar laser ranging (SLR/LLR). More recently, even optical astrometry has joined the list, with all-sky satellite measurements (Hipparcos) at the milliarcsecond level. The effects of relativity have been often treated as small corrections added to basically Newtonian developments. But it has become evident that the next generation of instrumentation and theory will require a more comprehensive approach, one that encompasses definitions of such basic concepts as coordinate systems, time scales, and units of measurement in a purely relativistic language. It may remain the case that, for certain applications, relativistic effects can either be ignored or handled as second-order corrections to Newtonian formulas. However, even in such pro-Newtonian cases, the establishment of a self-consistent relativistic framework has benefits – it, at least, allows the physical assumptions and the errors involved to be more clearly spelled out and understood.

In 1991, the IAU made a series of recommendations concerning how the theory of relativity could best be incorporated into positional astronomy. These recommendations and their implications were studied by several working groups in the 1990s and some deficiencies were noted. As a result, a series of new recommendations was proposed and discussed at IAU Colloquium 180 [Johnston *et al.*, 2000]. The new recommendations were passed by the IAU General Assembly in 2000 and they are given for the purpose of reference in appendix C of this book. It is these recommendations that are described briefly in this chapter.

In special relativity, the Newtonian idea of absolute time in all inertial reference systems is replaced by the concept that time runs differently in different inertial systems, in such a way that there is a fundamental speed c that has the same measured value in all of them. Operational realization of the universal invariant speed c is currently based on the speed of light in vacuum since there were no experiments showing violation of special relativity for vacuum electromagnetic fields. In future, the practical measurement of c may be extended to gravity and the other fields that also propagates in space with the fundamental speed. This interesting question is discussed in [Kopeikin, 2004, 2005; Low, 1999].

In both Newtonian physics and special relativity, inertial reference systems are privileged in the sense that physical laws are simple when written in terms of inertial coordinates [1)] In general relativity , however, the structure of spacetime is defined by the Einstein equations of gravitational field, and there are no privileged reference systems. One can use, in principle, any coordinates to model physical processes. For an infinitely small spacetime region around an observer (considered to be a massless point) in free fall, one can introduce a, so-called, locally-inertial reference system where, according to the Einstein's equivalence principle , all physical laws have the same form as in an inertial reference system in special relativity. Such locally inertial reference systems are used to describe observations taken by the point-like observer. In general-relativistic reference systems of finite spatial extent, the geometry of spacetime is defined by a metric tensor, a 4×4 matrix of mathematical expressions, that serves as an operator on two four-vectors. In its simplest application, the metric tensor directly yields the generalized (four-dimensional) distance between two neighboring spacetime events called the interval. The metric tensor (and other geometric objects depending on it) effectively enters the equations through which physics is described in the reference system. For further mathematical details and more precise definitions the reader is referred to Chapters 2 and 3 on special and general relativity.

Time in general relativity can be operationally understood as follows. An observer carries out an idealized clock and travels through spacetime along a timelike worldline. Each point (a spacetime event) on the worldline that he follows, can be characterized by a set of four numbers. These four numbers are the values of the four coordinates in four-dimensional spacetime. For the same worldline in a different coordinate system, the numbers will, in general, be different. Proper time of the observer is defined as the time kept by the observer's idealized clock, in whatever trajectory and gravity field it finds itself. Proper time is always measurable

1) The reader should not confuse the privileged role of the inertial coordinates with the existence of a preferred frame associated with æther. This point is briefly discussed in section 2.1.5.

irrespectively of existence of other observers who measure their own proper times. To measure properties of spacetime, the proper times of the observers must be synchronized. The synchronization is done with exchange of electromagnetic signals stamped by the observers with the proper-time identification tags. To reduce all the data to a single operational time network, proper times are converted to a coordinate time. Coordinate time is one of the four independent coordinates used to characterize a spacetime event. Coordinate time is *not* directly measurable with observer's clock but is connected to the clock's proper time by the spacetime coordinate transformation. On the other hand, the coordinate time of a reference system is the independent argument of the equations of motion of bodies in that reference system. For this reason, it can be measured through continuous astronomical observations of motions of the celestial bodies (planets, Moon, Earth's satellites). Though measurement of the proper time by an ensemble of atomic clocks synchronized through the Global Positioning System (GPS) is the most precise realization of time scale, the coordinate time continues to play a fundamental role in dynamical astronomy of the solar system bodies. The IAU resolutions on relativity passed in 2000 are concerned with two coordinate charts, one barycentric and one geocentric, and the coordinate times used in each one.

9.2.2
The BCRS and the GCRS

In resolution B1.3 of 2000, the IAU defined two coordinate charts for use in astronomy, one with its origin at the solar system barycenter and one with its origin at the geocenter. In current astronomical usage these are referred to as reference systems. The astronomical distinction between *reference systems* and reference frames is discussed in section 9.4. The two systems are the Barycentric Celestial Reference System (BCRS) and the Geocentric Celestial Reference System (GCRS). Harmonic coordinates are recommended for both systems, i.e., the harmonic gauge imposed on the metric tensor, is used. The resolution provides the specific forms of the metric tensors for the two coordinate systems and the four-dimensional transformation between them. The latter would reduce to a Lorentz transformation for a fictitious Earth moving with constant velocity in the absence of gravitational fields. The general forms of the gravitational potentials, which appear in the metric tensors, are also presented. In resolution B1.4, specific expansions of the Earth's gravitational potential in the GCRS are recommended. In resolution B1.5, the relationship between the coordinate time scales for the two reference systems, Barycentric Coordinate Time (TCB), and Geocentric Coordinate Time (TCG), is given. Each of the resolutions is mathematically detailed, and the formulas may be found in the text of the resolutions in the appendix C. For interested readers, the Chapters 4, 5, 6 of this book are highly recommended to understand the mathematical structure of the theory underlying the IAU resolutions. This section makes only very general comments on the BCRS and GCRS, although the time scales TCB and TCG are described in a bit more detail in section 9.3.

The BCRS is a global reference system in which the positions and motions of bodies outside the immediate environment of the Earth are to be expressed. It is the reference system appropriate for the solution of the equations of motion of solar system bodies (that is, the development of solar system ephemerides) and within which the positions and motions of galactic and extragalactic objects are most simply expressed. It is the system to be used for most positional-astronomy reference data, e.g., star catalogs. The GCRS is a local reference system for Earth-based measurements and the solution of the equations of motion of bodies in the near-Earth environment, e.g., artificial satellites. The time-varying position of the Earth's celestial pole is defined within the GCRS (resolution B1.7 of 2000). Precise astronomical observations involve both systems: the instrumental coordinates, boresights, baselines, etc., may be expressed in the GCRS, but in general one wants the astronomical results expressed in the BCRS where they are cataloged and easier to interpret. Thus it is unavoidable that data analysis procedures for precise techniques will involve both GCRS and BCRS quantities and the transformation between them. For example, the basic equation for VLBI delay (the time difference between wavefront arrivals at two antennas) explicitly involves vectors expressed in both systems – antenna-antenna baselines are given in the GCRS, while solar system coordinates and velocities and quasar directions are expressed in the BCRS. Various relativistic factors connect the two kinds of vectors [2)]

In the 2000 resolutions, the coordinate axes of the two reference systems do not have a defined orientation. They are described as kinematically non-rotating, which means that the axes have no systematic rotation with respect to distant objects in the universe, and specifically the radio sources that make up the ICRF – see section 9.4. Since the axis directions are not specified, one interpretation of the 2000 resolutions is that the BCRS and GCRS in effect define families of coordinate systems, the members of which differ only in overall orientation. The IAU Working Group on Nomenclature for Fundamental Astronomy (2003-2006) recommended that the directions of the coordinate axes of the BCRS be understood to be those of the International Celestial Reference System (ICRS) described in section 9.4. That recommendation was formalized into IAU resolution B2 (recommendation 3) of 2006. And, since the transformation between the BCRS and GCRS is specified in the resolutions, the directions of the GCRS axes are also implicitly defined by this understanding. Here are the definitions of the two systems recommended by the working group.

Barycentric Celestial Reference System (BCRS): A system of barycentric spacetime coordinates for the solar system within the framework of general relativity with metric tensor specified by the IAU 2000 resolution B1.3. Formally, the metric tensor of the BCRS does not fix the coordinates completely, leaving the final orientation of the spatial axes undefined. However, for all practical applications, unless otherwise stated, the BCRS is assumed to be oriented according to the ICRS axes.

Geocentric Celestial Reference System (GCRS): A system of geocentric spacetime coordinates within the framework of general relativity with metric

2) Section 7.7.3 discusses definitions and mathematics of VLBI technique.

tensor specified by the IAU 2000 resolution B1.3. The GCRS is defined such that the transformation between BCRS and GCRS spatial coordinates contains no rotation component, so that GCRS is kinematically non-rotating with respect to BCRS. The equations of motion of, for example, an Earth satellite with respect to the GCRS will contain relativistic Coriolis forces that come mainly from geodetic precession. The spatial orientation of the GCRS is derived from that of the BCRS, that is, unless otherwise stated, by the orientation of the ICRS.

Because, according to the last sentence of the GCRS definition, the orientation of the GCRS is determined by that of the BCRS, and therefore the ICRS, in this chapter the GCRS will often be described as the "geocentric ICRS". However, this sentence does not imply that the spatial orientation of the GCRS is the *same* as that of the BCRS (ICRS). The relative orientation of these two systems is embodied in the four-dimensional transformation given in resolution B1.3 of 2000, which are shown in the next section [3], is itself embodied in the algorithms used to compute observable quantities from BCRS (ICRS) reference data. From another perspective, the GCRS is just a rotation (or series of rotations) of the international geodetic system (discussed in section 9.7). The geodetic system rotates with the crust of the Earth, while the GCRS has no systematic rotation relative to extragalactic objects.

The above definition of the GCRS also indicates some of the subtleties involved in defining the spatial orientation of its axes. Without the kinematically non-rotating constraint, the GCRS would have a slow rotation with respect to the BCRS, the largest component of which is called geodetic precession also known as de Sitter-Fokker precession. This rotation, approximately 1.9 arcseconds per century, would be inherent in the GCRS if its axes had been defined as dynamically non-rotating rather than *kinematically* non-rotating. By imposing the latter condition, Coriolis terms must be added (via the inertial parts of the potentials in the metric; see notes to resolution B1.3 of 2000) to the equations of motion of bodies expressed in the GCRS. For example, as mentioned above, the motion of the celestial pole is defined within the GCRS, and geodetic precession appears in the precession-nutation theory rather than in the transformation between the GCRS and BCRS. Other barycentric-geocentric transformation terms that affect the equations of motion of bodies in the GCRS because of the axis-orientation constraint are described in Soffel *et al.* [2003, section 3.3] and Kopeikin and Vlasov [2004, section 6], and their derivation are summarized in Chapter 5 of this book.

9.2.3
Computing Observables

Ultimately, the goal of these theoretical formulations is to facilitate the accurate computation of the values of observable astrometric quantities: transit times(, zenith distances, focal plane coordinates, interferometric delays, etc.) at the time and place of observation, that is, in the proper reference system of the observer (that is called

3) See also Chapter 5.

the proper reference frame in section 2.6. There are some subtleties involved because in Newtonian physics and special relativity, observables are directly expressed in terms of some inertial coordinates, while according to the rules of general relativity, observables must be defined in a coordinate-independent manner allowing to use arbitrary coordinates.

In any case, to obtain observables, there are a number of reductions that must be performed. Typically, one predicts the value of an observable, based on theory and data in hand, then compare this value with what is actually measured; the difference yields information on various parameters of interest. The procedure begins with astrometric reference data: a pre-computed solar system ephemeris and, if a star (optical or radio source) is involved, a star catalog with positions and proper motions listed for a specified epoch. The phenomena listed below and that must be taken into account, as traditionally categorized, are, for stars:

- Proper motion: the 3-D space motion of the star, relative to that of the solar system barycenter, between the catalog epoch and the date of interest. It is assumed linear and computed from the catalog proper motion components, radial velocity, and parallax. Projected onto the sky, the motion amounts to less than 1 arcsecond per year (usually much less) except for a few nearby stars. For components of a binary or multiple system, the motion of the star relative to that of the system's center of mass must also be accounted for.
- Parallax: the change in one's perspective on stars in the solar neighborhood due to the varying position of the Earth in its orbit. Its magnitude in arcseconds is (distance in parsecs)$^{-1}$, and hence is always less than 1 arcsecond for any star.
- Gravitational light bending: the apparent deflection of the light path in the gravitational field of the Sun and (to a much lesser extent) the other planets (see sections 7.5.2 and 7.9.5). Although it reaches 1.75 arcsecond at the limb of the Sun, it falls to 0.05 arcsecond at 10° from the Sun, and amounts to no more than a few milliarcseconds over the hemisphere of the sky opposite the Sun.
- Aberration: the change in the apparent direction of light caused by the observer's velocity with respect to the solar system barycenter (see section 2.5.8). Independent of distance, it is equal approximately to $\beta = V/c$, expressed as an angle. Therefore, it can reach 21 arcseconds for observers on the surface of the Earth (V = 30 km/s), and somewhat more for instruments in orbit.
- Atmospheric refraction: the total angular change in the direction of the light path through the Earth's atmosphere; applies only to an observer on or near the surface of the Earth. The direction of refraction is usually assumed to be parallel to the local vertical and a function only of zenith distance (although these assumptions may not be true in all cases). At optical wavelengths, its magnitude is zero at the zenith, about 1 arcminute at a zenith distance of 45°, and 0.5° at the horizon. Refraction is roughly proportional to the atmospheric pressure at the observer, but it also depends on other atmospheric parameters and the observational wavelength of light.

The same effects are relevant to objects in the solar system, except that the proper motion calculation is replaced by a function that retrieves an object's barycentric position from its ephemeris, as part of an iterative light-time calculation. For the purposes of these calculations, extragalactic objects can be considered in most cases as having zero parallax and proper motion. The star or planet positions computed by considering all these effects obviously depend on the location of the observer; so that an observer on the surface of the Earth will see a slightly different position than a fictitious observer at the geocenter - the effect known as diurnal parallax, the differences being greater for solar system objects, especially nearby ones, reaching about $1°$ for the Moon.

Collectively, these calculations will be referred to in this chapter as the algorithms for proper place of an observed object - a term that refers explicitly to angular variables. The same effects must be accounted for in the time domain for VLBI and pulsar timing. Ranging observations within the solar system (LLR, radar, spacecraft Doppler tracking) involve two-way signals (often with different sending and receiving stations) and are therefore even more complex (see, for example, section 7.8 of this book).

For Earth-based observing systems, one must also account for precession, nutation, Earth rotation, and polar motion, which can be taken together and represented as a single rotation matrix applied at the end of the proper place calculation (for details, see sections 9.6 and 9.7). That allows the object coordinates to be expressed in one of the "equator of date" (now also called "intermediate") coordinate systems.

There are classical expressions for all these effects (except gravitational deflection), and relativity *explicitly* enters the procedure in only a few places, usually as added terms to the classical expressions, and in the formulas that link the various time scales used. In fact, the way these effects are categorized and visualized is based on a classical point of view. It has become common, then, to view this ensemble of calculations as being carried out entirely in a single reference system; or, two reference systems, barycentric and geocentric, that have parallel axes and differ only in the origin of coordinates (that is, they are connected by a Galilean transformation). For example, the coordinate system defined by the "mean equator and equinox of J2000.0", can be thought of as either barycentric or geocentric. The relativistic effects then are interpreted simply as "corrections" to the classical result.

While such a viewpoint may seem aesthetically tidy, it breaks down at high levels of accuracy and for some types of observations. Relativity theory leads to a more correct, albeit more subtle, interpretation for the same set of calculations. It is represented by the BCRS-GCRS paradigm wherein some of the astrometric quantities are expressed relative to the BCRS, and others are defined relative to the GCRS. The two systems are quite different in a number of ways, as described in the previous section as well as in sections 4.3 and 4.5 of this book. The situation is easiest to describe if one restricts the discussion to a fictitious observer at the center of mass of the Earth, that is, to observations referred to the geocenter. The transformation between the two systems is not explicit in the currently-adopted algorithms, but is embodied in the relativistic terms in the expressions used for aberration or VLBI delay. The distinction between the two systems is most obvious in the formulation for angular variables. There, the algorithms for space motion, parallax, light-time, and gravitational deflection [4] all use vectors expressed in the BCRS star catalogs(and solar system ephemerides are inherently BCRS), while the series for precession, nutation, Earth rotation, and polar motion (if applied in that order) starts with vectors expressed in the GCRS. In essence, the aberration calculation connects the two systems because it contains the transformation between them: its input is a pair of vectors in the BCRS and its output is a vector in the GCRS. In the VLBI case, aberration does not appear explicitly [5], but the conventional algorithm for the delay observable [McCarthy and Petit, 2004, Chapter 11] incorporates vectors expressed in both systems, with appropriate conversion factors obtained from the indexBCRS-GCRS transformation given, for example, in chapter 5.2.6 of this book.

For an observer on or near the Earth's surface, the calculations have to include the position and velocity of the observer relative to the geocenter. These are naturally expressed in the GCRS but for some of the calculations (parallax, light-time, gravitational light deflection, and aberration) they must be added to the position and velocity of the geocenter relative to the solar system barycenter, which are expressed in the BCRS. Thus, additional observer-to-GCRS transformation is indicated, although the velocity of the observer in the GCRS (i.e., relative to the geocenter) is sufficiently small that a Galilean transformation (simple vector addition) suffices for current observational accuracy. Correct use of the resulting vectors results in the values of the observables expressed, not in the GCRS, but in the proper (locally-inertial) reference frame of the observer, which is also known in astronomy as the topocentric frame. The reader is advised to read sections 8.3, 8.4, and 8.5 of the present book for the observer-GCRS transformation.

Preceding chapters of this book explore these considerations in detail and provides algorithms for computing observables within the context of general relativity and the IAU recommended reference systems. A shorter exposition of just the calculations needed for high precision angular observations from space is given by Brumberg *et al.* [1990]; Klioner [2003a]; Klioner and Kopeikin [1994].

9.2.4
Other Considerations

The 2000 IAU resolutions on relativity define a framework for future dynamical developments within the context of general relativity. However, there is much unfinished business. The apparently familiar concept of the ecliptic plane has not yet been defined in the context of relativity resolutions. A consistent relativistic theory of Earth rotation is still under development [Klioner *et al.*, 2009b]; the algorithms described in section 9.6 are not such a theory, although they contain all the main relativistic effects and are quite adequate for the current observational precision.

A local reference system similar to the GCRS can be easily constructed for any

4) In the case of a fictitious observer at the geocenter, one neglects the gravity field of the Earth itself in computing gravitational deflection.

5) Part of the expression for VLBI delay, in the time domain, accounts for what would be called aberration in the angular domain; it is possible to compute aberration from the VLBI time-delay algorithm as demonstrated in papers [Kaplan, 1998; Kopeikin, 1990; Kopeikin *et al.*, 1999].

body of an N-body system in exactly the same way as the GCRS, simply by changing the notation in resolutions B1.3-B1.5 of 2000 so that the subscript E denotes a body other than the Earth. In particular, a selenocentric reference system for the Moon plays an important role in lunar laser ranging [Kopeikin *et al.*, 2008; Xie and Kopeikin, 2010].

It is also worth noting that the 2000 resolutions do not describe the proper reference system of the observer – the local, or topocentric, system in which most measurements are actually taken. VLBI(observations are unique in that they exist only after data from various individual antennas are combined; therefore they are referred to the GCRS *ab initio*.) A kinematically non-rotating version of the proper reference system of the observer is just a simplified version of the GCRS: x^i_E should be understood to be the BCRS position of the observer (v^i_E and a^i_E are then the observer's velocity and acceleration) and one should neglect the internal potentials. See papers [Klioner, 2004; Kopejkin, 1991a] and chapter 8.

One final point: the 2000 IAU resolutions as adopted, apply specifically to Einstein's theory of gravity, i.e., the general theory of relativity. The scalar-tensor theory of gravity (see, e.g., Brans and Dicke [1961]; Damour and Esposito-Farese [1992]; Macías *et al.* [2001]) is more general, and possible modification of the IAU 2000 resolutions have been discussed by Kopeikin and Vlasov [2004] and Kopeikin [2010b] with the purpose to incorporate the parameters β and γ of this theory. In the IAU 2000 resolutions, it is assumed that the scalar-tensor theory parameters β and γ are both equal to 1.

One should notice that the parameterized post-Newtonian (PPN) formalism [Will, 1993] also includes parameters $\beta_{\rm PPN}$ and $\gamma_{\rm PPN}$, which should not be confused with β and γ parameters of the scalar-tensor theory of gravity. The difference is vanishing in case of a static, spherically-symmetric gravitational field. However, since the PPN formalism is not a theory, it fails to provide definitive (covariant) answers in a more complicated dynamic situation of N-body problem. An attempt to apply the PPN formalism for extension of the IAU 2000 resolutions was undertaken by Klioner and Soffel [2000] but the results are not consistent with the scalar-tensor theory for the above-said reason (see pages 313-314 of the paper by Kopeikin and Vlasov [2004] for further details).

In 2006, IAU Commission 52, "Relativity in Fundamental Astronomy", was established to pursue all these and similar issues.

9.3 Time Scales

Relevant IAU resolutions:

- A4.III, A4.IV, A4.V, A4.VI of 1991;
- C7 of 1994;
- B1.3, B1.5, B1.7, B1.8, B1.9, and B2 of 2000;
- B3 of 2006.

The IAU has not established any new time scales since 1991, but more recent IAU resolutions have redefined or clarified those already in use, with no loss of continuity. There are, at least, four major classes (groups) of time scales used in astronomy. They are based on: (1) the atomic SI second, (2) the rotation of the Earth, (3) the rotation of pulsars, and (4) the orbital motion of astronomical bodies (planets, pulsars). IAU accepts currently the atomic time and the time scale based on the rotation of the Earth. Orbital motion of the solar system bodies was used as a basis of ephemeris time (ET) but it was abandoned because of inability to compete with the atomic standards. Pulsar time scales based on the rotation and orbital motion of pulsars are maintained by various pulsar timing groups around the world but these time scales are not used (at least now) as time references for practical applications.

The SI second has a simple definition that allows it to be used (in practice and in theory) in any reference system. Time scales based on the SI second include TAI and TT for practical applications, and TCG and TCB for theoretical developments. The latter are to be used for relativistically correct dynamical theories in the geocentric and barycentric reference systems, respectively. Closely related to these are two time scales, TDB and T_{eph}, are implemented in the current generation of the solar system ephemerides. The dynamical time scales used to be called the ephemeris time, and they were widely used in astronomical practice before the advent of atomic clocks [Guinot and Seidelmann, 1988].

Time scales based on the rotation of the Earth include mean and apparent sidereal time and UT1. Because of irregularities in the Earth's rotation, and its tidal deceleration, Earth-rotation-based time scales do not advance at a uniform rate, and they increasingly lag behind the SI-second-based time scales. UT1 is now defined to be a linear function of a quantity called the Earth Rotation Angle, θ. In the formula for mean sidereal time, θ now constitutes the "fast term". The widely disseminated time scale UTC is a hybrid: it advances by SI seconds but is subject to one-second corrections called leap seconds, to keep it within $0\overset{s}{.}9$ of UT1. That procedure is now the subject of debate and there is a movement to eliminate leap seconds from UTC to keep its rate uniform.

Pulsar time scales [Foster and Backer, 1990; Ilyasov *et al.*, 1998; Kopeikin, 1997b; Petit and Tavella, 1996; Taylor, 1991] based on the intrinsic rotation (PT) of pulsars and their orbital motion in binary systems (BPT) are not practically-accepted time-keeping references. The study of stability and accuracy of PT and BPT scales continues at radio astronomical observatories as the pulsar timing data is accumulating. Current state of the art of the pulsar time scale indicates that it might be useful to maintain the uniformity of atomic time scales on a very long time interval exceeding a decade [Guinot and Petit, 1991; Hobbs, 2010; Kopeikin, 1999; Manchester, 2008] and for possible detection of low-frequency gravitational waves [Hobbs, 2008; Kopeikin, 1997a; Kramer and Wex, 2009; Sazhin, 1978; Verbiest *et al.*, 2009].

9.3.1
Different Flavors of Time

The phrase *time scale* is used quite freely in astronomical contexts, but there is sufficient confusion surrounding astronomical times scales that it is worthwhile revisiting the basic concept. A time scale is simply a well defined way of measuring time based on a specific periodic natural phenomenon. The definition of a time scale must provide a description of the phenomenon to be used (what defines a period, and under what conditions), the rate of advance (how many time units correspond to the natural period), and an initial epoch (the time reading at some identifiable event). For example, one could define a time scale where the swing of a certain kind of pendulum, in vacuum at sea level, defines one second, and where the time 00:00:00 corresponds to the transit of a specified star across a certain geographic meridian on an agreed-upon date.

As used in astronomy, a time scale is an idealization, a set of specifications written on a piece of paper. The instruments called clocks, no matter how sophisticated or accurate, provide some imperfect approximation to the time scale they are meant to represent. In this sense, time scales are similar to spatial reference systems (see section 9.4), which have precise definitions but various imperfect realizations. The parallels are not coincidental, since for modern high-precision applications one actually uses spacetime reference systems in the framework of general relativity (see section 9.2). All time scales get therefore associated with specific reference systems.

Operationally, two fundamentally different groups of time scales are currently used in the solar system ephemeris astronomy. The first group of time scales is based on the "atomic" second that is defined as part of the the Système International (SI), and the second group is based on the rotation of the Earth. The SI second is defined as $9,192,631,770$ cycles of the radiation corresponding to the ground state hyperfine transition of Cesium 133 [Göbel *et al.*, 2006], and provides a very precise and constant rate of time measurement, at least for observers local to the apparatus in which such seconds are counted. The rotation of the Earth defines the unit of time called length of day, and it is quite a different basis for time, since it is variable and has unpredictable components. It must be continuously monitored through astronomical observations, now done primarily with very long baseline [radio] interferometry (VLBI). The SI-based time scales are relatively new in the history of time-keeping metrology, since they rely on atomic clocks first put into regular use in the 1950s. Before that, all time scales were tied to the rotation of the Earth. Crystal oscillator clocks in the 1930s were the first artificial time-keeping mechanisms to exceed the accuracy of the Earth itself but they were quickly superseded by atomic clocks and had been used in time-keeping metrology only for a short while. As one will see, the ubiquitous use of SI-based time for modern applications has led to a conundrum about what the relationship between the two kinds of time should be in the future. Both kinds of time scales can be further subdivided into those that are represented by actual clock systems and those that are idealized theoretical constructs including the geometrodynamical clock by Marzke and Wheeler [1964].

General reviews of astronomical time scales are given in [Guinot, 1979, 1986; Guinot and Seidelmann, 1988; Seidelmann and Fukushima, 1992] and Chapter 2 of the Explanatory Supplement to the Astronomical Almanac [Seidelmann and Urban, 2010].

9.3.2 Time Scales Based on the SI Second

Let us first consider the time scales based on the SI second. As a simple count of cycles of microwave radiation from a specific atomic transition, the SI second can be implemented, at least in principle, by an observer anywhere. Thus, SI-based proper time scales can be constructed or hypothesized on the surface of the Earth, on other celestial bodies, on spacecraft, or at theoretically interesting locations in space, such as the solar system barycenter. According to relativity theory, clocks advancing by SI seconds as measured by a local, comoving observer may not appear to advance by SI seconds to an observer on another spacetime trajectory. In general, there will be an observed difference in rate and possibly higher-order or periodic differences, depending on the relative trajectory of the clocks and the remote observer and the gravitational fields involved. The precise conversion formulas [Brumberg, 1991; Brumberg and Kopeikin, 1990; Harada and Fukushima, 2003; McCarthy and Petit, 2004] can be mathematically complex, involving the positions and velocities not just of the clock and observer but also those of an ensemble of massive bodies (Earth, Sun, Moon, planets).

These considerations also apply to coordinate time scales established for specific reference systems in the context of general relativity. However, the definition of an "SI second" as something that is measurable, strictly applies only to proper time scales; it may be better to say that coordinate time scales are expressed in "SI-induced seconds" [Klioner *et al.*, 2009a]. This is a metrological and semantic distinction without any computational consequences. The time-scale conversions are taken from the general four-dimensional spacetime transformation between the reference systems given by relativity theory as described in sections 9.2 and 5.

Two SI-second-based times have already been mentioned in section 9.2: these are the coordinate time scales for theoretical developments based on the Barycentric Celestial Reference System (BCRS) or the Geocentric Celestial Reference System (GCRS). These time scales are called, respectively, Barycentric Coordinate Time (TCB) and Geocentric Coordinate Time (TCG). With respect to a time scale based on SI second as measured on the surface of the Earth, TCG advances at a rate 6.97×10^{-10} faster, while TCB advances at a rate 1.55×10^{-8} faster. TCB and TCG are not likely to come into common use for practical applications, but they appear as the independent arguments [6] for theoretical developments in dynamical astronomy (e.g., Moisson and Bretagnon [2001]). However, none of the current IAU recommended models used in the analysis of astrometric data use TCB or TCG as a basis. This simply reflects the fact that there has not been enough time or motivation for a new generation of dynamical models to be fully developed within the IAU-recommended relativistic paradigm.

For practical applications on or near the surface of the Earth, International Atomic Time (TAI) is a commonly used time scale. It is based on the SI second on the Earth's surface at sea level (specifically, the rotating geoid). TAI is the most precisely determined time scale that is now available for astronomical use. This scale results from analyses by the Bureau International des Poids et Mesures (BIPM) in Sèvres, France, of data from atomic time standards of many countries, according to an agreed-upon algorithm. Although TAI was not officially introduced until 1972, atomic time scales have been available since 1956, and TAI may be extrapolated backwards to the period 1956-1971 [7]. An interesting discussion of whether TAI should be considered a *coordinate time* or a kind of modified *proper time* [8] in the context of general relativity has been given by Guinot [1986]. In any event, TAI is readily available as an integral number of seconds offset from UTC, which is extensively disseminated; UTC is discussed at the end of this section. The TAI offset from UTC is designated ΔAT = TAI-UTC. For example, in 2009 and 2010, ΔAT = 34 s. ΔAT increases by 1 s whenever a positive *leap second* is introduced into UTC (see below). The history of ΔAT values can be found on page K9 of each issue of *The Astronomical Almanac* [2010] and the current value can be found at the beginning of Bulletin A of the Central Bureau of the IERS [2010] and Bulletins A & B of the IERS Rapid Service/Prediction Centre [2010].

The astronomical time scale called Terrestrial Time (TT), used widely for geocentric and topocentric ephemerides such as in *The Astronomical Almanac* [2010], is defined to run at a rate of $(1 - \mathrm{L}_G)$ times that of TCG, where the rate factor $\mathrm{L}_G = 6.969290134\times10^{-10}$. The rate factor applied to TCG to create TT means that TT runs at the same rate as a time scale based on SI seconds on the surface of the Earth. L_G is now considered a defining constant, not subject to further revision. Since TCG is a theoretical time scale that is not kept by any real clock, for practical purposes, TT can be considered an idealized form of TAI with an epoch offset: TT = TAI + $32\overset{\mathrm{s}}{.}184$. This expression for TT preserves continuity with previously-used (now obsolete) "dynamical" time scales, Terrestrial Dynamical Time (TDT) and Ephemeris Time (ET). That is, ET $\rightarrow$ TDT $\rightarrow$ TT can be considered as a single continuous time scale.

Important Note: The "standard epoch" for modern astrometric reference data, designated J2000.0, is expressed as a TT instant: J2000.0 is 2000 January 1, 12^{h} TT (JD 2451545.0 TT) at the geocenter.

6) They are denoted respectively as t and u in Chapters 4-8 of this book.
7) See Nelson *et al.* [2001] for a history of TAI.
8) These terms are described in section 9.2, p. 696.

The fundamental solar system ephemerides from the Jet Propulsion Laboratory (JPL) that are the basis for many of the tabulations in *The Astronomical Almanac* [2010] and other national almanacs were computed in a barycentric reference system and are distributed with the independent argument being a coordinate time scale called T_{eph} (section 9.5 describes the JPL ephemerides). T_{eph} differs in rate (by about 10^{-8} at the geocenter) from that of TCB, the IAU recommended time scale for barycentric developments; the rate of T_{eph} matches that of TT, on average, over the time span of the ephemerides. For many years, T_{eph} was treated as being functionally equivalent to Barycentric Dynamical Time (TDB), defined by the IAU in 1976 and 1979. Both are meant to be "time scales for equations of motion referred to the barycenter of the solar system" yet (loosely speaking) match TT in average rate. The original IAU definition of TDB specified that "there be only periodic variations" with respect to what is now called TT; however, Brumberg and Kopeikin [1990] and Standish [1998b] pointed out that this condition cannot be rigorously fulfilled in practice. Therefore, IAU resolution B3 of 2006 redefined TDB to be a linear function of TCB that is equivalent to T_{eph} for the JPL DE405 ephemeris. Barycentric and heliocentric data derived from the JPL ephemerides are often tabulated with TDB shown as the time argument in *The Astronomical Almanac* [2010], and TDB is the specified time argument for many of the equations presented in this section.

The time scales TCB and TCG are the "natural" coordinate times for the BCRS and GCRS reference systems as explained in sections 4.3 and 4.5, while TDB (or T_{eph}) and TT are defined to be linear functions of TCB and TCG, respectively. TDB (or T_{eph}) and TT are, therefore, sometimes referred to as being "scaled" – that is, the rate has been adjusted for convenience – whereas TCB and TCG are "unscaled". An advantage of the former is that their long-term rates closely match those of proper time scales of observers on the Earth's surface as well as TAI. An advantage of the latter is that fundamental physical constants have the same numerical values in all reference systems that use unscaled coordinate time. When either TDB or TT is used, there must be a corresponding scaling of the spatial coordinates and constants [Brumberg and Kopejkin, 1989a; Irwin and Fukushima, 1999]. Specifically, because T_{eph} = TDB has a rate offset with respect to TCB, the values of parameters determined from or consistent with the JPL ephemerides will, in general, require scaling to convert them to TCB-compatible quantities. This includes the length of the astronomical unit [Klioner, 2008]. Dimensionless quantities such as mass ratios are unaffected. Fundamental solar system ephemerides with TCB as the independent time argument would not, of course, require such scaling. Lists of astronomical constants now often contain separate "TDB compatible" and "TCB compatible" values.

The problem of defining relativistic time scales in the solar system has been treated by Brumberg and Kopeikin [1990], which is quite general but pre-dates the current terminology. Soffel and Brumberg [1991] and Seidelmann and Fukushima [1992] discuss the necessity of introducing the relativistic time scales in dynamical astronomy operating with a rather simple theoretical arguments. Klioner *et al.* [2009a] describes some of the subtleties related to units of time in the IAU relativistic context and suggests some terminology.

9.3.3
Time Scales Based on the Rotation of the Earth

Time scales that are based on the rotation of the Earth are frequently used in astronomical applications, such as telescope pointing, that depend on the geographic location of the observer. Greenwich sidereal time is the hour angle of the equinox measured with respect to the Greenwich meridian. Local sidereal time is the local hour angle of the equinox, or the Greenwich sidereal time plus the longitude (east positive) of the observer, expressed in time units. Sidereal time appears in two forms, mean and apparent, depending on whether the mean or true equinox is the reference point. The position of the mean equinox is affected only by precession while the true equinox is affected by both precession and nutation. The difference between true and mean sidereal time is the *equation of the equinoxes*, which is a complex periodic function with a maximum amplitude of about 1 s. Of the two forms, apparent sidereal time is more relevant to actual observations, since it includes the effect of nutation. Greenwich (or local) apparent sidereal time can be observationally obtained from the right ascensions of celestial objects transiting the Greenwich (or local) meridian.

Universal Time (UT) is also widely used in astronomy, and now almost always refers to the specific time scale UT1. Historically, Universal Time (formerly, Greenwich Mean Time) has been obtained from Greenwich Sidereal Time using a standard expression. In 2000, the IAU redefined UT1 to be a linear function of the *Earth Rotation Angle*, θ, which is the geocentric angle between two directions in the equatorial plane called, respectively, the Celestial Intermediate Origin (CIO) and the Terrestrial Intermediate Origin (TIO) (resolution B1.8 of 2000[9)]). The TIO rotates with the Earth, while the CIO has no instantaneous rotation around the Earth's axis, so that θ is a direct measure of the Earth's rotational motion: conceptually, $\dot{\theta} = \omega$, the Earth's average angular velocity of rotation. See section 9.7 for a more complete description of these new reference points. The definition of UT1 based on sidereal time is still widely used, but the definition based on θ is becoming more common for precise applications. In fact, the two definitions are equivalent, since the expression for sidereal time as a function of UT1 is itself now based on θ.

Since they are mathematically linked, sidereal time, θ, and UT1 are all affected by variations in the Earth's rate of rotation (length of day), which are unpredictable and must be routinely measured through astronomical observations. The lengths of the sidereal and UT1 seconds are therefore not precisely constant when expressed in a uniform time scale such as TT. The accumulated difference in time measured by a clock keeping SI seconds on the geoid from that measured by the rotation of the Earth is ΔT = TT-UT1. A table of observed and extrapolated values of ΔT is given in *The Astronomical Almanac* [2010] on page K9. The long-term trend for ΔT is to increase gradually because of the tidal deceleration of the Earth's rotation, which causes UT1 to lag increasingly behind TT.

9) In the resolution, these points are called the Celestial Ephemeris Origin (CEO) and the Terrestrial Ephemeris Origin (TEO). The change in terminology to CIO and TIO was adopted at the 2006 IAU General Assembly.

In predicting the precise times of topocentric phenomena, like solar eclipse contacts, both TT and UT1 come into play. Therefore, assumptions have to be made about the value of ΔT at the time of the phenomenon. Alternatively, the circumstances of such phenomena can be expressed in terms of an imaginary system of geographic meridians that rotate uniformly about the Earth's axis (ΔT is assumed zero, so that UT1=TT) rather than with the real Earth; the real value of ΔT then does not need to be known when the predictions are made. The zero-longitude meridian of the uniformly rotating system is called the *ephemeris meridian*. As the time of the phenomenon approaches and the value of ΔT can be estimated with some confidence, the predictions can be related to the real Earth: the uniformly rotating system is 1.002738 ΔT east of the real system of geographic meridians. (The 1.002738 factor converts a UT1 interval to the equivalent Earth Rotation Angle – i.e., the sidereal/solar time ratio.)

9.3.4
Coordinated Universal Time (UTC)

The worldwide system of civil time is based on Coordinated Universal Time (UTC), which is now ubiquitous and tightly synchronized. This is the de facto situation; though many nations' legal codes do not mention UTC specifically. UTC is a hybrid time scale, using the SI second on the geoid as its fundamental unit, but subject to occasional 1-second adjustments to keep it within $0\overset{s}{.}9$ of UT1. Such adjustments, called *leap seconds*, are normally introduced at the end of June or December, when necessary, by international agreement. Tables of the remaining difference, UT1-UTC, for various dates are published by the International Earth Rotation Service (IERS) [IERS Rapid Service/Prediction Centre, 2010]. Both past observations and predictions are available. DUT1, an approximation to UT1-UTC, is transmitted in code with some radio time signals, such as those from WWV. As previously discussed in the context of TAI, the difference ΔAT = TAI-UTC is an integral number of seconds, a number that increases by 1 whenever a (positive) leap second is introduced into UTC. That is, UTC and TAI share the same seconds ticks, they are just labeled differently.

Clearly UT1-UTC and ΔT must be related, since they are both measures of the natural "error" in the Earth's angle of rotation at some date. The relation is ΔT = $32\overset{s}{.}184$ + ΔAT - (UT1-UTC).

For the user, then, UTC, which is widely available from GPS, radio broadcast services, and the Internet, is the practical starting point for computing any of the other time scales described above. For the SI-based time scales, one simply adds the current value of ΔAT to UTC to obtain TAI. TT is then just $32\overset{s}{.}184$ seconds ahead of TAI. The theoretical time scales TCG, TCB, TDB, and T_{eph} can be obtained from TT using the appropriate mathematical formulas. For the time scales based on the rotation of the Earth, one again starts with UTC and add the current value of UT1-UTC to obtain UT1. The various kinds of sidereal time can then be computed from UT1 using standard formulas.

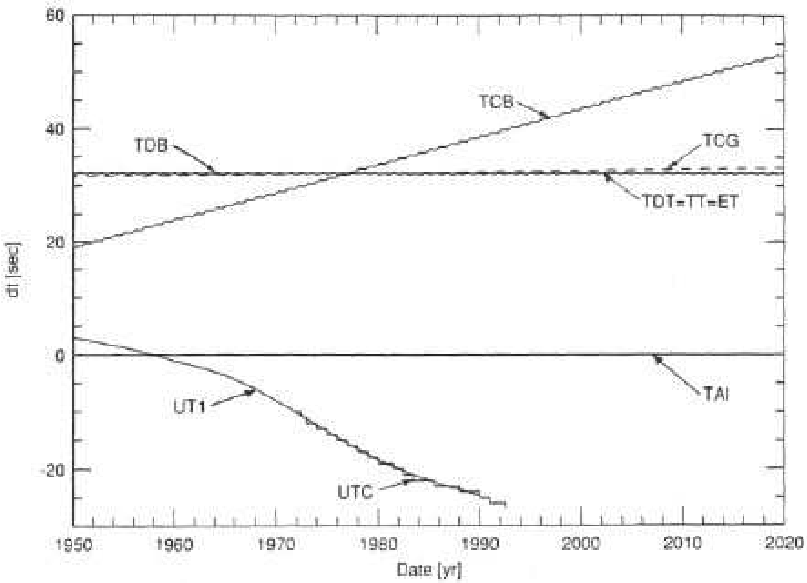

Figure 9.1 Differences in readings of various time scales compared to International Atomic Time (TAI). TT and its predecessors, TDT and ET, are all shown as TAI+32.184 s. The periodic terms of TCB and TDB are exaggerated by a factor of 100. The "stair-step" appearance of UTC is due to the leap seconds inserted into that time scale so that it tracks UT1. TT and the "steps" of UTC are parallel to the TAI line because they are all based on the SI second on the geoid. TDB (or T_{eph}) tracks TT on average over the time span of the specific ephemeris to which it applies. Note the instant at the beginning of 1977 when TT, TCB, and TCG all had the same value. The figure is from [Seidelmann and Fukushima, 1992].

9.3.5
To Leap or Not to Leap

Because of the widespread and increasing use of UTC for applications not considered three decades ago – such as precisely time-tagging electronic fund transfers and other networked business transactions – the addition of leap seconds to UTC at unpredictable intervals creates technical problems and legal issues for service providers. There is now a movement to relax the requirement that UTC remains within 0.9 seconds of UT1. The issue is compounded by the unavoidable scientific fact that the Earth's rotation is slowing due to tidal friction, so that the rate of addition of leap seconds to UTC must inevitably increase. Aside from monthly, annual, and decadal variations, the Earth's angular velocity of rotation is decreasing linearly (because of the tidal friction), which means that the accumulated lag in UT1 increases quadratically; viewed over many centuries, the ΔT curve is roughly a parabola. The formulas for sidereal time, and length of the *ephemeris second* to which the SI second was originally calibrated, are based on the average (assumed fixed) rate of Earth rotation of the mid-1800s [Nelson *et al.*, 2001]. All modern time-keeping systems are ultimately based on what the Earth was doing a century and a half ago!

An IAU Working Group on the Redefinition of Universal Time Coordinated (UTC) was established to consider the leap second problem and recommend a solution, working with the IERS, the International Union of Radio Science (URSI), the Radiocommunication Sector of the International Telecommunications Union (ITU-R), the International Bureau for Weights and Measures (BIPM), and the relevant navigational agencies (IAU resolution B2 of 2000). The IAU Working Group ended its work in 2006 without making a specific recommendation for change. Possibilities that were considered included: using TAI for technical applications instead of UTC; allowing UT1 and UTC to diverge by a larger amount (e.g., 10 or 100 seconds) before a multi-second correction to UTC is made; making a variable correction to UTC at regularly scheduled dates; eliminating the corrections to UTC entirely and allowing UTC and UT1 to drift apart; or changing the definition of the SI second. No solution is ideal (including the status quo) and each of these possibilities has its own problems. For example, if one keeps leap seconds, or a less frequent multi-second correction, can current systems properly time-tag the date and time of an event that occurs *during* the correction? Does a time scale that diverges from UT1 provide a legally acceptable representation of civil time? If corrections are made less frequently, will the possibility of technical blunders increase? If leap seconds are eliminated, won't natural phenomena such as sunrise and sunset eventually fall out of sync with civil time? How does one find all the existing computer code that assumes $|\text{UT1-UTC}| \leq 0.9\,\text{s}$? The matter is now in the hands of the ITU-R, where a working group has proposed eliminating all future leap seconds and letting UTC diverge from UT1. At the end of 2009 it was unclear whether there is the necessary international consensus (within the rules of the ITU) to make this change. In any event, it would take a number of years for any proposed change to take place because of the many institutions and international bodies that would have to be involved.

For scientific instrumentation, the use of TAI – which is free of leap seconds – has much to recommend it. Its seconds can be easily synchronized to those of UTC (only the labels of the seconds are different). It is straightforward to convert from TAI to any of the other time scales. Use of TAI provides an internationally recognized time standard and avoids the need to establish an instrument-specific time scale when continuity of time tags is a requirement.

9.3.6
Formulas

9.3.6.1 *Formulas for Time Scales Based on the SI Second*

For the SI-based time scales, the event tagged 1977 January 1, 00:00:00 TAI (JD 2443144.5 TAI) at the geocenter is special. At that event, the time scales TT, TCG, and TCB all read 1977 January 1, 00:00:32.184 (JD 2443144.5003725). (The $32^{\mathrm{s}}.184$ offset is the estimated difference between TAI and the old Ephemeris Time scale.) This event will be designated t_0 in the following; it can be represented in any of the time scales, and the context will dictate which time scale is appropriate.

From the perspective of a user, the starting point for computing all the time scales is Coordinated Universal Time (UTC). From UTC, one can immediately get International Atomic Time (TAI)

$$\mathrm{TAI} = \mathrm{UTC} + \Delta\mathrm{AT}\,, \tag{9.1}$$

where ΔAT, an integral number of seconds, is the accumulated number of leap seconds applied to UTC.

The astronomical time scale Terrestrial Time (TT) is defined by the epoch t_0 and its IAU-specified rate with respect to Geocentric Coordinate Time (TCG)

$$\frac{d\mathrm{TT}}{d\mathrm{TCG}} = 1 - \mathrm{L}_G\,, \tag{9.2}$$

where $\mathrm{L}_G = 6.969290134\times10^{-10}$ exactly, from which one obtains

$$\mathrm{TT} = \mathrm{TCG} - \mathrm{L}_G\,(\mathrm{TCG} - t_0)\,. \tag{9.3}$$

However, TCG is a theoretical time scale, not kept by any real clock system, so in practice,

$$\mathrm{TT} = \mathrm{TAI} + 32^{\mathrm{s}}.184\,, \tag{9.4}$$

and one obtains TCG from TT.

The relationship between TCG and Barycentric Coordinate Time (TCB) is more complex. TCG and TCB are both coordinate time scales, to be used with the geocentric and barycentric reference systems (the GCRS and BCRS), respectively. The exact formula for the relationship between TCG and TCB is given in resolution B1.5 of 2000, recommendation 2. For a given TCB epoch, one has [10)]

$$\mathrm{TCG} = \mathrm{TCB} - \frac{1}{c^2}\int_{t_0}^{\mathrm{TCB}}\Big(\frac{v_E^2}{2} + U_{ext}(\mathbf{x}_E)\Big)dt - \frac{\boldsymbol{v}_E}{c^2}\cdot(\boldsymbol{x} - \boldsymbol{x}_E) + \cdots\,, \tag{9.5}$$

10) This is equation (5.92) from section 5.2.6 with u =TCG and t =TCB.

where c is the fundamental speed (= speed of light), $\boldsymbol{x}_E$ and $\boldsymbol{v}_E$ are the position and velocity vectors of the Earth's center with respect to the solar system barycenter, and U_{ext} is the Newtonian potential of all solar system bodies apart from the Earth. The integral is carried out in TCB since the positions and motions of the Earth and other solar system bodies are represented (ideally) as functions of TCB. The last term on the right contains the barycentric position vector of the point of interest, $\boldsymbol{x}$, and will be zero for the geocenter, as would normally be the case. The omitted terms are of maximum order c^{-4} as shown in equation (5.92). Note that the transformation is ephemeris-dependent, since it is a function of the time series of $\boldsymbol{x}_E$ and $\boldsymbol{v}_E$ values. The result is a "time ephemeris" [Harada and Fukushima, 2003] associated with every spatial ephemeris of solar system bodies expressed in TCB. It is to be expected that ephemeris developers will supply appropriate time conversion algorithms (software) to allow the positions and motions of solar system bodies to be retrieved for epochs in conventional time scales such as TT or TAI. It is unlikely that ordinary ephemeris users will have to compute equation (9.5) on their own.

The functional form of the above expressions may seem backwards for practical applications; that is, they provide TCG from TCB and TT from TCG. These forms make sense, however, when one considers how an ephemeris of a solar system body (or bodies) or a spacecraft is developed. The equations of motion for the body (or bodies) of interest are expressed in either the barycentric or geocentric system as a function of some independent coordinate time argument. For barycentric equations of motion, expressed in SI units, one wouldbe tempted immediately to identify this time argument with TCB. Actually, however, the association of the time argument with TCB is not automatic; it comes about only when the solution of the equations of motion is made to satisfy the boundary conditions set by the ensemble of real observations of various kinds. Generally, these observations will be time-tagged in UTC, TAI, or TT all of which are based on the SI second on the geoid, and these time tags must be associated with the time argument of the ephemeris. The above formulas can be used to make that association, which then allows the ephemeris to be fit to the observations. More precisely, the spacetime coordinates of the observation events must be transformed to the BCRS. As a consequence, the time argument of the ephemeris becomes a realization of TCB. The fit of the computed ephemeris to observations usually proceeds iteratively, and every iteration of the spatial ephemeris produces a new time ephemeris. With each iteration, the spatial coordinates of the ephemeris become better grounded in reality, as represented by the observations, and the time coordinate becomes a better approximation to TCB. Viewed from this computational perspective, the ephemeris and its time argument are the starting point of the process and the sequence TCB → TCG → TT makes sense.

One can compute an ephemeris and fit it to observations using other formulas for the time scale conversions. A completely valid and precise ephemeris can be constructed in this way, but its independent time argument could not be called TCB. The values of various constants used in, or derived from, such an ephemeris would also not be TCB-based, and a conversion factor would have to be applied to convert them to or from TCB-compatible units. Such is the case with the solar system Development Ephemeris (DE) series from the Jet Propulsion Laboratory. DE405 is now the consensus standard for solar system ephemerides and is described in section 9.5. The DE series dates back to the 1960s, long before TCB and TCG were defined, and its independent time argument is now called $\mathrm{T_{eph}}$. $\mathrm{T_{eph}}$ can be considered to be TCB with a rate factor applied. Or, as mentioned above, $\mathrm{T_{eph}}$ can be considered to be functionally equivalent to the time scale called TDB. Both $\mathrm{T_{eph}}$ and TDB advance, on average, at the same rate as TT. This arrangement makes accessing the DE ephemerides straightforward, since for most purposes, TT can be used as the input argument with little error. The total error in time in using TT as the input argument is <2 ms, which for the geocentric position of the Moon would correspond to an angular error of <1 mas. When more precision is required, the following formula can be used

$$\begin{aligned}
\mathrm{T_{eph}} \approx \mathrm{TDB} \approx \mathrm{TT} \quad &+ \quad 0.001657\ \sin(628.3076\,T + 6.2401)\\
&+ \quad 0.000022\ \sin(575.3385\,T + 4.2970)\\
&+ \quad 0.000014\ \sin(1256.6152\,T + 6.1969)\\
&+ \quad 0.000005\ \sin(606.9777\,T + 4.0212)\\
&+ \quad 0.000005\ \sin(52.9691\,T + 0.4444)\\
&+ \quad 0.000002\ \sin(21.3299\,T + 5.5431)\\
&+ \quad 0.000010\,T\,\sin(628.3076\,T + 4.2490)\ + \cdots\ ,
\end{aligned} \tag{9.6}$$

where the coefficients are in seconds, the angular arguments are in radians, and T is the number of Julian centuries of TT from J2000.0

$$T = \frac{\mathrm{JD(TT)} - 2451545.0}{36525}\ . \tag{9.7}$$

The above is a truncated form of a much longer and more precise series given by Fairhead and Bretagnon [1990]. The maximum error in using the above formula is about 10 μs from 1600 to 2200; that is, its precision is more than two orders of magnitude better than the approximation $\mathrm{T_{eph}}\approx\mathrm{TDB}\approx\mathrm{TT}$. For even more precise applications, the series expansion by Harada and Fukushima [2003] is recommended.

A word of caution: The idea that "$\mathrm{T_{eph}}$ and TDB advance, on average, at the same rate as TT" is problematic. The independent time argument of a barycentric ephemeris (whether considered to be $\mathrm{T_{eph}}$, TDB, or TCB) has a large number of periodic components with respect to TT. Some of the periods are quite long, and may extend beyond the time period of the ephemeris. Thus, the "average rate" of the time argument of the ephemeris, with respect to TT, depends on the averaging method and the time span considered. Differences in rate of some tens of microseconds per century are possible [Fairhead and Bretagnon, 1990]. These rate ambiguities are probably unimportant (amounting to fractional errors of only $\sim 10^{-14}$) for retrieving positions and velocities from the ephemeris but do affect pulsar timing models that are reduced to the barycentric time scale (see section 7.7.2).

9.3.6.2 *Formulas for Time Scales Based on the Rotation of the Earth*

For the time scales that are based on the rotation of the Earth, one again starts with UTC. One has

$$\mathrm{UT1} = \mathrm{UTC} + (\mathrm{UT1\text{-}UTC}) \tag{9.8}$$

$$\approx \mathrm{UTC} + \mathrm{DUT1} \,, \tag{9.9}$$

where DUT1 is a broadcast approximation to UT1-UTC (precision $\pm 0\overset{s}{.}1$). One also has

$$\mathrm{UT1} = \mathrm{TT} - \Delta T \,, \tag{9.10}$$

where $\Delta T = 32\overset{s}{.}184 + \Delta\mathrm{AT} - (\mathrm{UT1\text{-}UTC})$. The most recent values of UT1-UTC and ΔAT are listed in Bulletin A of the Central Bureau of the IERS [2010] and Bulleins A & B of the IERS Rapid Service/Prediction Centre [2010]. Values of ΔT are listed in *The Astronomical Almanac* [2010] on page K9.

The Earth Rotation Angle, θ, is

$$\theta = 0.7790572732640 + 1.00273781191135448\, D_U \,, \tag{9.11}$$

where D_U is the number of UT1 days from 2000 January 1, 12^{h} UT1: $D_U = \mathrm{JD(UT1)} - 2451545.0$. The angle θ is given in terms of rotations (units of 2π radians or 360°). The above rate coefficient gives an Earth rotation period of 86164.0989036903511 seconds of UT1. If one considers this number to be equivalent to the adopted average rotation period of the Earth in SI seconds, it is consistent with the nominal mean angular velocity of Earth rotation, $\omega = 7.292115\times10^{-5}$ radian s^{-1}, used by the International Association of Geodesy. The above expression is taken directly from note 3 to IAU resolution B1.8 of 2000. An equivalent form of this expression (if the integral number of rotations is neglected) that is usually more numerically precise is

$$\theta = 0.7790572732640 + 0.00273781191135448\, D_U + \mathrm{frac[JD(UT1)]} \,, \tag{9.12}$$

where frac[JD(UT1)] is the fractional part of the UT1 Julian date, i.e., JD(UT1) modulus 1.0. Then, Greenwich mean sidereal time (GMST) in seconds is

$$\begin{aligned}\mathrm{GMST} &= 86400 \cdot \theta + (0.014506 + 4612.156534\, T + 1.3915817\, T^2 \\ &- 0.00000044\, T^3 - 0.000029956\, T^4 - 0.0000000368\, T^5)/15 \,,\end{aligned} \tag{9.13}$$

where T is the number of centuries of TDB (equivalently for this purpose, TT) from J2000.0

$$T = \frac{\mathrm{JD(TDB)} - 2451545.0}{36525} \,. \tag{9.14}$$

The polynomial in parentheses is the accumulated precession of the equinox in right ascension, in arcseconds, as given for the P03 solution (equation 42) in paper by Capitaine *et al.* [2003]. Note that two time scales are now required to compute sidereal time: in the "fast term", θ is a function of UT1, while in the remaining terms, T is expressed in TDB (or TT).

To obtain Greenwich apparent sidereal time (GAST) in seconds, one must add the *equation of the equinoxes*

$$\mathrm{GAST} = \mathrm{GMST} + \mathcal{E}_\Upsilon/15\,, \tag{9.15}$$

which accounts for the motion of the equinox due to nutation. An extended series is now used for the equation of the equinoxes. The new series includes so-called *complementary terms* and more fully accounts for the accumulated effect of combined precession and nutation on the position of the equinox. The equation of the equinoxes in arcseconds is

$$\begin{aligned}
\mathcal{E}_\Upsilon = \quad & \Delta\psi \cos\epsilon \\
+ \quad & 0.00264096 \,\sin(\Omega) \\
+ \quad & 0.00006352 \,\sin(2\Omega) \\
+ \quad & 0.00001175 \,\sin(2F - 2D + 3\Omega) \\
+ \quad & 0.00001121 \,\sin(2F - 2D + \Omega) \\
- \quad & 0.00000455 \,\sin(2F - 2D + 2\Omega) \\
+ \quad & 0.00000202 \,\sin(2F + 3\Omega) \\
+ \quad & 0.00000198 \,\sin(2F + \Omega) \\
- \quad & 0.00000172 \,\sin(3\Omega) \\
- \quad & 0.00000087\, T \,\sin(\Omega) + \cdots \,,
\end{aligned} \tag{9.16}$$

where $\Delta\psi$ is the nutation in longitude, in arcseconds; ϵ is the mean obliquity of the ecliptic ; and F, D, and Ω are fundamental luni-solar arguments. All of these quantities are functions of TDB (or TT); see section 9.6 for expressions (especially equations (9.42), (9.46), & (9.50)). The above series is a truncated form of a longer series given in the IERS Conventions [McCarthy and Petit, 2004], but should be adequate for almost all practical applications.

Local mean sidereal time (LMST) and local apparent sidereal time (LAST) in seconds can then be computed respectively from

$$\mathrm{LMST} = \mathrm{GMST} + \left(\frac{3600}{15}\right)\lambda \tag{9.17}$$

$$\mathrm{LAST} = \mathrm{GAST} + \left(\frac{3600}{15}\right)\lambda\,, \tag{9.18}$$

where λ is the longitude of the place of interest, in degrees, positive for places east of Greenwich.

In the above, "Greenwich" actually refers to a plane containing the geocenter, the Celestial Intermediate Pole (CIP), and the point called the Terrestrial Intermediate Origin (TIO). These concepts are described in sections 9.6 and 9.7. Loosely, the CIP is the rotational pole, defined by the precession and nutation theories. For astronomical purposes, the TIO can be considered to be a point on the rotational equator (the plane orthogonal to the CIP) essentially fixed at geodetic longitude 0. Strictly, then, the longitude λ should be measured around the axis of the CIP from the TIO to the location of interest. Because of polar motion, the pole of the conventional system of geodetic coordinates is not at the CIP so the longitude needed, is not quite the same as the geodetic longitude. The longitude, in degrees, to be used in equation (9.18) is

$$\lambda = \lambda_G + (x_p \sin\lambda_G + y_p \cos\lambda_G) \tan\phi_G \;/3600 \,, \tag{9.19}$$

where λ_G and ϕ_G are the usual geodetic longitude and latitude of the place, with λ_G in degrees (north latitudes and east longitudes are positive); and x_p and y_p are the coordinates of the pole (CIP) with respect to the geodetic system, in arcseconds (x_p and y_p can be a few tenths of an arcsecond). The geodetic system is formally the International Terrestrial Reference System (ITRS), which matches WGS-84 system(available from GPS) to several centimeters. The local meridian assumed by the formula for LAST, using the longitude λ, passes through the local zenith (orthogonal to the local surface of the WGS-84 ellipsoid) and the north and south celestial poles – close to but not through the local geodetic north and south points. This is the meridian that all stars with apparent topocentric right ascension equal to LAST will pass over at time UT1. More information can be found in sections 9.7.4, 9.7.5.5, and 9.7.5.7.

The above formulas are entirely geometric. Not described are astronomical latitude and longitude, which are based on the local direction of gravity. Astronomical latitude and longitude are affected by the deflection of the vertical caused by permanent gravitational anomalies and, at a much lower level, semidiurnal tides. Astronomical latitude and longitude must be corrected for such effects to obtain geodetic latitude and longitude.

9.4 The Fundamental Celestial Reference System

Reference data for positional astronomy, such as the data in astrometric star catalogs or barycentric planetary ephemerides, are now specified within the International Celestial Reference System (ICRS). The ICRS is a coordinate system whose origin is at the solar system barycenter and whose axis directions are effectively defined by the adopted coordinates of an ensemble of extragalactic radio sources determined by very long-baseline radio interferometry (VLBI). These radio sources (quasars and active galactic nuclei (AGN)) are assumed to have no observable intrinsic angular motions [11]. Thus, the ICRS is a "space-fixed" system (more precisely, a kinematically non-rotating coordinate system) without an associated epoch. However, the ICRS closely matches the conventional dynamical system defined by the Earth's mean equator and equinox of J2000.0; the alignment difference is at the 0.02 arcsecond level, negligible for many applications. The ICRS replaced the FK5 as the IAU-recommended fundamental astronomical reference system in 1998.

Strictly speaking, the ICRS is somewhat of an abstraction, a set of coordinate directions that perfectly satisfies a list of theoretical criteria. The list of radio source positions that define it for practical purposes is called the International Celestial Reference Frame (ICRF), now in its second generation [Ma *et al.*, 1998, 2009]. In the terminology that is now commonly used, a *reference system* like the ICRS is "realized" by a *reference frame* like the ICRF, and there can be more than one such realization. In the case of the ICRS, there is, in fact, a lower-accuracy realization for work at optical wavelengths, called the Hipparcos Celestial Reference Frame (HCRF). The HCRF is composed of the positions and proper motions of the astrometrically "well-behaved" stars in the Hipparcos catalog.

Astrometric data referred to the ICRS are becoming more common, with new catalogs now available in the optical, infrared, and radio.

The ICRS defines the spatial orientation of the Barycentric Celestial Reference System, which incorporates the relativistic metric specified in IAU resolution B1.3 of 2000 for solar system barycentric coordinate systems. The BCRS metric is explained in more detail in section 4.3.2 of this book.

9.4.1 The ICRS, the ICRF, and the HCRF

The fundamental celestial reference system for astronomical applications is now the International Celestial Reference System (ICRS), as provided in IAU resolution B2 of 1997. The ICRS is a coordinate system with its origin at the solar system barycenter and axis directions that are fixed with respect to distant objects in the universe; it is to be used to express the fundamental positions and motions of stars, planets, and other celestial objects. Unlike geocentric or topocentric reference systems that

11) This assumption is subsequently checked and the sources which do not satisfy it, are excluded from ICRS [Hughes *et al.*, 1991; Jacobs *et al.*, 2003; Ma *et al.*, 2009].

are used for expressing the apparent directions of celestial objects for observational purposes, the ICRS is a system in which the equations of motion of these objects are relatively simple (it is close to being a locally inertial system), and is, therefore, suitable for basic reference data, such as star catalogs or barycentric planetary ephemerides derived from numerical integration. The ICRS has replaced the previous FK5 system but is consistent with it to within the FK5's errors.

The specifications for what became the ICRS were established by IAU working groups in the late 1980s and early 1990s, and most particularly in IAU resolution A2 of 1991 (Recommendations VI and VII). These specifications can be summarized as follows:

- Origin is at solar system barycenter;
- Defining objects are unresolved, stationary, and stable extragalactic radio sources (i.e., with minimal structural variations), including a large proportion with "well identified optical counterparts";
- Positions of defining objects are determined by long baseline radio interferometry;
- Direction of axes are fixed in space – do not rotate with respect to defining objects;
- To be "as near as possible" to system defined by the Earth mean equator and equinox of J2000.0 and to the FK5;
- Independent of time and specific realizations (if the objects that define it, eventually are changed).

The relativistic basis of the ICRS is that of the Barycentric Celestial Reference System (BCRS) as defined by IAU resolution B1.3 of 2000, which is described in sections 4.3 and 9.2 of this book. Strictly speaking, according to IAU resolution B2 (recommendation 2) of 2006, the ICRS is the spatial orientation of the BCRS, and as such its axes are *kinematically non-rotating* (see section 4.3.1). Because the ICRS and BCRS are so closely linked, the two terms are often used interchangeably. The two terms refer to different aspects of the fundamental barycentric reference system and are not, therefore, independent. A simple way of understanding the connection is that BCRS vectors are expressed with respect to the ICRS spatial axes and ICRS data are based on the BCRS metric.

To establish the ICRS as a practical system, in 1997 the IAU specified a particular set of distant benchmark objects, observable at radio wavelengths, whose adopted coordinates effectively define the directions of the ICRS axes. This first "realization" of the ICRS, called the International Celestial Reference Frame (ICRF or, now, ICRF1), is a set of high accuracy positions of extragalactic radio sources measured by Very Long Baseline Interferometry (VLBI) [Ma *et al.*, 1998; Ma and Feissel, 1997]. In 2009, the IAU replaced the original list of sources and coordinates with a revised and enlarged list [Ma *et al.*, 2009], referred to as ICRF2 (IAU resolution B3 of 2009). The ICRS is realized at optical wavelengths – but at lower accuracy – by the Hipparcos Celestial Reference Frame (HCRF), consisting of the *Hipparcos Catalogue* [Perryman and ESA, 1997] of star positions and motions, with certain exclusions (IAU resolution B1.2 of 2000). The areal density of the two ICRS realizations is quite different: the ICRF2 consists of about 3400 radio sources, while the HCRF contains about 100,000 stars. The coordinates of the ICRF radio sources and HCRF stars are given relative to the ICRS origin at the solar system barycenter, and a number of transformations are required to obtain the coordinates that would be observed from a given location on or near the Earth at a specific date and time. Observations taken from or near the Earth are expressed in the Geocentric Celestial Reference System (GCRS). The GCRS is briefly described in section and 9.2.2, as is the transformation between the BCRS and GCRS. Detailed description of GCRS is given in sections 4.5 and 8.3. The standard algorithms for computing geocentric "apparent places" of stars or planets incorporate the BCRS-to-GCRS transformation given in section 5.2.6.

Although the directions of the ICRS coordinate axes are not defined by the kinematics of the Earth, the ICRS axes (as implemented by the ICRF1, ICRF2, and HCRF) closely approximate the axes that would be defined by the mean Earth equator and equinox of J2000.0 (to within about 0.02 arcsecond), if the latter is considered to be a barycentric system. Because the ICRS axes are meant to be "space fixed", i.e., kinematically non-rotating, there is no date associated with the ICRS. Furthermore, since the defining radio sources are assumed to be so distant that their angular motions, seen from Earth, are negligible, there is no epoch associated with the ICRF. It is technically incorrect, then, to say that the ICRS is a "J2000.0 system", even though for many current data sources, the directions in space defined by the equator and equinox of J2000.0 and the ICRS axes are the same to within the uncertainties of the data.

The IAU Working Group on Nomenclature for Fundamental Astronomy, which worked from 2003 to 2006, recommended the following definitions for the ICRS and ICRF:

International Celestial Reference System (ICRS): The idealized barycentric coordinate system to which celestial positions are referred. It is kinematically non-rotating with respect to the ensemble of distant extragalactic objects. It has no intrinsic orientation, but was aligned close to the mean equator and dynamical equinox of J2000.0 for continuity with previous fundamental reference systems. Its orientation is independent of epoch, ecliptic or equator and is realized by a list of adopted coordinates of extragalactic sources.

International Celestial Reference Frame (ICRF): A set of extragalactic objects whose adopted positions and uncertainties realize the ICRS axes and give the uncertainties of the axes. It is also the name of the radio catalog whose 212 defining sources are currently the most accurate realization of the ICRS. Note that the orientation of the ICRF catalog was carried over from earlier IERS radio catalogs and was within the errors of the standard stellar and dynamic frames at the time of adoption. Successive revisions of the ICRF are intended to minimize rotation from its original orientation. Other realizations of the ICRS have specific names (e.g., Hipparcos Celestial Reference Frame).

As previously noted, the ICRF has been superseded by the ICRF2; there are now 295 defining sources.

9.4.2
Background: Reference Systems and Reference Frames

The terminology that has become standard over the past several decades distinguishes between a *reference system* and a *reference frame*. A reference system is the complete theoretical specification of how a celestial coordinate system is to be formed. Both the origin and the orientation of the fundamental planes (or axes) are defined. A reference system also incorporates a specification of the fundamental models needed to construct the system; that is, the basis for the algorithms used to transform between observable quantities and reference data in the system. A reference frame, on the other hand, consists of a set of identifiable fiducial points on the sky along with their coordinates, which serve as the practical realization of a reference system.

For example, the fundamental plane of an astronomical reference system has conventionally been the extension of the Earth's equatorial plane, at some date, to infinity. *Declination* is the angular distance north or south of this plane, and *right ascension* is the angular distance measured eastward along the equator from some defined reference point (see figure 9.2). This reference point, the right ascension origin, has traditionally been the *equinox*: the point at which the Sun, in its yearly circuit of the celestial sphere, crosses the equatorial plane moving from south to north. The Sun's apparent yearly motion lies in the *ecliptic*, the plane of the Earth's orbit. The equinox, therefore, is a direction in space along the nodal line defined by the intersection of the ecliptic and equatorial planes; equivalently, on the celestial sphere, the equinox is at one of the two intersections of the great circles representing these planes. Because both of these planes are moving, the coordinate systems that they define must have a date associated with them; such a reference system must therefore be specified as "the equator and equinox of (some) date".

Of course, such a reference system is an idealization, because the theories of motion of the Earth that define how the two planes move are imperfect. In fact, the very definitions of these planes are problematic for high precision work. Even if the fundamental planes of a reference system are defined without any reference to the motions of the Earth, there is no way magically to paint them on the celestial sphere at any particular time. Therefore, in practice, one uses a specific reference frame – a set of fiducial objects with assigned coordinates – as the practical representation of an astronomical reference system. The scheme is completely analogous to how terrestrial reference systems are established using survey control stations (geodetic reference points) on the Earth's surface.

Most commonly, a reference frame consists of a catalog of precise positions (and motions, if measurable) of stars or extragalactic objects, as seen from the solar system barycenter at a specific epoch (now usually "J2000.0", which is 12^{h} TT on 1 January 2000). Each object's instantaneous position, expressed as right ascension and declination, indicates the object's angular distance from the catalog's equator and origin of right ascension. (A catalog's right ascension origin was formerly referred to as the *catalog equinox*, a now-obsolete term.) Therefore, any two such objects in the catalog – if they are not coincident or antipodal – uniquely orient a spherical

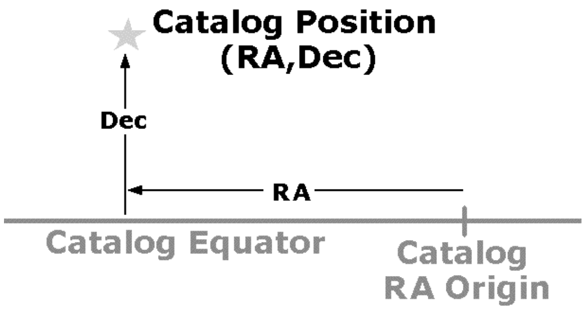

Figure 9.2 Schematic of the celestial coordinates of a catalog object. Celestial coordinates of objects in a catalog are expressed with respect to the catalog equator and right ascension origin, which generally are close approximations to the mean Earth equator and equinox at some date (e.g., J2000.0). Conversely, the tabulated coordinates of all the objects in a catalog define, in an average sense, the location of the catalog equator and equinox on the sky.

coordinate system on the sky – a reference frame.

To be useful, a reference frame must be implemented at the time of actual observations, and this requires the computation of the apparent coordinates of the catalog objects at arbitrary dates and times (see section 9.2.3). Astrometric star catalogs list *proper motions*, which are the projection of each star's space motion onto the celestial sphere, expressed as an angular rate in right ascension and declination per unit time (usually now milliarcseconds/year). For stars or other galactic objects, the proper motions allow the reference frame to be established at epochs other than that of the positions listed in the catalog.

The positions of solar system objects can also be used to define a reference frame. For each solar system body involved, an *ephemeris* (plural *ephemerides*) is used, which is simply a table of the celestial coordinates of the body as a function of time (or an algorithm that yields such a table). A reference frame defined by the ephemerides of one or more solar system bodies is called a *dynamical reference frame*. Because the ephemerides used incorporate the motion of the Earth as well as that of the other solar system bodies, dynamical reference frames embody in a very fundamental way the moving equator and ecliptic, hence the equinox. They have, therefore, been used to correct the orientation of star catalog reference frames (the star positions were systematically adjusted) on the basis of simultaneous observations of stars and planets. In a sense, the solar system is used in this way as a gyrocompass with planets playing a role of a "rotating wheel" maintaining the precise orientation in the inertial space with respect to "stars". However, dynamical reference frames are not very practical for establishing a coordinate system for day-to-day astronomical observations, because the sky density of solar system objects is low, many are resolved, and they tend to be clustered near the ecliptic plane.

Descriptions of reference frames and reference systems often refer to three coordinate axes, which are simply the set of right-handed Cartesian axes that correspond to the usual celestial spherical coordinate system. The xy-plane is the equator, the z-axis points toward the north celestial pole, and the x-axis points toward the origin of right ascension. Although, in principle, this allows us to specify the position of any celestial object in rectangular coordinates, the distance scale (based on stellar parallaxes) is not established to high precision beyond the solar system. What a reference system actually defines is the way in which the two conventional astronomical *angular* coordinates - right ascension and declination - relate to real observable points in the sky. See equations (9.24) & (9.25) for the conversion between rectangular and spherical celestial coordinates.

9.4.3 The Effect of Catalog Errors on Reference Frames

A modern astrometric catalog contains data on a huge number of celestial objects (N), so the coordinate system it defines, is vastly over-determined. The quality of the reference frame provided by a catalog depends on the extent to which the coordinates of all possible pairs of the objects ($\approx N^2/2$) serve to define the identical equator and right ascension origin, within the expected random errors. Of course, in practice, the catalog's equator and right ascension origin are effectively defined by averages over the objects that are used for a specific application. Therefore, random errors in the coordinates of the selected objects limit the accuracy with which the reference frame approximates the intended reference system for the particular application; generally, the accuracy will be proportional to the square root of the number of objects used.

However, the object' positions listed in catalogs also may contain systematic errors, which are errors in the coordinates that are similar for objects that are in the same area of the sky, or are of the same magnitude (flux) or color (spectral index). Systematic errors in position mean that the reference frame is warped (distorted), or is effectively different for different classes of objects. Systematic errors do not obey square-root-of-N statistics, and may be unrecognized, so they tend to be more pernicious in effect than random errors. Obviously, minimizing systematic errors, when a catalog is constructed, is at least as important as minimizing the random errors, although much more difficult.

Typical applications of reference frames involve sets of astronomical observations taken at various dates. To establish a reference frame based on stars or other galactic objects at the dates of the observations, the proper motions in the catalog must be used to update the positions. The accuracy of the proper motions in the catalog largely determines how well the reference frame can be established for dates that are many years beyond (or prior to) the catalog's epoch. Because the tabulated proper motions are never perfect, any celestial reference frame based on galactic objects deteriorates with time. Just as for the positions, errors in proper motion can be both random and systematic.

Random errors in proper motion will naturally result from the random errors in the measurements of an object's positions at various epochs (see figure 9.3), but they can also result from an incomplete model of the object's motion. Many stars are components of binary or multiple systems. Ideally, the catalog's proper motion should

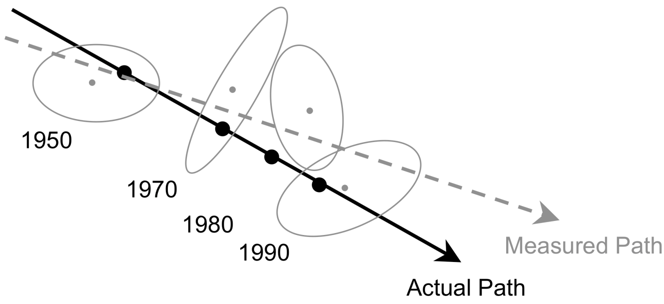

Figure 9.3 Errors in proper motion resulting from the random errors in the measurement of position at various epochs. The true path of a star (black line), with its positions at several epochs marked. The measured positions of the star (gray points) on the same dates are also shown, with their error ellipses. The catalog proper motion of the star (which determines its computed path across the sky) derived from these measurements is in error; therefore, the predicted positions of the star will deteriorate in accuracy with time. Since all the stars in a catalog share this problem to some extent, the reference frame that the catalog defines slowly degrades. If the proper motion errors have systematic trends across the sky, the catalog's reference frame may develop warps or exhibit a small, spurious net rotation.

reflect the linear space motion of the center of mass of the system, and the orbit of the star about that center of mass would then have to be taken into account as part of the computation of the star's position for any arbitrary date. Alternatively, stars may be listed with second- or higher-order terms in their motion (as in the Hipparcos catalog). However, in practice, a relatively small number of binary systems have well determined orbits, or are even recognized as such. The orbital motion of many stars is unknown and contaminates their measured proper motion. This is especially true of catalogs based on a relatively short time span of observations, where the slight curvature of the star's path may be less than the observational uncertainty [Makarov and Kaplan, 2005]. Under-sampling of short-period binary orbits has a similar effect. A related problem arises from measurements of the positions of the blended images of binaries that are unresolved to the instrument(s) used in a catalog's construction. However, such binaries may be resolved with other instruments or the blended image may appear displaced at other wavelength bands if the two stars are of different colors. For the individual stars involved, such modeling or observational problems would be considered systematic errors, although the net effect of many such errors on the reference frame as a whole would be random.

Additionally, there may be systematic errors in a catalog's proper motions. Such errors arise when the causes of the systematic errors in position are themselves functions of time. Systematic errors in the proper motions can produce time-dependent warping and spurious rotations of the frame. Therefore, the accuracy and consistency of the proper motions are critical to the overall quality, utility, and longevity of reference frames defined by stars.

Even reference frames defined by extragalactic objects, which are usually considered to have zero proper motion (i.e., unmeasurably small [12]), may deteriorate because many of these objects show small apparent motions at radio wavelengths that are artifacts of their emission mechanisms. Furthermore, if an object is even partially resolved, observations at wavelengths other than that used in the construction of the catalog may show an image center displaced from the catalog's position, arising from a color or spectral index gradient across the image. The variable radio flux and image structure of these objects may mean that some of them eventually must be replaced as reference frame benchmarks [Titov, 2007]. The systematic errors of reference frames based on extragalactic objects can be caused by the *secular aberration* [Kopeikin and Makarov, 2006] or asymmetry in the radio source distribution between hemispheres [Titov and Malkin, 2009].

9.4.4
Late 20th Century Developments

The establishment of celestial reference systems is coordinated by the IAU. The previous astronomical reference system was based on the equator and equinox of J2000.0 determined from observations of planetary motions, together with the IAU (1976) System of Astronomical Constants and related algorithms [Kaplan, 1981, 2005c]. The reference frame that embodied this system for practical purposes was the Fifth Fundamental Catalogue (FK5). The FK5 is a catalog of 1535 bright stars (to magnitude 7.5), supplemented by a fainter extension of 3117 additional stars (to magnitude 9.5), compiled at the Astronomische Rechen-Institut in Heidelberg and the U.S. Naval Observatory in Washington [Fricke *et al.*, 1991, 1988]. The FK5 was the successor to the FK3 and FK4 catalogs, all of which were based on meridian observations taken in the visual band – many such observations were, in fact, taken by eye. The formal uncertainties in the star positions of the FK5 at the time of its publication in 1988 were about 30-40 milliarcseconds over most of the sky, but the errors are considerably worse when systematic trends are taken into account.

Beginning in the 1970s, the most precise wide-angle astrometry was conducted, not in the optical regime, but at radio wavelengths, involving the techniques of Very Long Baseline Interferometry (VLBI) and pulsar timing. Uncertainties of radio source positions listed in good all-sky VLBI catalogs are now typically a fraction of a milliarcsecond. Furthermore, because these radio sources are very distant extragalactic objects (mostly optically faint quasars) that are not expected to show measurable intrinsic motion, a reference frame defined by VLBI positions should be "more inertial" (less subject to spurious rotation) than a reference frame defined by galactic objects, such as stars or pulsars. The origin of right ascension in VLBI catalogs is somewhat arbitrary, because VLBI has little sensitivity to the ecliptic plane and the location of the equinox is not easily determined. The VLBI origin of right ascension has effectively been carried over from one catalog to the next; it was originally based

12) See, however, [MacMillan, 2005].

on the right ascension of the source 3C273B based on lunar occultations.

Because of the accuracy and stability of radio reference frames, since the mid 1980s, astronomical measurements of the Earth's rotation – from which universal time (UT1) is determined – have depended heavily on VLBI, with classical methods based on star transits being phased out. Hence, the situation evolved to where the definition of the fundamental astronomical reference frame (the FK5) became irrelevant to some of the most precise and important astrometric measurements. VLBI revealed, in addition, that the models of the Earth's precession and nutation that were part of the old system were inadequate for modern astrometric precision. In particular, the "constant of precession"– a measurement of the long-term rate of change of the orientation of the Earth's axis in space – had been overestimated by about 0.3 arcsecond/century. Moreover, the success of the European Space Agency's Hipparcos astrometric satellite, launched in 1989, promised to provide a new, very accurate set of star coordinates in the optical regime.

Thus, beginning in 1988, a number of IAU working groups began considering the requirements for a new fundamental astronomical reference system [Hughes *et al.*, 1991; Lieske and Abalakin, 1990]. The resulting series of IAU resolutions, passed in 1991, 1994, 1997, and 2000 effectively form the specifications for the ICRS. The axes of the ICRS are defined by the adopted positions of a specific set of extragalactic objects, which are assumed to have no measurable proper motions. The ICRS axes are consistent, to about 0.02 arcsecond, with the equator and equinox of J2000.0 defined by the dynamics of the Earth. However, the ICRS axes are meant to be regarded as fixed directions in space that have an existence independent of the dynamics of the Earth or the particular set of objects used to define them at any given time.

Shortly after the adoption of the ICRS, Feissel and Mignard [1998] wrote a concise review of the the implications of that change. A few years later, Seidelmann and Kovalevsky [2002] published a broader review of the ICRS and the new IAU Earth orientation models. Although now somewhat dated, both publications give a good idea of how the IAU resolutions of 1997 and 2000 created a fundamental break with past practice.

The promotion, maintenance, extension, and use of the ICRS are the responsibilities of IAU Division 1 (Fundamental Astronomy).

9.4.5
ICRS Implementation

9.4.5.1 *The Defining Extragalactic Frame*

The original International Celestial Reference Frame (ICRF or ICRF1) is a catalog of positions of 608 extragalactic radio sources observed with VLBI, all strong (>0.1 Jy) at S and X bands (wavelengths 13 and 3.6 cm) [Ma and Feissel, 1997; Souchay and R., 2010]. The catalog was adopted by the IAU in 1997 as the primary realization of the ICRS. Most of the sources in the catalog have faint optical counterparts (typically $m_V \gg 18$) and the majority are quasars. Of these objects, 212 are *defining sources* that establish the orientation of the ICRS axes, with origin at the solar system barycenter. Typical position uncertainties for the defining sources are of order 0.5 milliarcsecond; the orientation of the axes is defined from the ensemble to an accuracy of about 0.02 milliarcsecond. As described in section 9.4.7, these axes correspond closely to what would conventionally be described as "the mean equator and equinox of J2000.0".

The International Earth Rotation and Reference Systems Service (IERS) monitors the radio sources involved in the ICRF. This monitoring is necessary because, at some level, most of the sources are variable in both flux and structure and the centers of emission can display spurious motions, which may not be linear or constant (see discussion in [Ma *et al.*, 1998, section 8]). As a result of additional observations and analysis, ICRF Extensions 1 and 2 were developed [Fey *et al.*, 2004]; the positions of the non-defining sources were improved and new sources were added. The positions of the 212 defining sources were left unchanged.

By 2006, it was recognized that a major update of the ICRF was needed to maintain the accuracy and fixed orientation of the overall frame, and an IAU working group was established to update the list of sources and coordinates. The working group presented a revised and extended list of sources and coordinates. The new list was adopted by the IAU in 2009 as the "Second Realization of the International Celestial Reference Frame," or ICRF2 [Ma *et al.*, 2009], superseding the original in defining the spatial orientation of the ICRS at S and X bands. The ICRF2 has 295 defining sources, chosen from a solution for the positions of 3414 sources. Only 97 of the defining sources are also defining sources in ICRF1, reflecting the results of the ongoing analysis of source stability and the working group's goal of mitigating source position variations. The positional uncertainties have been reduced considerably and the new list is more evenly distributed across the sky, especially in the south. Typical ICRF2 defining source position errors, all things considered, are not expected to exceed 0.1 milliarcseconds; the "noise floor" is 0.040 milliarcseconds. The overall orientation of the axes is estimated to be stable within 0.010 milliarcseconds and is consistent with that of ICRF1.

9.4.5.2 *The Frame at Optical Wavelengths*

The ICRS is realized at optical wavelengths by a subset of the Hipparcos catalog of 118,218 stars, some as faint as visual magnitude 12 [Perryman and ESA, 1997]. Only stars with uncomplicated and well-determined proper motions (e.g., no known binaries) are used for the ICRS realization. This subset, referred to as the Hipparcos Celestial Reference Frame (HCRF), comprises 85% of the stars in the Hipparcos catalog.[13] Hipparcos star coordinates and proper motions are given within the ICRS coordinate system, but are listed for epoch J1991.25. It means that the catalog effectively represents a snapshot of the positions and motions of the stars taken on 2 April 1991, a date that is near the mean epoch of the Hipparcos observations. At the catalog epoch, Hipparcos uncertainties for stars brighter than 9th magnitude have median values somewhat better than 1 milliarcsecond in position and 1 mil-

13) The HCRF excludes the stars in the Hipparcos catalog with astrometric solutions flagged C, G, O, V, or X.

liarcsecond/year in proper motion [Mignard, 1997; Perryman and ESA, 1997]. The overall alignment to the ICRF at that epoch is estimated to be within 0.6 milliarcsecond, with any spurious rotations or distortions less than 0.25 milliarcsecond/year. Projected to epoch 2010, typical position errors for the brighter Hipparcos stars are approximately 20 milliarcseconds.

A major reanalysis of the original Hipparcos observations [van Leeuwen, 2007a,b] has resulted in a new Hipparcos catalog with substantially improved astrometric data. However, the IAU has not taken any action that officially replaces the original Hipparcos catalog as the basis for the HCRF.

Extensions of the ICRS to fainter magnitudes are described in section 9.4.8.

9.4.6
Standard Algorithms

Sections 9.2, 9.3, 9.6, and 9.7 of this Chapter describe IAU-sanctioned algorithms used in the construction, maintenance, and use of the ICRS.

The 2000 IAU resolutions provide the relativistic metric tensors for what it called the Barycentric Celestial Reference System (BCRS) and the Geocentric Celestial Reference System (GCRS), as well as expressions for the transformation between the two systems; see section 9.2 and IAU resolution B1.3 of 2000. IAU resolution 2 of 2006 provides that the orientation of the BCRS axes be understood to be that of the ICRS/ICRF (now ICRF2).

In 2000, the IAU also adopted new models for the computation of the Earth's instantaneous orientation, which affect the analysis of VLBI observations that are used to define and maintain the ICRS, as well as the calculation of various observable quantities from ICRS-compatible reference data. The new models include what is referred to as the *IAU 2000A precession-nutation model*, a new definition of the celestial pole, and two new reference points in the equatorial plane for measuring the rotational angle of the Earth around its instantaneous axis. Despite the IAU action in 2000, some aspects of the models were not finalized until late 2002 (mid-2005 for agreement on the final precession expressions). These algorithms are described in sections 9.6 and 9.7 of this Chapter and in the McCarthy and Petit [2004].

The new Earth orientation models are, of course, relevant only to fundamental observations made from the surface of the Earth. Astrometric observations taken from space platforms, or those that are differential in nature (based on reference objects all located within a small field), do not use these models. There are, of course, other effects that must be taken into account in analyzing astrometric observations – e.g., proper motion, parallax, aberration, and gravitational light-bending – and algorithms for these may be found in Volumes 1 and 3 of the Hipparcos catalog documentation [Perryman and ESA, 1997]. For analysis of very high accuracy satellite observations, see the development by Klioner [2003a].

As described in section 9.1.3, there are two collections of general-purpose computer subroutines that implement the new IAU-sanctioned algorithms for practical applications: the Standards of Fundamental Astronomy (SOFA) [IAU SOFA Center, 2009], and the Naval Observatory Vector Astrometry Subroutines (NOVAS) [Kaplan *et al.*, 2009]. NOVAS also implements many of the Hipparcos algorithms, or the equivalent.

For ground-based applications requiring accuracies of no better than 50 milliarcseconds between about 1990 and 2010, the algorithms described in Chapter 3 of the Seidelmann and Urban [2010] can still be used with ICRS data. For such purposes, ICRS data can be treated as being on the dynamical equator and equinox of J2000.0. A major revision of the Explanatory Supplement to reflect the adoption of the ICRS and all the new models was expected to be complete in 2010.

9.4.7 Relationship to Other Systems

The orientation of the ICRS axes is consistent with the equator and equinox of J2000.0 represented by the FK5, within the errors of the latter [14)]. Systematically, the FK5 equator is tilted by 22 milliarcseconds and its origin of right ascension is offset by 23 milliarcseconds with respect to the ICRS. But the uncertainties of the FK5 equator and right ascension system with respect to the dynamical equator and equinox of J2000.0 are 50 and 80 milliarcseconds, respectively. Since, at J2000.0, the errors of the FK5 are significantly worse than those of Hipparcos, the ICRS (as realized by the HCRF) can be considered to be a refinement of the FK5 system [Perryman and ESA, 1997] at (or near) that epoch. For a more complete discussion of FK5 errors, see [Mignard and Frœschlé, 2000].

The ICRS can also be considered to be a good approximation (at least as good as the FK5) to the conventionally defined dynamical equator and equinox of J2000.0 [Feissel and Mignard, 1998], if the latter system is considered to be barycentric. This follows from the IAU resolution passed in 1991 (see section 9.4.1) that provided the original specifications for the new fundamental astronomical reference system based on extragalactic objects – what became the ICRS. In fact, the equator is well determined fundamentally from the VLBI observations that are the basis for the ICRS, and the ICRS pole is within 20 milliarcseconds of the dynamical pole. The reason that the ICRS pole is not perfectly aligned with the dynamical pole is complex. The ICRF was created from almost 20 years of VLBI observations from which a grand solution was made for the directions to the extragalactic radio sources and the changing position of the celestial pole. A specific decision made in that analysis [15)] resulted in an offset of the ICRF (hence ICRS) pole at J2000.0. With respect to the ICRS X and Y axes, the dynamical mean pole has coordinates on the unit sphere, in milliarcseconds, of approximately (-16.6, -6.8).

The zero point of VLBI-derived right ascensions is arbitrary, but traditionally has been set by assigning to the right ascension of source 3C 273B a value derived from lunar occultation timings – the Moon's ephemeris thus providing an indirect link to the dynamical equinox.[16)] The ICRS origin of right ascension was made to be

14) See [Feissel and Mignard, 1998] for a short discussion.
15) See Figure 9.5 and discussion related to it.
16) The original right ascension value, with respect to the equinox of 1950.0, was given by Hazard *et al.*

consistent with that in a group of VLBI catalogs previously used by the IERS, each of which had been individually aligned to the lunar occultation right ascension of 3C 273B. The difference between the ICRS origin of right ascension and the dynamical equinox has been independently measured by two groups that used different definitions of the equinox, but in both cases the difference found was less than 0.1 arcsecond.

Because of its consistency with previous reference systems, implementation of the ICRS will be transparent to any applications with accuracy requirements of no better than 0.1 arcseconds near epoch J2000.0. That is, for applications of this accuracy, the distinctions between the ICRS, FK5, and dynamical equator and equinox of J2000.0 are not significant.

9.4.8
Data in the ICRS

Although the ICRF2 and HCRF are its basic radio and optical realizations, the ICRS is gradually being extended to fainter magnitudes and other wavelengths. Thus, an increasing amount of fundamental astronomical data is being brought within the new system. A number of projects for the densification of the ICRS have been completed or are in progress.

As described above, the ICRF2 consists of the adopted positions of 3414 extragalactic radio sources, less than a tenth of which are defining sources. Because all of the observational data were part of a common catalog solution, the adopted coordinates of all the sources are in the ICRS. The non-defining sources tend to have positional uncertainties that are larger than those of the defining sources. Many of the non-defining sources are part of the VLBA Calibrator Survey (described below) and have a very limited observational base – 1966 of them were observed in only one session. Thus, their long-term stability has not been evaluated.

The VLBA Calibrator Survey is a list of radio sources, with positions in the ICRS, to be used as calibrators for the Very Long Baseline Array and the Very Large Array. The original list was prepared by Beasley *et al.* [2002]; the list has been extended several times and the current version is known as VCS6 [Petrov *et al.*, 2008]. The ICRS is also being established at higher radio frequencies (8, 24, 32, and 43 GHz); see, e.g., papers by Fey *et al.* [2009]; Jacobs *et al.* [2003]; Jacobs and Sovers [2008].

In the optical regime, the Tycho-2 Catalog [Høg *et al.*, 2000] (which supersedes the original Tycho Catalog and the ACT Reference Catalog) combines a re-analysis of the Hipparcos star mapper observations with data from 144 ground-based star catalogs. The ground-based catalogs include the Astrographic Catalog (AC), a large photographic project carried out near the beginning of the 20th century, involving 20 observatories worldwide. Tycho-2 contains 2,539,913 stars, to about magnitude 12, and combines the accuracy of the recent Hipparcos position measurements with proper motions derived from a time baseline of almost a century. Proper motion uncertainties are 1-3 milliarcseconds/year. The epoch J2000.0 Tycho-2 positions of

[1971]; the equivalent value in the J2000.0 system was first computed by Kaplan *et al.* [1982].

stars brighter than 9th magnitude are typically in error by 20 milliarcseconds. However, the positional accuracy degrades quite rapidly for magnitudes fainter than 9, so that 12th magnitude stars generally have a median J2000.0 position error exceeding 100 milliarcseconds.

Also in the optical band, the U.S. Naval Observatory CCD Astrographic Catalog (UCAC) provides ICRS-compatible positions and proper motions for stars as faint as visual magnitude 16. See Zacharias *et al.* [2004] for information on the second release of UCAC data; the third version was released on August 10, 2009 at the IAU General Assembly in Rio de Janeiro. The accuracy of the positions at epoch 2010 is estimated to be in the 20-100 milliarcsecond range, depending on magnitude. Proper motion errors are about 1-3 milliarcseconds/year for stars to 12th magnitude, and about 4-7 milliarcseconds/year for fainter stars to 16th magnitude.

The Large Quasar Reference Frame (LQRF) is another representation of the ICRS at faint optical magnitudes [Andrei *et al.*, 2009]. It contains the coordinates of 100,165 quasars, well distributed around the sky, accurate to about 100 milliarcseconds.

The ICRS has been extended to the near infrared through the Two Micron All Sky Survey (2MASS) [Cutri *et al.*, 2003; Zacharias *et al.*, 2005]. This ground-based program provides positions for 471 million point sources, most of which are stars, observed in the J, H, and K_s infrared bands. 2MASS is a single epoch survey without proper motions; positions are listed for J2000.0, which is within the four-year span of observations. Astrometric accuracy is around 80 milliarcseconds in the K_s magnitude range 9-14, with larger errors at both brighter and fainter magnitudes.

The Jet Propulsion Laboratory DE405/LE405 planetary and lunar ephemerides (usually just referred to as DE405) [Standish, 1998a] have been aligned to the ICRS. These ephemerides provide the positions and velocities of the nine major planets and the Moon with respect to the solar system barycenter, in rectangular coordinates. The data are represented in Chebyshev's polynomial series form, and Fortran subroutines are provided to read and evaluate the series for any date and time. DE405 spans the years 1600 to 2200; a long version, DE406, spans the years −3000 to +3000 with lower precision. More recent ephemerides from JPL, such as DE421, are also aligned to the ICRS. See section 9.5 for further details.

The barycentric data tabulated in *The Astronomical Almanac* [2010] are in the ICRS; this began with the *Astronomical Almanac* 2003 edition. Planetary and lunar ephemerides are derived from DE405/LE405. *The Astronomical Almanac* for 2006 is the first edition fully to support the new ICRS-related algorithms, including the new IAU Earth rotation models. Geocentric coordinates are based, therefore, on the GCRS: the reference system based on the true equator and equinox of date is just a rotation from the GCRS.

9.4.9
Formulas

A rotational matrix $\mathbb{B}$ is required to convert ICRS data to the dynamical mean equator and equinox of J2000.0 (the "J2000.0 system"), the latter considered for this purpose to be a barycentric system. The same matrix is used in the geocentric transformations described in sections 9.6 and 9.7 to adjust vectors in the GCRS (the "geocentric ICRS") so that they can be operated on by the conventional precession and nutation matrices. The matrix $\mathbb{B}$ is called the *frame bias matrix*, and it corresponds to a fixed set of very small rotations. In the barycentric reference system, it is used as follows

$$\boldsymbol{r}_{\rm mean}(J2000.0) = \mathbb{B}\,\boldsymbol{r}_{\rm ICRS} \tag{9.20}$$

where $\boldsymbol{r}_{\rm ICRS}$ is a BCRS vector with components measured with respect to the ICRS and $\boldsymbol{r}_{\rm mean}(J2000.0)$ is a BCRS vector with respect to the dynamical mean equator and equinox of J2000.0. Both of the $\boldsymbol{r}$'s are column vectors and, if they represent a direction on the sky [17], are of the general form of a unit vector

$$\boldsymbol{r} = \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} = \begin{bmatrix} \cos\delta\,\cos\alpha \\ \cos\delta\,\sin\alpha \\ \sin\delta \end{bmatrix} \tag{9.21}$$

where $n_x^2 + n_y^2 + n_z^2 = 1$, α is the right ascension and δ is the declination, with respect to the ICRS, or the dynamical system of J2000.0, as appropriate.

In the geocentric case, $\boldsymbol{r}_{\rm ICRS}$ is replaced by $\boldsymbol{r}_{\rm GCRS}$, and $\boldsymbol{r}_{\rm mean(2000)}$ is then a geocentric vector. This transformation must be carried out, for example, before precession is applied to GCRS vectors, since the precession algorithm assumes a dynamically based coordinate system. That is, the above transformation is a necessary step in obtaining coordinates with respect to the mean equator and equinox of date, if one starts with ICRS reference data. See section 9.6 for more information.

The matrix $\mathbb{B}$ is, to first order,

$$\mathbb{B} = \begin{bmatrix} 1 & d\alpha_0 & -\xi_0 \\ -d\alpha_0 & 1 & -\eta_0 \\ \xi_0 & \eta_0 & 1 \end{bmatrix} \tag{9.22}$$

where $d\alpha_0 = -14.6$ mas, $\xi_0 = -16.6170$ mas, and $\eta_0 = -6.8192$ mas; all converted to radians (divide the numbers by 206 264 806.247). The values of the three small angular offsets are taken from the IERS Conventions [McCarthy and Petit, 2004]. They can be considered adopted values; previous investigations of the dynamical-ICRS relationship obtained results that differ at the mas level or more, depending on the technique and assumptions [18]. The angles ξ_0 and η_0 are the ICRS pole offsets, and $d\alpha_0$ is the offset in the ICRS right ascension origin with respect to the dynamical equinox of J2000.0, as measured in an inertial (non-rotating) reference system.

The above matrix can also be used to transform vectors from the ICRS to the FK5 system at J2000.0. For this purpose, one should simply substitute $d\alpha_0 = -22.9$, $\xi_0 = 9.1$, and $\eta_0 = -19.9$ (milliarcseconds). However, there is also a time-dependent rotation of the FK5 with respect to the ICRS (i.e., a slow spin), reflecting the non-inertiality of the FK5 proper motions; see [Mignard and Frœschlé, 2000].

17) The length of the vector $\boldsymbol{r}$ is defined by the Euclidean norm as explained in section 2.3.2.1.
18) See the discussion in [Hilton and Hohenkerk, 2004].

Although the above matrix is adequate for most applications, a more precise result can be obtained by using the second-order version

$$\mathbb{B} = \begin{bmatrix} 1 - \frac{1}{2}(d\alpha_0^2 + \xi_0^2) & d\alpha_0 & -\xi_0 \\ -d\alpha_0 - \eta_0\xi_0 & 1 - \frac{1}{2}(d\alpha_0^2 + \eta_0^2) & -\eta_0 \\ \xi_0 - \eta_0 d\alpha_0 & \eta_0 + \xi_0 d\alpha_0 & 1 - \frac{1}{2}(\eta_0^2 + \xi_0^2) \end{bmatrix} \tag{9.23}$$

The above matrix, from the paper by Slabinski [2005], is an excellent approximation to the set of rotations $\boldsymbol{R}_1(-\eta_0)\boldsymbol{R}_2(\xi_0)\boldsymbol{R}_3(d\alpha_0)$, where $\boldsymbol{R}_1$, $\boldsymbol{R}_2$, and $\boldsymbol{R}_3$ are standard rotations about the x, y, and z axes, respectively (see appendix , “Abbreviations and Symbols Frequently Used” for precise definitions).

9.5
Ephemerides of the Major Solar System Bodies

The de facto standard source of accurate data on the positions and motions of the major solar system bodies is currently the ephemeris designated DE405/LE405 (or simply DE405) developed at the Jet Propulsion Laboratory. This ephemeris provides instantaneous position and velocity vectors of the eight major planets, Pluto, and the Earth's Moon, with respect to the solar system barycenter, for any date and time between 1600 and 2201. Lunar rotation angles (libration)are also provided. The ephemeris has been the basis for the tabulations in the *Astronomical Almanac* since the 2003 edition. The DE405 coordinate system has been aligned to the ICRS.

IAU-standard data on the sizes, shapes, rotational parameters, and latitude-longitude systems for the major planets and their satellites are given in the reports of the IAU/IAG Working Group on Cartographic Coordinates and Rotational Elements, issued every three years.

9.5.1
The JPL Ephemerides

A list of positions of one or more solar system bodies as a function of time is called an *ephemeris* (pl. *ephemerides*). An ephemeris can take many forms, including a printed tabulation, a sequential computer file, or a piece of software that, when interrogated, computes the requested data from series approximations or other mathematical schemes. As used in this section, an ephemeris represents the 3-D spatial coordinates of one or more solar system bodies, expressed in the BCRS as a function of a coordinate time scale such as TCB.

Ephemerides of the major solar system bodies, with respect to the solar system barycenter, have been calculated for many years at the Jet Propulsion Laboratory (JPL) to support various spacecraft missions. These ephemerides have been widely distributed and, because of their quality, have become the de facto standard source of such data for applications requiring the highest accuracy. Between the early 1980s and about 2000, the JPL ephemeris designated DE200/LE200 was most frequently used for such applications; it was the basis for the tabulations in the *Astronomical Almanac* from the 1984 to 2002 editions. A more recent JPL ephemeris, DE405/LE405, created in 1997, has been in widespread use since about 2000; for example, it has been the basis for the *Astronomical Almanac* since the 2003 edition. These ephemerides are usually referred to as just DE200 and DE405, respectively. Neither DE200 nor DE405 were ever formally adopted by the IAU, although they have been frequently reported on at various IAU-sponsored meetings and DE405 is a recommended standard of the IERS [McCarthy and Petit, 2004]. DE405 is now the basis of the barycentric time scale TDB (see below) as well as the precession of the ecliptic in the IAU-recommended P03 precession development. A comparison of DE405 with DE200, with an estimate of their errors, has been given by Standish [2004].

The JPL ephemerides are computed by an N-body numerical integration, carried out in a barycentric reference system which is consistent, except for the time scale used, with the Barycentric Celestial Reference System (BCRS) described in section 9.2, and in more theoretical details in section 4.3. The equations of motion [19], the method of integration, and the techniques used to adjust the starting conditions of the integration so that the results are an optimal fit to observations are described in Chapter 5 of the Seidelmann and Urban [2010]. That chapter specifically describes DE200, but the basic procedures are largely the same for all of the JPL ephemerides.

The position and velocity data provided by the JPL ephemerides represent the centers of mass of each planet-satellite system, although data for the Earth and Moon can be extracted separately. Therefore, the positions, when converted to geocentric apparent places – angular coordinates of an object as seen from Earth – do not precisely indicate the center of the apparent planetary disk. Displacements can amount to a few tens of milliarcseconds for Jupiter and Saturn, a few milliarcseconds for Uranus and Neptune, and about 0.1 arcsecond for Pluto.

9.5.2 DE405

The JPL DE405/LE405 ephemeris provides the coordinates and velocities of the major planets, Pluto, the Sun, and the Earth's Moon for the period 1600 to 2200 [Standish, 1998a]. The position and velocity 3-vectors are in equatorial rectangular coordinates referred to the solar system barycenter. The reference frame for the DE405 is the ICRF; the alignment onto this frame, and therefore onto the ICRS, has an estimated accuracy of a few milliarcseconds, at least for the inner-planet data. Optical, radar, laser, and spacecraft observations were analyzed to determine starting (initial) conditions for the numerical integration of the EIH equations of motion 6.82 and values of fundamental constants such as the Earth/Moon mass ratio and the length of the astronomical unit in meters. In addition to the planetary and lunar coordinates and velocities, the ephemerides, as distributed, include the nutation angles of the Earth and the rotation (libration) angles of the Moon. Note, however, that the nutation angles are not derived from the IAU 2000A theory described in section 9.6.

As described in section 9.3.2, DE405 was computed in a barycentric reference system using a coordinate time called T_{eph} as the independent argument . T_{eph} is similar to the IAU-defined coordinate time TCB (used in the BCRS) but differs in rate. For many years, users of the JPL ephemerides assumed that T_{eph} was simply another name for TDB, a coordinate time for barycentric developments that was defined by the IAU in 1976 and 1979; both T_{eph} and TDB are meant to match the rate of TT at the geocenter. However, the original definition of TDB was problematic when used with numerically integrated ephemerides [Standish, 1998b]. Therefore, IAU resolution B3 of 2006 redefined TDB to be a linear function of TCB that is effectively equivalent to T_{eph} for DE405. However, because T_{eph} = TDB differs in rate from that of TCB, many astronomical constants obtained from the JPL ephemerides must be scaled to be compatible with BCRS quantities based on TCB.

19) More exactly, the post-Newtonian equations of motion, are given in section 6.3.4 of this book.

The ephemerides are distributed by JPL as plain-text (ASCII) computer files of Chebyshev's polynomial series coefficients and Fortran source code. Third-party C versions of the code are also available and, for Unix users, the data files can be downloaded in binary form. Once the system is installed on a given computer, a Fortran subroutine named PLEPH can be called to provide the position and velocity of any requested body, at any date and time; PLEPH supervises the process of reading the Chebyshev file and evaluating the appropriate series. Normally the position and velocity vectors returned are expressed in units of AU and AU/day, respectively. The date/time requested must be expressed as a T_{eph} or TDB Julian date. An entry named DPLEPH is provided that allows the input Julian date to be split into two parts for greater precision. And, since $|TT - T_{eph}| < 0.002$ s, a TT Julian date may be used for applications not requiring the highest accuracy. The data and software files can be obtained on CD-ROM from Willmann-Bell, Inc. [Standish *et al.*, 2010], or downloaded from a JPL ftp server ftp://ssd.jpl.nasa.gov/pub/eph/export/. A "README" file provides export information and software documentation (available separately at Folkner [2007]).

An extended version of DE405/LE405, called DE406/LE406, is available that spans the years -3000 to +3000, but with coordinates given to lower precision (they are represented by shorter Chebyshev series). The nutation angles and the lunar rotation angles are also omitted from the DE406 files. DE406 is provided only in Unix binary format. These files are about 1/3 the size of those for DE405 for a given span of time. The additional error in the coordinates (DE406 - DE405) may amount to 25 m for the planets and 1 m for the Moon, which may be significant for some applications.

The NOVAS software package mentioned in the introduction to this chapter provides an interface to an existing DE405 or DE406 installation through Fortran subroutine SOLSYS (version 2) or C function *ephemeris*.

9.5.3
Recent Ephemeris Development

As this book is being written, DE405 has been in use for over ten years and the question of a successor has arisen, given new observational data from spacecraft made in the last decade. As of the date of this writing, DE421 is the most recent ephemeris produced by the JPL ephemeris group. That group is headed by Dr. William Folkner, and his report to IAU Commission 4 (Ephemerides) in August 2009 [20] summarized the ongoing work as follows:

"In the past three years the accuracy of the ephemerides of Venus and Saturn have improved dramatically due to observations from the ESA Venus Express mission and the NASA Cassini mission. The Mars ephemeris accuracy is being maintained by continuing observations of Mars-orbiting spacecraft... The outer planet ephemerides are being slowly improved by continuing ground-based astrometric observations,

20) See also the IPN Progress Report 42-178 on http://ipnpr.jpl.nasa.gov/progress_report/42-178/178C.pdf

and some older observations are being re-processed against modern ICRF-based star catalogs.

"The most recent JPL ephemeris released is DE421. The next release is planned for September 2009, primarily to include data from two encounters of Mercury by the MESSENGER spacecraft and in preparation for its third Mercury encounter in November 2009. Following that, the Mars spacecraft VLBI data set will be re-reduced against the recently released ICRF 2, followed by a new ephemeris fit in spring 2010."

It is worth noting here that the accuracy of modern ephemerides of bodies in the inner solar system is largely limited by perturbations by asteroids of unknown mass [Standish and Fienga, 2002]. Well determined masses are available for only a few of the largest asteroids, along with a few smaller ones that have been visited by spacecraft; the rest essentially represent a source of noise in the perturbing functions. In ephemeris computations, more sophisticated representation of groups of asteroids, by taxonomic or orbital family, can help, but until the masses of many more individual asteroids are known (particularly those with close approaches to the inner planets), the problem will remain. Continuing observations are needed to bound the runoff.

Two other ephemeris groups are now producing high quality solar system ephemerides comparable to the JPL products: the Institut de Mécanique Céleste et de Calcul des Éphémérides (IMCCE), associated with the Paris Observatory, and the Institute of Applied Astronomy (IAA) at St. Petersburg, Russia. Both groups use simultaneous N-body numerical integrations fitted to a large number of observations from a wide variety of sources, including recent in situ measurements by spacecraft. The equations of motion include general relativity terms consistent with the definition of the BCRS. The IMCCE series of ephemerides are referred to as INPOP(n) (Intégration Numérique Planétaire d'Observatoire de Paris; n is a serial number) and the IAA ephemerides are referred to as EPMy (Ephemerides of the Planets and Moon; y is a year). Both the INPOP and EPM ephemerides are available in versions based on TDB and TCB as the time argument.

Having three sets of fundamental planetary and lunar ephemerides (from JPL, IMCCE, and IAA), possibly joined by others in the future, is a healthy development for the field and will undoubtedly spur improvements. In 2010, a project was begun by IAU Commission 4 (Ephemerides) to provide a detailed comparison of these ephemerides, covering the models, numerical procedures, observation sets, and results, with the objective of providing users with information that would enable an informed choice for specific applications.

9.5.4
Sizes, Shapes, and Rotational Data

The IAU/IAG[21)] Working Group on Cartographic Coordinates and Rotational Elements [Seidelmann *et al.*, 2006] produces a report every three years for each IAU

21) IAG = International Association of Geodesy

General Assembly, giving the best estimates for the dimensions and rotational parameters of the planets, satellites, and asteroids, as far as is known. The working group is also responsible for establishing latitude-longitude coordinate systems for these bodies. The rotational elements given in the report for the 2006 General Assembly [Seidelmann *et al.*, 2007] serve to orient these coordinate systems within the ICRS as a function of time. The working group's reports are the basis for the physical ephemerides of the planets given in *The Astronomical Almanac* [2010].

Although the rotational elements of the Earth and Moon are given in each report for completeness, the expressions given there provide only an approximation to the known motions and should not be used for precise work. Lunar rotation (libration) angles can be obtained from DE405, and sections 9.6 and 9.7 of this chapter describe algorithms for the precise instantaneous alignment of the terrestrial coordinate system within the GCRS (the "geocentric ICRS").

9.6 Precession and Nutation

Precession and nutation are really two aspects of a single phenomenon, the overall response of the spinning, oblate, elastic Earth to external gravitational torques from the Moon, Sun, and planets. As a result of these torques, the orientation of the Earth's rotation axis is constantly changing with respect to a space-fixed (locally inertial) reference system. The motion of the celestial pole among the stars is conventionally described as consisting of a smooth long-term motion called precession upon which is superimposed a series of small periodic components called nutation.

The algorithms for precession used generally from about 1980 through 2000 (in the *Astronomical Almanac* from the 1984 through 2005 editions) were based on the IAU (1976) value for the rate of general precession in ecliptic longitude (5029.0966 arcseconds per Julian century at J2000.0). Nutation over most of the same time period was given by the 1980 IAU Theory of Nutation. However, not long after these algorithms were widely adopted, it became clear that the IAU (1976) rate of precession had been overestimated by approximately 3 milliarcseconds per year. Further observations also revealed periodic errors of a few milliarcseconds in the 1980 IAU Theory of Nutation. For many applications these errors are negligible, but they are significant at the level of the best ground-based astrometry and geodesy.

As part of the 2000 IAU resolutions, the IAU 2000A precession-nutation model was introduced, based on an updated value for the rate of precession and a completely new nutation theory. As before, the model actually consists of two parts, a precession algorithm describing the smooth secular motion of the celestial pole and a nutation algorithm describing the small periodic variations in the pole's position. The precession algorithm consists of short polynomial series for the values of certain angles. The sines and cosines of these angles, in combination, then define the elements of a precession matrix, $\mathbb{P}$. The nutation algorithm consists of a rather long series expansion in Fourier terms for the angular offsets, in ecliptic longitude and latitude, of the actual celestial pole (as modeled) from the precession-only pole (true pole - mean pole). The sines and cosines of these offsets, in combination, then define the elements of a nutation matrix, $\mathbb{N}$. The $\mathbb{P}$ and $\mathbb{N}$ matrices are applied to the coordinates of celestial objects, expressed as 3-vectors, to transform them from the equator and equinox of one epoch to the equator and equinox of another.

9.6.1 Aspects of Earth Rotation

The Earth is a relatively well-behaved rotating body, and illustrates the three basic elements of classical spin dynamics: precession, nutation, and Eulerian wobble. In fact, to first order, the Earth can be considered to be a rigid "fast top", and very good approximations to its rotational motion can be obtained from elementary developments. Although the effects of the Earth's liquid core, elastic mantle, and oceans are not negligible for modern observations, they can be considered to be small perturbations on the rigid-body motion. Since the Earth is nearly spherical and experiences relatively weak torques, its axis of rotation moves slowly with respect to both the stars and the body of the Earth itself.

The orientation of any rotating rigid body can be *theoretically* described as a time series of three Euler angles that relate a body-fixed coordinate system to a space-fixed coordinate system. If the body-fixed coordinate system can be defined such that the rate of change of one of the three Euler angles is much greater than that of the other two – as is the case for the Earth – then the rotational kinematics are usually described in terms of the slowly changing orientation of an axis of rotation passing through the body's center of mass. One can equivalently speak of the kinematics of the pole: one of the points where the axis of rotation intersects the body's surface or, extended to infinity, the "celestial sphere" [22)]. For this kinematic construction to work well, the angular motion of the axis or pole should be small and nearly linear over one rotation, predictable from theory, and observable.

However, as was pointed out by Smith and Turcotte [1993], when one *practically* uses such an axis or pole, one needs five angles, not three, to fully describe the instantaneous orientation of the body: two angles to describe the orientation of the body rotational axis with respect to the body-fixed coordinate system, one to describe the angle of the body's rotation about the axis, and two more to describe the orientation of the body-fixed coordinate system with respect to the non-rotating inertial space. For the Earth, these five angles correspond to the five standard parameters of Earth orientation disseminated by organizations such as the IERS: the coordinates of the pole, x_p and y_p, measured in a terrestrial coordinate system; the Universal Time difference, UT1-UTC; and the celestial pole offsets, $d\psi$ and $d\epsilon$, measured in a celestial coordinate system. Phenomenologically, the parameters divide up as follows: x_p and y_p describe polar motion, the variations in the position of the pole with respect to the Earth's crust; UT1-UTC measures the integrated variation in , the departure from a constant angular rate of rotation; and $d\psi$ and $d\epsilon$ are the errors in the computed position of the celestial pole, reflecting deficiencies in the adopted algorithms for precession and nutation.

What one calls polar motion corresponds, in rigid-body rotation, to the free Eulerian wobble of the figure axis about the rotation axis. On the real Earth, the phenomenon is not that simple. From an Earth-fixed (rotating) frame of reference, polar motion is a 10-meter (0.3 arcsecond) quasi-circular excursion in the pole position, with principal periods of 12 and 14 months. The 14-month component corresponds to the Eulerian wobble, as modified by the Earth's elasticity, while the 12-month component undoubtedly is a seasonal effect. Smaller, quasi-random variations are not well understood. None of the components is regular enough to permit reliable predictions, and polar motion must be obtained from observations.

Variations in the Earth's rotation rate are due to several causes. There are fortnightly, monthly, semiannual, and annual tidal effects, and other short-term and seasonal changes are largely due to exchange of angular momentum with the atmo-

22) Only the former definition of the pole is accepted in the relativistic framework because the rotational theory of a body (see section 6.2.2) is formulated in the local coordinates which do not spread out to infinity.

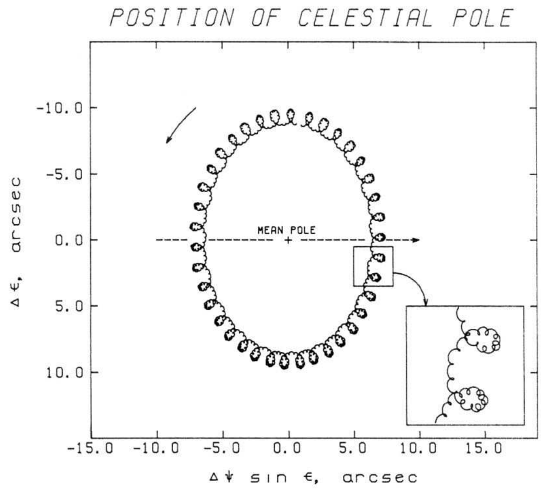

Figure 9.4 The path of the true celestial pole on the sky, over an 18-year period, compared to the mean pole. The mean pole moves along a smooth arc at a rate of 20 arcseconds per year due to precession only. The complex epicyclic motion of the true pole is nutation. The inset shows the detail of one year's motion.

sphere. Longer-term variations (decade fluctuations) are less well understood. For a discussion of time scales that are based on the variable rotation of the Earth, see sections 9.3.1 and 9.3.3.

Precession and nutation refer to the changing orientation of the Earth's axis, with respect to a space-fixed (kinematically non-rotating) system, in response to external torques. The torques are due to the gravitational attraction of the Moon and Sun and, to a much lesser extent, the planets on the equatorial bulge of the Earth. Precession and nutation are really different aspects of a single physical phenomenon, and it has become more common in recent years to write "precession-nutation". Precession is simply the secular term in the response of Earth to the external gravitational torque, while nutation is the set of periodic terms. Due to the precession the celestial pole traces out a circle on the celestial sphere, about 23° in radius, centered on the ecliptic pole (the direction orthogonal to the ecliptic plane), taking about 26,000 years to complete one circuit (≈20 arcseconds/year). Precession theory describes this smooth, long-term motion, and the precessional pole is referred to as the *mean pole* (the orthogonal plane is the *mean equator*). But the pole also undergoes a hierarchy of small epicyclic motions, the largest of which is a 14×18 arcsecond ellipse traced out every 18.6 years (see Figure 9.4). Nutation theory describes this and other periodic motions with unprecedented precision. To get the path of the *true pole* on the celestial sphere (i.e., the direction of the Earth's axis in space), it is necessary to compute both precession and nutation; conventionally, they are described by separate time-dependent rotation matrices, $\mathbb{P}(t)$ and $\mathbb{N}(t)$, which are either multiplied together or applied sequentially.

9.6.2
Which Pole?

In theoretical developments of Earth rotation, the first issue that must be confronted is the definition of the celestial pole. If the Earth were a rigid oblate spheroid, there would be three possible axes, and corresponding poles, to choose from: the angular momentum axis; the rotation axis, defined by the instantaneous angular velocity vector; and the figure axis, which is the body-fixed axis orthogonal to the geometric equator and along the unique eigenvector of the Earth's inertia tensor. The distinctions among the axes arise from the physics of rotation. For example, as previously noted, in a rotating rigid body, the free Eulerian wobble describes the motion of the figure axis with respect to the rotation axis (or vice versa). The analog of this on the real Earth is polar motion.

In the previous section, precession-nutation was described as the changing orientation of the Earth's axis in response to external torques, expressed in a space-fixed (non-rotating) frame of reference. Which axis? The principal components of the response are rather large, amounting to many arcseconds over the course of a year, and are nearly the same for all three axes. However, the three axes cannot coincide in the presence of external torques. For a rigid Earth, the forced oscillations of the figure and rotation axes differ by about 10 milliarcseconds, and those of the angular momentum and rotation axes differ by only about 1 milliarcsecond. Until the mid-20th century, observations were not accurate enough to distinguish between the axes, so the choice of the best axis for theory was academic. But with improving observational accuracy and new techniques coming online in the 1960s and 1970s, the question of which axis should be used for the theoretical developments became important. After considerable discussion, the consensus emerged that the forced motion of the figure axis was the most relevant for observations, and therefore also for theory [23].

At about the same time, new theoretical work was being undertaken based on Earth models that were triaxial and contained a liquid core and elastic mantle. Such theories complicate the axis question considerably, because the inertia tensor varies with time as the Earth's shape responds to tidal forces, and the tidal deformation results in large daily excursions of the Earth's axis of figure. These excursions do not, in general, reflect the changing overall orientation of the Earth's crust in space, which is relevant to astronomical observations. That is, for the elastic Earth, the figure axis as classically defined is not an astronomically useful axis. The solution is to construct a rotating Cartesian coordinate system tied to the elastic, rotating Earth in such a way that (1) the net angular momentum of the tidal deformation, relative to this system, is always zero; and (2) for zero tidal deformation, the axes correspond to the principal axes of the Earth's mantle. These axes are the "Tisserand mean axes of the body" [Munk and MacDonald, 1960], and the Tisserand axis of

23) Section 6.2.2 demonstrates unequivocally that the external torque affects directly motion of the angular momentum (spin) of the body. However, the practical measurement of the direction of the angular momentum is impossible.

the maximum moment of inertia is referred to in resolution B1.7 of 2000 as "the mean surface geographic axis". Almost all modern theories of nutation refer to the principal Tisserand axis; in the previously used 1980 IAU Theory of Nutation it was referred to as axis B, and the corresponding pole called the "Celestial Ephemeris Pole".

However, even if one has chosen an axis that best reflects the overall rotation of stations (observatories) on the Earth's surface, a further complication arises as the observations and theoretical developments become more sensitive to short-period motions. The problem is the small but non-negligible circular components of nutation or polar motion with periods near one day. One can imagine the geometric confusion that arises when the pole undergoes a circular motion in one rotation period; in fact, it becomes difficult to disentangle the various effects, and the conventional labels become nearly meaningless. For example, any prograde nearly-diurnal nutation is equivalent to a long-period variation in polar motion, and any retrograde nearly-diurnal polar motion appears as a long-period nutation component [Capitaine, 2000]. In practice, this means a potential "leakage" or "crossover" from the Earth orientation parameters x_p and y_p to $d\psi$ and $d\epsilon$ or vice versa. The only practical solution is an explicit (although somewhat arbitrary) cutoff in the periods of what is considered precession-nutation, embodied in the definition of the celestial pole.

Therefore, the new IAU definition of the celestial pole to be used for the new precession-nutation model (IAU resolution B1.7 of 2000) is defined by the motions of Tisserand mean axis of the Earth with periods greater than two days in the celestial reference system. This pole is called the *Celestial Intermediate Pole* (CIP). The position of the CIP is given by the adopted precession-nutation model plus observational corrections. The word *intermediate* reminds us that the definition of the pole is merely a convention, serving to impose a division between what one calls precession-nutation (the Earth orientation angles measured in the celestial system) and polar motion (the Earth orientation angles measured in the terrestrial system). The CIP is the true pole, orthogonal to the true equator of date. Its motion is defined within the Geocentric Celestial Reference System (GCRS) – see section 9.2. Therefore, the geometric transformations described in this section (as well as those in section 9.7) all apply within a geocentric system. The GCRS can be described loosely as the "geocentric ICRS", since its axis directions are obtained from those of the ICRS.

9.6.3
The New Models

The variables $d\psi$ and $d\epsilon$ are the small angular offsets on the sky expressing the difference between the position of the celestial pole that is observed and the position predicted by the conventional precession and nutation theories. These angles are just the differential forms of the angles $\Delta\psi$ and $\Delta\epsilon$ in which nutation theories are conventionally expressed ($d\psi$ and $d\epsilon$ are sometimes labeled $\Delta\Delta\psi$ and $\Delta\Delta\epsilon$). $\Delta\psi$ and $\Delta\epsilon$ are in turn differential forms of the ecliptic coordinates of the celestial pole (see Figure 9.4).

Obviously the time series of $d\psi$ and $d\epsilon$ values, if they show systematic trends, can be used to improve the theories of precession and nutation. In fact, 20 years of $d\psi$ and $d\epsilon$ values from VLBI show significant patterns as shown in figure 9.5. Most obvious is the overall downward slope in longitude and an annual periodicity in both longitude and obliquity, suggesting the need for substantial corrections to the precession rate as well to the annual nutation term. A long-period sinusoid is also evident, and spectral analysis reveals the presence of a number of periodic components. The Figure 9.5 indicates the origin of the ICRS "frame bias" discussed in section 9.4. The pole offsets shown are taken from the solution for the ICRF catalog. The ICRS frame biases in longitude and obliquity are essentially the values, at J2000.0 (Time=100), of the two curves fitted to the data. The data was arbitrarily zeroed near the beginning of the data span, which led to non-zero values at J2000.0.

Other techniques, particularly lunar laser ranging (LLR), confirm the general trends. As a result, there has been a major multinational effort to improve the precession and nutation formulation and obtain interesting geophysical information in the process. This project, coordinated by an IAU/IUGG[24] working group, has involved dozens of investigators in several fields, and the resulting algorithms, taken together, are referred to as the IAU 2000A precession-nutation model.

The VLBI observations of $d\psi$ and $d\epsilon$ indicate the error in the computed position of the pole with respect to a space-fixed system defined by the positions of extragalactic objects. However, the conventional expressions for precession and nutation have used angles measured with respect to the ecliptic, a plane to which VLBI is not sensitive. The ecliptic plane has a slow precessional movement of its own due to planetary perturbations on the heliocentric orbital motion of the Earth-Moon barycenter.[25] In the theoretical developments it is necessary to distinguish between *precession of the equator* and *precession of the ecliptic*, which were formerly called, respectively, *lunisolar precession* and *planetary precession*. Both types of precession are measured with respect to a space-fixed system. The algorithms for precession and nutation provide the motion of the equator, as appropriate for most observations, but generally use a moving ecliptic as a reference plane for at least some of the angles involved (there are different formulations of precession using different angle sets). This allows the precession and nutation transformations to properly account for the motion of the equinox as well as that of the equator. The precession of the ecliptic is obtained from theory (although indirectly tied to observations through the JPL DE405 ephemeris), as are the high-order (unobserved) components of the precession of the equator. However, because of the mix of theory and observation that is involved in the final expressions, raw corrections to rates of precession from VLBI observations will not in general propagate exactly to the familiar precession quantities.

The changes in the amplitudes of the nutation components are also not directly taken from these observations; instead, a new nutation theory is developed and fit to observations by allowing a small number of geophysical constants to be free pa-

24) IUGG = International Union of Geodesy and Geophysics

25) The *mean ecliptic* is always implied. This is the smoothly moving plane that does not undergo the periodic oscillations of the instantaneous orbital plane of the Earth.

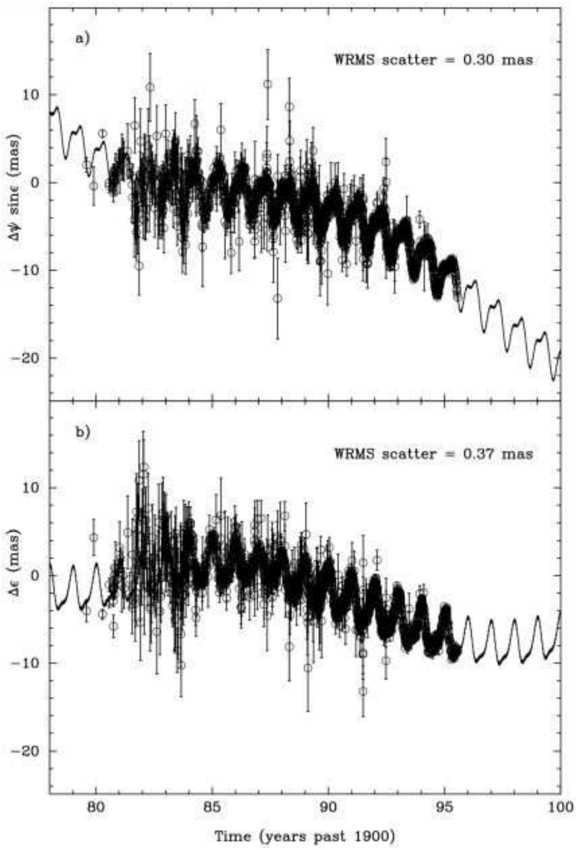

Figure 9.5 Observed values of celestial pole offsets from VLBI data. Offsets in longitude have been multiplied by the sine of the obliquity to allow the same scale to be used for both components. Circled points with error bars represent the offset of the observed pole with respect to the computed pole, and the solid line in each plot is a curve fitted to the data. The computed pole is given by Lieske *et al.* [1977] precession expressions and the 1980 IAU Theory of Nutation. These plots are from [Ma *et al.*, 1998].

Table 9.1 Precession-Nutation: Old & New. Values are given in arcseconds at J2000.0

Quantity	Old value	New value	New-Old
General precession in longitude (/cen)	5029.0966	5028.796195	−0.3004
Mean obliquity	84381.448	84381.406	−0.042
Mean obliquity rate (/cen)	−46.8150	−46.836769	−0.0218
In-phase nutation amplitudes:			
18.6-year longitude	−17.1966	−17.2064161	−0.0098
18.6-year obliquity	9.2025	9.2052331	0.0027
9.3-year longitude	0.2062	0.2074554	0.0013
9.3-year obliquity	−0.0895	−0.0897492	−0.0002
annual longitude	0.1426	0.1475877	0.0050
annual obliquity	0.0054	0.0073871	0.0020
semiannual longitude	−1.3187	−1.3170906	0.0016
semiannual obliquity	0.5736	0.5730336	−0.0006
122-day longitude	−0.0517	−0.0516821	0.0000
122-day obliquity	0.0224	0.0224386	0.0000
monthly longitude	0.0712	0.0711159	−0.0001
monthly obliquity	−0.0007	−0.0006750	0.0000
semimonthly longitude	−0.2274	−0.2276413	−0.0002
semimonthly obliquity	0.0977	0.0978459	0.0001

rameters. These parameters are constants in a "transfer function" that modifies the amplitudes of the terms from a rigid-Earth nutation development. Since there are fewer solved-for geophysical constants than the number of terms with observed amplitudes, the fit cannot be perfect. For the IAU 2000A model, seven geophysical parameters were determined based on the observed amplitudes of 21 nutation terms (prograde and retrograde amplitudes for each) together with the apparent change in the rate of precession in longitude. Note that the number of observational constraints and the number of free parameters in the model are both quite small compared to the 1365 terms in the new, full nutation series.

Table 9.1 compares the old and new values, at epoch J2000.0, of some of the primary quantities involved in the precession and nutation algorithms. In the table, all quantities are in arcseconds, and the rates (marked /cen) are per Julian century of TDB (or TT). The longitude components should be multiplied by the sine of the obliquity (≈0.3978) to obtain the corresponding motion of the pole on the celestial sphere. The new mean obliquity at J2000.0 is $23^\circ\,26'\,21''\!.406$. The theories from which the values are taken are:

- Old precession: given by [Lieske *et al.*, 1977], based on the IAU (1976) values for general precession and the obliquity at J2000.0, shown in the table 9.1
- Old nutation: 1980 IAU Theory of Nutation by Wahr [1981], based on the theory by Kinoshita [1977]; see report of the IAU working group by Seidelmann [1982]
- New precession: P03 solution by Capitaine *et al.* [2003]; see report of the IAU

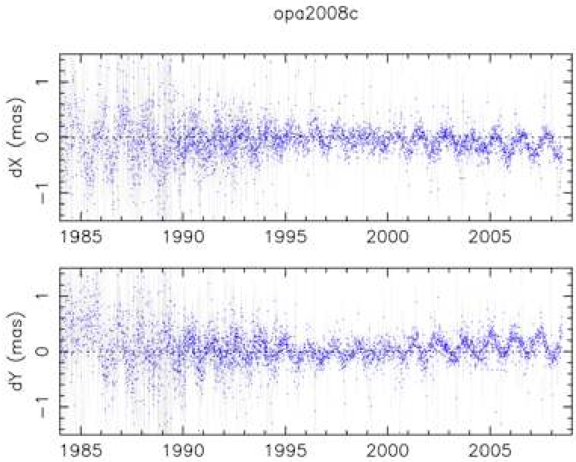

Figure 9.6 Observed values of celestial pole offsets, similar to Figure 9.5, but comparing recent VLBI data with the currently recommended precession-nutation theories (P03 precession and MHB nutation). The ordinates dX and dY correspond closely to the ordinates $\Delta\psi \sin\epsilon$ and $\Delta\epsilon$, respectively, used in the previous figure (but note the change of scale). The periodic signature that remains is due mainly to the free core nutation, which has a period of about 430 days but an arbitrary amplitude and phase. These plots are from Capitaine *et al.* [2009].

working group by Hilton *et al.* [2006]

- New nutation: worked out by Buffett *et al.* [2002] (often referred to as MHB), based on the work by Souchay *et al.* [1999]; series listed in [McCarthy and Petit, 2003a]

The MHB nutation was adopted in IAU resolution B1.6 of 2000, even though the theory had not been finalized at the time of the IAU General Assembly of that year. The new precession development was adopted in IAU resolution B2 of 2006, although it had been available since 2003. Used together, these two developments yield the computed path of the Celestial Intermediate Pole (CIP) as well as that of the true equinox. The formulas given below are based on these two developments.

A comparison of the new theories, taken together, with modern high accuracy observations is shown in figure 9.6.

It should be noted that the new precession model uses a slightly different value of the obliquity of the ecliptic than was assumed for the nutation model, and it also incorporates a rate of change in the Earth's dynamical form-factor (J_2) not considered previously. As explained by Capitaine *et al.* [2005], when the new nutation model is used with the new precession model (as would normally be the case) the nutation series terms require some correction to account for these differences in the underlying dynamical model. In practice, the required corrections affect only a few terms, and the changes are at the microarcsecond level at the current epoch. The corrections have not been formally adopted by the IAU, and are not discussed or implicit in any of the following developments. However, a corrected nutation series is available as part of the SOFA subroutines.

9.6.4 Formulas

In the development below, precession and nutation are represented as 3×3 rotation matrices that operate on column 3-vectors. The latter are position vectors in a specific celestial coordinate system – which must be stated or understood – with components that are Cartesian (rectangular) coordinates. They have the general form

$$\boldsymbol{r} = \begin{bmatrix} r_x \\ r_y \\ r_z \end{bmatrix} = \begin{bmatrix} r\cos\delta\cos\alpha \\ r\cos\delta\sin\alpha \\ r\sin\delta \end{bmatrix} \tag{9.24}$$

where r_x, r_y, and r_z are the three rectangular components of vector $\boldsymbol{r}$, α is the right ascension, δ is the declination, and $r = (r_x^2 + r_y^2 + r_z^2)^{1/2}$ is the radial coordinate distance from the specified origin. The distance r is an important observable for the objects lying within the solar system. However, for stars and other objects "at infinity" (beyond the solar system), r is not so important, or simply unmeasurable, and often set to 1. The celestial coordinate system being used will be indicated by a subscript, e.g., $\boldsymbol{r}_{\text{GCRS}}$. If one has the vector $\boldsymbol{r}$ in some coordinate system, then the right ascension and declination in that coordinate system can be obtained from

$$\alpha = \arctan\left(\frac{r_y}{r_x}\right), \tag{9.25a}$$

$$\delta = \arctan\left(\frac{r_z}{\sqrt{r_x^2 + r_y^2}}\right). \tag{9.25b}$$

A two-argument arctangent function (often called "atan2") will return the correct quadrant for α, if r_y and r_x are provided separately.

In the context of traditional equatorial celestial coordinate systems, the adjective *mean* is applied to quantities (pole, equator, equinox, coordinates) affected only by precession, while *true* describes quantities affected by both precession and nutation. This is a computational distinction only, since precession and nutation are simply different aspects of the same physical phenomenon. Thus, it is the *true* quantities that are directly relevant to observations; *mean* quantities now usually represent an intermediate step in the computations, or the final step where only very low accuracy is needed (10 arcseconds or worse) and nutation can be ignored.

Thus, a precession transformation is applied to celestial coordinates to convert them from the mean equator and equinox of J2000.0 to the mean equator and equinox of another date, t. Nutation is applied to the resulting coordinates to transform them to the true equator and equinox of t. These transformations should be understood to be inherently geocentric rotations and they originate in dynamical theories. Generally, one will be starting with celestial coordinates in the GCRS, which are obtained from basic ICRS data by applying the usual algorithms for proper place.

Computing *proper place* involves adjusting the catalog place of a star or other extra-solar system object for proper motion and parallax (where known), gravitational light deflection within the solar system, and aberration due to the observer's and Earth's motions. For a solar system object there are comparable adjustments to its position vector taken from a barycentric dynamical ephemeris [26]. In conventional usage, an *apparent place* can be considered to be a proper place that has been transformed to the true equator and equinox of date. The details of the proper place computations are beyond the scope of this Chapter but are described in detail in many textbooks on positional astronomy, and in papers by Smith *et al.* [1989,?]; Yallop *et al.* [1989], and Klioner [2003a].

As discussed in section 9.4, the ICRS is not based on a dynamically defined equator and equinox and so neither is the GCRS. Therefore, before one applies precession and nutation – and if one requires a final accuracy of better than 0.02 arcsecond – one must first apply the frame bias correction (section 9.4.9) to transform the GCRS coordinates to the dynamical mean equator and equinox of J2000.0. Schematically, these transformations sequence is

$$
\begin{array}{c}
\boxed{\text{GCRS}} \\
\| \\
\textit{frame bias} \\
\Downarrow \\
\boxed{\text{mean equator \& equinox of J2000.0}} \\
\| \\
\textit{precession} \\
\Downarrow \\
\boxed{\text{mean equator \& equinox of } t} \\
\| \\
\textit{nutation} \\
\Downarrow \\
\boxed{\text{true equator \& equinox of } t}
\end{array}
$$

Mathematically, this sequence can be represented as the matrix transform

$$\boldsymbol{r}_{\text{true}}(t) = \mathbb{N}(t)\,\mathbb{P}(t)\,\mathbb{B}\,\boldsymbol{r}_{\text{GCRS}} \tag{9.26}$$

where $\boldsymbol{r}_{\text{GCRS}}$ is a direction vector with respect to the GCRS and $\boldsymbol{r}_{\text{true}}$ is the equivalent vector with respect to the true equator and equinox of t. In equation (9.26), $\mathbb{N}(t)$ and $\mathbb{P}(t)$ are the nutation and precession rotation matrices, respectively. The remainder

26) See section 9.2.3.

of this section shows how to compute the elements of these matrices. The quantity $\mathbb{B}$ in equation (9.26) is the constant frame-bias matrix given in section 9.4.9.

The transformation from the mean equator and equinox of J2000.0 to the mean equator and equinox of t is simply

$$\boldsymbol{r}_{\text{mean}}(t) = \mathbb{P}(t)\,\boldsymbol{r}_{\text{mean}}(J2000.0) \tag{9.27}$$

and the reverse transformation is

$$\boldsymbol{r}_{\text{mean}}(J2000.0) = \mathbb{P}^{\mathrm{T}}(t)\,\boldsymbol{r}_{\text{mean}}(t) \tag{9.28}$$

where $\mathbb{P}^{\mathrm{T}}(t)$ is the transpose of matrix $\mathbb{P}(t)$.

What is described above are all conventional, equinox-based transformations. The equinox is the traditional origin of right ascension. An alternative transformation has been developed based on another point on the celestial equator called the Celestial Intermediate Origin (CIO); see resolution B1.8 of 2000. The alternative scheme has been introduced because the equinox is based on a barycentric concept (the ecliptic, the Earth's mean orbit) that is not relevant to a geocentric system; furthermore, the equinox suffers from ambiguity of definition below the 0.1 arcsecond level. Additionally, the new scheme cleanly separates different aspects of Earth orientation in the overall transformation between the terrestrial and celestial coordinate systems. The conventional transformations are described in this section because of widespread current usage. The angles needed for these transformations are provided by the newest theories of precession and nutation. The transformation, which combines frame bias, precession, and nutation into a single matrix, is given in section 9.6.7 and is described more fully in section 9.7, where the CIO is introduced and explained.

The true celestial pole of date t – the Celestial Intermediate Pole (CIP) – has, by definition, unit vector coordinates (0,0,1) with respect to the true equator and equinox of date. Therefore one can obtain the computed coordinates of the CIP with respect to the GCRS by simply reversing the transformation of equation (9.26)

$$\boldsymbol{r}_{\text{(CIP)GCRS}} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathbb{B}^{\mathrm{T}}\,\mathbb{P}^{\mathrm{T}}(t)\,\mathbb{N}^{\mathrm{T}}(t) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} (\mathbb{NPB})_{31} \\ (\mathbb{NPB})_{32} \\ (\mathbb{NPB})_{33} \end{bmatrix} \tag{9.29}$$

where one has used notation $\mathbb{NPB} = \mathbb{N}(t)\mathbb{P}(t)\mathbb{B}$, and where the sub-indices indicate the corresponding components of the matrix $\mathbb{NPB}$. Daily values of the elements of the combined matrix $\mathbb{NPB}$ are listed in *The Astronomical Almanac* [2010].

McCarthy and Petit [2004] list series expansions that directly provide X and Y, the two most rapidly changing components of the pole position unit vector. Daily values of X and Y are also listed in *The Astronomical Almanac* [2010]. The values of X and Y are given in arcseconds and are converted to dimensionless unit vector components simply by dividing them by the number of arcseconds in one radian, $648000''/\pi \simeq 206264.806247$. As soon as X and Y components are known, then, $Z = \sqrt{1 - X^2 - Y^2}$. The values of X and Y are used in the new transformation scheme discussed in section 9.6.7 and in several places in section 9.7.

9.6.5
Formulas for Precession

To construct the precession matrix for the transformation of coordinates from one date to another, one must evaluate short polynomials for the angles involved. The expressions for these angles in the IAU 2000A model, given below, have only a single time argument, since precession from or to J2000.0 (actually, the TDB equivalent of J2000.0) is assumed. As used in this Chapter (and *The Astronomical Almanac* [2010]), the matrix $\mathbb{P}(t)$ always denotes precession from J2000.0 (TDB) to another date, t. To precess in the opposite direction, the angles are the same but the transpose of the precession matrix, $\mathbb{P}^{\mathrm{T}}(t)$, is used. To precess coordinates from one arbitrary date, t_1, to another, t_2, it is necessary to precess them from t_1 to J2000.0 (using $\mathbb{P}^{\mathrm{T}}(t_1)$), then from J2000.0 to t_2 (using $\mathbb{P}(t_2)$). Where high accuracy is not required, and t_1 and t_2 are not more than a few years apart, a simpler procedure for precession from t_1 to t_2 is available and is given at the end of this subsection.

All expressions given in this subsection are from section 7 (P03 solution) of the work by Capitaine *et al.* [2003] and all coefficients are expressed in arcseconds. This is the theory of precession recommended by the IAU Working Group on Precession and the Ecliptic [Hilton *et al.*, 2006].

For a given TDB date and time t, let T be the number of Julian centuries of TDB since 2000 Jan 1, 12^{h} TDB. If the dates and times are expressed as Julian dates, then $T = (t - 2451545.0)/36525$. TT dates and times can be used equally well – the resulting error in precession is only a few $\times 10^{-9}$ arcseconds. Then, the mean obliquity of the ecliptic at J2000.0 (or the equivalent TDB date) is $\epsilon_0 = 84381.406$ arcseconds, and the angles, characterizing the precession are as follows

$$\psi_A = 5038.481507\,T - 1.0790069\,T^2 - 0.00114045\,T^3 + 0.000132851\,T^4 - 0.0000000951\,T^5\,, \tag{9.30}$$

$$\omega_A = \epsilon_0 - 0.025754\,T + 0.0512623\,T^2 - 0.00772503\,T^3 - 0.000000467\,T^4 + 0.0000003337\,T^5\,, \tag{9.31}$$

$$\chi_A = 10.556403\,T - 2.3814292\,T^2 - 0.00121197\,T^3 + 0.000170663\,T^4 - 0.0000000560\,T^5\,, \tag{9.32}$$

or equivalently, in notation appropriate for computer programs,

$$
\begin{aligned}
\psi_A = ((((\;& - \quad 0.0000000951 \; T \\
& + \quad 0.000132851 \;) T \\
& - \quad 0.00114045 \;) T \\
& - \quad 1.0790069 \;) T \\
& + \quad 5038.481507 \;) T \,, \\
\omega_A = ((((\;& + \quad 0.0000003337 \; T \\
& - \quad 0.000000467 \;) T \\
& - \quad 0.00772503 \;) T \\
& + \quad 0.0512623 \;) T \\
& - \quad 0.025754 \;) T + \epsilon_0 \,, \\
\chi_A = ((((\;& - \quad 0.0000000560 \; T \\
& + \quad 0.000170663 \;) T \\
& - \quad 0.00121197 \;) T \\
& - \quad 2.3814292 \;) T \\
& + \quad 10.556403 \;) T \,.
\end{aligned}
\tag{9.33}
$$

The precession matrix is, then, a product of the four matrices

$$\mathbb{P}(t) = \mathbb{R}_3(\chi_A)\,\mathbb{R}_1(-\omega_A)\,\mathbb{R}_3(-\psi_A)\,\mathbb{R}_1(\epsilon_0) \,, \tag{9.34}$$

where $\mathbb{R}_1$ and $\mathbb{R}_3$ are standard counterclockwise rotations about the x and z axes, respectively (see appendix "Abbreviations and Symbols Frequently Used" for precise definitions). This four-angle precession formulation is comprised of the following elements:

1) A rotation from the mean equator and equinox of J2000.0 to the mean ecliptic and equinox of J2000.0. This is a rotation around the x-axis (the direction toward the mean equinox of J2000.0) by the angle ϵ_0, the mean obliquity of J2000.0. After the rotation, the fundamental plane is the ecliptic of J2000.0.
2) A rotation around the new z-axis (the direction toward the ecliptic pole of J2000.0) by the angle $-\psi_A$, the amount of precession of the equator from J2000.0 to t.
3) A rotation around the new x-axis (the direction along the intersection of the mean equator of t with the ecliptic of J2000.0) by the angle $-\omega_A$, the obliquity of the mean equator of t with respect to the ecliptic of J2000.0. After the rotation, the fundamental plane is the mean equator of t.
4) A rotation around the new z-axis (the direction toward the mean celestial pole of t) by the angle χ_A, accounting for the precession of the ecliptic along the mean equator of t. After the rotation, the new x-axis is in the direction of the mean equinox of date.

If one denotes

$$S_1 = \sin(\epsilon_0) \,, \quad S_2 = \sin(-\psi_A) \,, \quad S_3 = \sin(-\omega_A) \,, \quad S_4 = \sin(\chi_A) \,, \tag{9.35}$$

$$C_1 = \cos(\epsilon_0)\,, \quad C_2 = \cos(-\psi_A)\,, \quad C_3 = \cos(-\omega_A)\,, \quad C_4 = \cos(\chi_A)\,, \tag{9.36}$$

then, the overall four-angle precession matrix $\mathbb{P}(t)$ can also be written as follows

$$\mathbb{P}(t) = \tag{9.37}$$
$$\begin{bmatrix} C_4C_2-S_2S_4C_3 & C_4S_2C_1+S_4C_3C_2C_1-S_1S_4S_3 & C_4S_2S_1+S_4C_3C_2S_1+C_1S_4S_3 \\ -S_4C_2-S_2C_4C_3 & -S_4S_2C_1+C_4C_3C_2C_1-S_1C_4S_3 & -S_4S_2S_1+C_4C_3C_2S_1+C_1C_4S_3 \\ S_2S_3 & -S_3C_2C_1-S_1C_3 & -S_3C_2S_1+C_3C_1 \end{bmatrix}$$

Existing applications that use the 3-angle precession formulation of Newcomb and Lieske can be easily modified for the IAU 2000A precession, by replacing the current polynomials for the angles ζ_A, z_A, and θ_A with the following

$$\begin{aligned} \zeta_A &= 2.650545 + 2306.083227\,T + 0.2988499\,T^2 + 0.01801828\,T^3 \\ &\quad -0.000005971\,T^4 - 0.0000003173\,T^5\,, \end{aligned} \tag{9.38}$$

$$\begin{aligned} z_A &= -2.650545 + 2306.077181\,T + 1.0927348\,T^2 + 0.01826837\,T^3 \\ &\quad -0.000028596\,T^4 - 0.0000002904\,T^5\,, \end{aligned} \tag{9.39}$$

$$\begin{aligned} \theta_A &= 2004.191903\,T - 0.4294934\,T^2 - 0.04182264\,T^3 \\ &\quad -0.000007089\,T^4 - 0.0000001274\,T^5\,, \end{aligned} \tag{9.40}$$

The 3-angle precession matrix is

$$\mathbb{P}(t) = \mathbb{R}_3(-z_A)\,\mathbb{R}_2(\theta_A)\,\mathbb{R}_3(-\zeta_A)\,, \tag{9.41}$$

but any existing correct construction of $\mathbb{P}$ using these three angles can still be used.

The expression for the mean obliquity of the ecliptic (the angle between the mean equator and ecliptic, or, equivalently, between the ecliptic pole and mean celestial pole of date) is

$$\begin{aligned} \epsilon &= \epsilon_0 - 46.836769\,T - 0.0001831\,T^2 + 0.00200340\,T^3 \\ &\quad -0.000000576\,T^4 - 0.0000000434\,T^5\,, \end{aligned} \tag{9.42}$$

where, as stated above, $\epsilon_0 = 84381.406$ arcseconds. This expression arises from the precession formulation but is actually used only for nutation. Almost all of the obliquity rate – the term linear in T in equation (9.42) – is due to the precession of the ecliptic.

Where high accuracy is not required, the precession between two dates, t_1 and t_2, not too far apart (i.e., where $|t_2 - t_1| \ll 1$ century), can be approximated using the rates of change of right ascension and declination with respect to the mean equator and equinox of date. These rates are respectively

$$m \simeq 4612.16 + 2.78\,T \tag{9.43}$$

$$n \simeq 2004.19 - 0.86\,T \tag{9.44}$$

where the values are in arcseconds per century, and T is the number of centuries between J2000.0 and the midpoint of t_1 and t_2. If the dates are expressed as Julian dates, $T = [(t_1 + t_2)/2 - 2451545.0]/36525$. Then, denoting the celestial coordinates at t_1 by (α_1, δ_1) and those at t_2 by (α_2, δ_2),

$$\begin{aligned}\alpha_2 &\simeq \alpha_1 + \tau\,(m + n\,\sin\alpha_1\,\tan\delta_1)\\ \delta_2 &\simeq \delta_1 + \tau\,(n\,\cos\alpha_1)\end{aligned} \tag{9.45}$$

where $\tau = t_2 - t_1$, expressed in centuries. These formulas deteriorate in accuracy at high (or low) declinations and should not be used at all for coordinates close to the celestial poles (how close depends on the accuracy requirement and the value of τ) .

9.6.6 Formulas for Nutation

Nutation is conventionally expressed as two small angles, $\Delta\psi$, the nutation in longitude, and $\Delta\epsilon$, the nutation in obliquity. These angles are measured in the ecliptic system of date, which is developed as part of the precession formulation. The angle $\Delta\psi$ is the small change in the position of the equinox along the ecliptic due to nutation, so the effect of nutation on the ecliptic coordinates of a fixed point in the sky is simply to add $\Delta\psi$ to its ecliptic longitude. The angle $\Delta\epsilon$ is the small change in the obliquity of the ecliptic due to nutation. The *true obliquity* of date is $\epsilon' = \epsilon + \Delta\epsilon$. Nutation in obliquity reflects the orientation of the equator in space and does not affect the ecliptic coordinates of a fixed point on the sky.

The angles $\Delta\psi$ and $\Delta\epsilon$ can also be thought of as small shifts in the position of the celestial pole (CIP) with respect to the ecliptic and mean equinox of date. In that coordinate system, and assuming positive values for $\Delta\psi$ and $\Delta\epsilon$, the nutation in longitude shifts the celestial pole westward on the sky by the angle $\Delta\psi\,\sin\epsilon$, decreasing the pole's mean ecliptic longitude by $\Delta\psi$. Nutation in obliquity moves the celestial pole further from the ecliptic pole, i.e., southward in ecliptic coordinates, by $\Delta\epsilon$. Negative values of $\Delta\psi$ and $\Delta\epsilon$ move the pole eastward and northward in ecliptic coordinates.

The effect of nutation on the equatorial coordinates (α,δ) of a fixed point in the sky is more complex and is best dealt with through the action of the nutation matrix, $\mathbb{N}(t)$, on the equatorial position vector, $\boldsymbol{r}_{\text{mean}}(t)$. Where high accuracy is not required, formulas that directly give the changes to α and δ as a function of $\Delta\psi$ and $\Delta\epsilon$ are available and are given at the end of this subsection.

The values of $\Delta\psi$ and $\Delta\epsilon$ are obtained by evaluating rather lengthy trigonometric series, of the general form

$$\Delta\psi = \sum_{i=1}^{N}\left[(S_i + \dot{S}_i T)\sin\,\Phi_i + C_i'\cos\,\Phi_i\right] \tag{9.46a}$$

$$\Delta\epsilon = \sum_{i=1}^{N}\left[(C_i + \dot{C}_i T)\cos\,\Phi_i + S_i'\sin\,\Phi_i\right]\,, \tag{9.46b}$$

where, in each term

$$\Phi_i = \sum_{j=1}^{K} M_{i,j}\, \phi_j(T)\,, \tag{9.47}$$

with $M_{i,j}$ being well-determined numerical multipliers listed in the the IAU 2000A model, for which N=1365 and K=14. The 14 $\phi_j(T)$ are the fundamental arguments, which are, except for one, orbital angles. The main time dependence of the nutation series enters through these arguments. The expressions given below are all taken from the paper by Simon *et al.* [1994], and all coefficients are in arcseconds.

The first eight fundamental arguments are the mean heliocentric ecliptic longitudes of the planets Mercury through Neptune

$$\begin{aligned}
\phi_1 &= 908103.259872 + 538101628.688982\, T \\
\phi_2 &= 655127.283060 + 210664136.433548\, T \\
\phi_3 &= 361679.244588 + 129597742.283429\, T \\
\phi_4 &= 1279558.798488 + 68905077.493988\, T \\
\phi_5 &= 123665.467464 + 10925660.377991\, T \\
\phi_6 &= 180278.799480 + 4399609.855732\, T \\
\phi_7 &= 1130598.018396 + 1542481.193933\, T \\
\phi_8 &= 1095655.195728 + 786550.320744\, T
\end{aligned} \tag{9.48}$$

In all of these expressions, T is the number of Julian centuries of TDB since 2000 Jan 1, 12^{h} TDB (or, with negligible error, the number of Julian centuries of TT since J2000.0). In some implementations it may be necessary to reduce the resulting angles, which are expressed in arcseconds, to radians in the range 0-2π. The ninth argument is an approximation to the general precession in longitude

$$\phi_9 = 5028.8200\, T + 1.112022\, T^2 \tag{9.49}$$

The last five arguments are the same fundamental luni-solar arguments used in previous nutation theories, but with updated expresssions. They are, respectively, l, the mean anomaly of the Moon; l', the mean anomaly of the Sun; F, the mean argument of latitude of the Moon; D, the mean elongation of the Moon from the Sun, and Ω, the mean longitude of the Moon's mean ascending node

$$\begin{aligned}
\phi_{10} &= l = 485868.249036 + 1717915923.2178\, T + 31.8792\, T^2 + 0.051635\, T^3 - 0.00024470\, T^4 \\
\phi_{11} &= l' = 1287104.79305 + 129596581.0481\, T - 0.5532\, T^2 + 0.000136\, T^3 - 0.00001149\, T^4 \\
\phi_{12} &= F = 335779.526232 + 1739527262.8478\, T - 12.7512\, T^2 - 0.001037\, T^3 + 0.00000417\, T^4 \\
\phi_{13} &= D = 1072260.70369 + 1602961601.2090\, T - 6.3706\, T^2 + 0.006593\, T^3 - 0.00003169\, T^4 \\
\phi_{14} &= \Omega = 450160.398036 - 6962890.5431\, T + 7.4722\, T^2 + 0.007702\, T^3 - 0.00005939\, T^4
\end{aligned}$$

The first step in evaluating the series for nutation for a given date is to compute the values of all 14 fundamental arguments for the date of interest. This is done only once. Then the nutation terms are evaluated one by one. For each term i, first compute Φ_i according to equation (9.47), using the 14 integer multipliers, $M_{i,j}$, listed for the term; i.e., sum over $M_{i,j} \times \phi_j$ (where j=1-14). Then the cosine and sine components for the term can be evaluated, as per equation (9.46), using the listed values of the coefficients S_i, $\dot{S}_i$, C'_i, C_i, $\dot{C}_i$, and S'_i for the term. Generally it is good practice to sum the terms from smallest to largest to preserve precision in the sums.

Table 9.2 (at the very end of this section) lists 40 sample terms from the IAU 2000A nutation series. The entire nutation series is listed at the end of USNO Circular 179 [Kaplan, 2005b]. About the first half of the series consists of lunisolar terms, which depend only on l, l', F, D, and Ω (= ϕ_{10} to ϕ_{14}). In all of these terms, the first nine multipliers are all zero. The generally smaller planetary terms comprise the remainder of the series. As an example of how the individual terms are computed according to equations (9.46) and (9.47), term 6 would be evaluated

$$\begin{aligned}
\Delta\psi_6 &= (-0.0516821 + 0.0001226\,T)\sin\Phi_6 - 0.0000524\cos\Phi_6 \\
\Delta\epsilon_6 &= (\ \ 0.0224386 - 0.0000667\,T)\cos\Phi_6 - 0.0000174\sin\Phi_6 \\
&\quad \text{where}\quad \Phi_6 = \phi_{11} + 2\phi_{12} - 2\phi_{13} + 2\phi_{14}
\end{aligned}$$

since $M_{6,1}$ through $M_{6,10}$ are zero, and only ϕ_{11} through ϕ_{14} are therefore relevant for this term. It is assumed that all the ϕ_j have been pre-computed (for all terms) using the appropriate value of T for the date and time of interest. A printed version of a 1365-term nutation series is obviously not the most convenient form for computation. As noted earlier, the series is available as a pair of plain-text computer files at McCarthy and Petit [2003a], and the SOFA and NOVAS software packages (IAU SOFA Center [2009], Kaplan *et al.* [2009]) include subroutines for evaluating it. There are also shorter series available where the highest precision is not required. The IERS web site provides, in addition to the full IAU 2000A series, an IAU 2000B series, which has only 77 terms and duplicates the IAU 2000A results to within a milliarcsecond for input times between 1995 and 2050. NOVAS also provides a subroutine that evaluates a truncated series, with 488 terms, that duplicates the full series to 0.1 milliarcsecond accuracy between 1700 and 2300.

Once the nutation series has been evaluated and the values of $\Delta\psi$ and $\Delta\epsilon$ are available, the nutation matrix can be constructed. The nutation matrix is simply $\mathbb{N}(t) = \mathbb{R}_1(-\epsilon')\,\mathbb{R}_3(-\Delta\psi)\,\mathbb{R}_1(\epsilon)$, where, again, $\mathbb{R}_1$ and $\mathbb{R}_3$ are standard rotations about the x and z axes, respectively (see "Abbreviations and Symbols Frequently Used" for precise definitions), and $\epsilon' = \epsilon + \Delta\epsilon$ is the true obliquity (compute ϵ using equation (9.42)). This formulation is comprised of

1) A rotation from the mean equator and equinox of t to the mean ecliptic and equinox of t. This is simply a rotation around the x-axis (the direction toward the mean equinox of t) by the angle ϵ, the mean obliquity of t. After the rotation, the fundamental plane is the ecliptic of t.
2) A rotation around the new z-axis (the direction toward the ecliptic pole of t) by the angle $-\Delta\psi$, the amount of nutation in longitude at t. After the rotation, the new x-axis is in the direction of the true equinox of t.

3) A rotation around the new x-axis (the direction toward the true equinox of t) by the angle $-\epsilon'$, the true obliquity of t. After the rotation, the fundamental plane is the true equator of t, orthogonal to the computed position of the CIP at t.

If one let

$$\begin{aligned} S_1 &= \sin(\epsilon) \\ S_2 &= \sin(-\Delta\psi) \\ S_3 &= \sin(-\epsilon-\Delta\epsilon) \end{aligned}$$

$$\begin{aligned} C_1 &= \cos(\epsilon) \\ C_2 &= \cos(-\Delta\psi) \\ C_3 &= \cos(-\epsilon-\Delta\epsilon) \end{aligned} \tag{9.51}$$

then the nutation matrix can also be written

$$\mathbb{N}(t) = \begin{bmatrix} C_2 & S_2C_1 & S_2S_1 \\ -S_2C_3 & C_3C_2C_1 - S_1S_3 & C_3C_2S_1 + C_1S_3 \\ S_2S_3 & -S_3C_2C_1 - S_1C_3 & -S_3C_2S_1 + C_3C_1 \end{bmatrix} \tag{9.52}$$

Where high accuracy is not required, coordinates corrected for nutation in right ascension and declination can be obtained from

$$\begin{aligned} \alpha_{\mathrm{t}} &\approx \alpha_{\mathrm{m}} + \Delta\psi\,(\cos\epsilon' + \sin\epsilon' \sin\alpha_{\mathrm{m}} \tan\delta_{\mathrm{m}}) - \Delta\epsilon\,\cos\alpha_{\mathrm{m}} \tan\delta_{\mathrm{m}} \\ \delta_{\mathrm{t}} &\approx \delta_{\mathrm{m}} + \Delta\psi\,\sin\epsilon' \cos\alpha_{\mathrm{m}} + \Delta\epsilon\,\sin\alpha_{\mathrm{m}} \end{aligned} \tag{9.53}$$

where $(\alpha_{\mathrm{m}}, \delta_{\mathrm{m}})$ are coordinates with respect to the mean equator and equinox of date (precession only), $(\alpha_{\mathrm{t}}, \delta_{\mathrm{t}})$ are the corresponding coordinates with respect to the true equator and equinox of date (precession + nutation), and ϵ' is the true obliquity. Note the $\tan\delta_{\mathrm{m}}$ factor in right ascension that makes these formulas unsuitable for use close to the celestial poles.

The traditional formula for the equation of the equinoxes (the difference between apparent and mean sidereal time) is $\Delta\psi\cos\epsilon'$, but in recent years this has been superceded by the more accurate version given in equation (9.16).

9.6.7
Alternative Combined Transformation

The following matrix, $\mathbb{C}(t)$, combines precession, nutation, and frame bias and is used to transform vectors from the GCRS to the Celestial Intermediate Reference System (CIRS). The CIRS is defined by the equator of the CIP and an origin of right ascension called the Celestial Intermediate Origin (CIO). The CIO is discussed extensively in section 9.7. There, the CIRS is symbolized E_σ; it is analogous to the true equator and equinox of date, but with a different right ascension origin.

The matrix $\mathbb{C}(t)$ is used in the sense

$$\boldsymbol{r}_{\text{CIRS}} = \mathbb{C}(t)\, \boldsymbol{r}_{\text{GCRS}} \tag{9.54}$$

and the components of C(t), as given in the McCarthy and Petit [2004] and *The Astronomical Almanac* [2010], are

$$\mathbb{C}(t) = \mathbb{R}_3(-s) \begin{bmatrix} 1 - bX^2 & -bXY & -X \\ -bXY & 1 - bY^2 & -Y \\ X & Y & 1 - b(X^2 + Y^2) \end{bmatrix} \tag{9.55}$$

where X and Y are the dimensionless coordinates of the CIP in the GCRS (unit vector components), $b = 1/(1 + Z)$, $Z = \sqrt{1 - X^2 - Y^2}$, and s is the *CIO locator*, a small angle described in section 9.7. All of these quantities are functions of time. $\mathbb{R}_3$ is a standard rotation around the z axis; see appendix , "Abbreviations and Symbols Frequently Used" for a precise definition.

9.6.8 Observational Corrections to Precession-Nutation

The IERS still publishes daily values of the observed celestial pole offsets, despite the vast improvement to the pole position predictions given by the IAU 2000A precession-nutation model. The offsets now have magnitudes generally less than 1 mas. The fact that they are non-zero is due in part to an effect of unpredictable amplitude and phase called the free core nutation (or nearly diurnal free wobble), caused by the rotation of the fluid core of the Earth inside the ellipsoidal cavity that it occupies. The free core nutation appears as a very small nutation component with a period of about 430 days (see Figure 9.6). Any other effects not accounted for in the adopted precession-nutation model will also appear in the celestial pole offsets. In any event, the celestial pole offsets are now so small that many users may now decide to ignore them. However, it is worth noting again that, by definition, the Celestial Intermediate Pole (CIP) includes these observed offsets.

The IERS now publishes celestial pole offsets with respect to the IAU 2000A precession-nutation model only as dX and dY – corrections to the pole's computed unit vector components X and Y in the GCRS (see equation (9.29) and following notes). The IERS pole offsets are published in units of milliarcseconds but they can be converted to dimensionless quantities by dividing them by the number of milliarcseconds in one radian, 206264806.247... . Then, the observationally corrected values of X and Y are

$$X_{\text{cor}} = X + dX \qquad \text{and} \qquad Y_{\text{cor}} = Y + dY \tag{9.56}$$

The corrected values, expressed as dimensionless quantities (unit vector components), are used, e.g., in the matrix $\mathbb{C}$ given in sections 9.6.7 and 9.7.5.6. That is, in equations (9.55) and (9.77), assume $X = X_{\text{cor}}$ and $Y = Y_{\text{cor}}$.

The ecliptic-based pole celestial offsets, $d\psi$ and $d\epsilon$, which are used to correct the nutation theory's output angles $\Delta\psi$ and $\Delta\epsilon$, are no longer supplied (actually, they are supplied but only for the old pre-2000 precession-nutation model). Software that has not been coded to use X and Y directly – which includes all software developed prior to 2003 – will need a front-end to convert the IERS dX and dY values to $d\psi$ and $d\epsilon$. A derivation of a conversion algorithm and several options for its implementation (depending on the accuracy desired) are given by Kaplan [2003]. Succinctly, given dimensionless dX and dY values for a given date t, let

$$\begin{bmatrix} dX' \\ dY' \\ dZ' \end{bmatrix} = \mathbb{P}(t) \begin{bmatrix} dX \\ dY \\ dZ \end{bmatrix} \tag{9.57}$$

where $\mathbb{P}(t)$ is the precession matrix from J2000.0 to date t, and one can set $dZ = 0$ in this approximation, which holds for only a few centuries around J2000.0. Then one computes the ecliptic-based correction angles in radians using

$$d\psi = dX' / \sin\epsilon \qquad \text{and} \qquad d\epsilon = dY' \tag{9.58}$$

where ϵ is the mean obliquity of the ecliptic of date t, computed according to equation (9.42). The observationally corrected values of $\Delta\psi$ and $\Delta\epsilon$ are obtained simply by adding $d\psi$ and $d\epsilon$, respectively

$$\Delta\psi_{\rm cor} = \Delta\psi + d\psi \qquad \text{and} \qquad \Delta\epsilon_{\rm cor} = \Delta\epsilon + d\epsilon \tag{9.59}$$

where care must be taken to ensure that all angles are expressed in the same units. The corrected values are used in forming the nutation matrix $\mathbb{N}(t)$ and in other nutation-related expressions. That is, in equations (9.51) and (9.53), assume $\Delta\psi = \Delta\psi_{\rm cor}$ and $\Delta\epsilon = \Delta\epsilon_{\rm cor}$. At the same time, the corrected value of $\Delta\psi$ should be used in forming the equation of the equinoxes using equation (9.16).

The following table 9.2 lists 40 terms from the IAU2000A nutation series adopted by the IAU, developed by Buffett *et al.* [2002] (MHB): the first 20 lunisolar terms and the first 20 planetary terms (these terms are among the largest in the series). The entire series, in the same format, is printed in Kaplan [2005b]. The series is also available from the IERS as a pair of plain-text computer files at McCarthy and Petit [2003a], although the arrangement of the columns differs from what is presented here. The IERS also provides a Fortran subroutine for evaluating the nutation series, written by P. Wallace, at McCarthy and Petit [2003c]. The NOVAS software package includes this subroutine, and the SOFA package contains the same code in a subroutine of a different name. There are also subroutines available that evaluate only a subset of the series terms for applications that do not require the highest accuracy.

There are 1365 terms in the complete series, 678 lunisolar terms and 687 planetary terms. (The term numbers are arbitrary and are not involved in the computation.) In the lunisolar terms, the only fundamental argument multipliers that are non-zero are $M_{i,10}$ through $M_{i,14}$, corresponding to the arguments l, l', F, D, and Ω, respectively. In the planetary terms, there are no rates of change of the coefficients, i.e., $\dot{S}_i$ and $\dot{C}_i$ are zero.

The formulas for evaluating the series are given in section 9.6.6; see equations (9.46)-(9.47) and the text that follows them.

Table 9.2 Nutation Series – Sample Terms.

Term	Fundamental Argument Multipliers $M_{i,j}$														$\Delta\psi$ Coefficients		
i	j= 1	2	3	4	5	6	7	8	9	10	11	12	13	14	S_i	$\dot{S}_i$	C'_i
															″	″	″
1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-17.2064161	-0.0174666	0.0033386
2	0	0	0	0	0	0	0	0	0	0	0	2	-2	2	-1.3170906	-0.0001675	-0.0013696
3	0	0	0	0	0	0	0	0	0	0	0	2	0	2	-0.2276413	-0.0000234	0.0002796
4	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0.2074554	0.0000207	-0.0000698
5	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0.1475877	-0.0003633	0.0011817
6	0	0	0	0	0	0	0	0	0	0	1	2	-2	2	-0.0516821	0.0001226	-0.0000524
7	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0.0711159	0.0000073	-0.0000872
8	0	0	0	0	0	0	0	0	0	0	0	2	0	1	-0.0387298	-0.0000367	0.0000380
9	0	0	0	0	0	0	0	0	0	1	0	2	0	2	-0.0301461	-0.0000036	0.0000816
10	0	0	0	0	0	0	0	0	0	0	-1	2	-2	2	0.0215829	-0.0000494	0.0000111
11	0	0	0	0	0	0	0	0	0	0	0	2	-2	1	0.0128227	0.0000137	0.0000181
12	0	0	0	0	0	0	0	0	0	-1	0	2	0	2	0.0123457	0.0000011	0.0000019
13	0	0	0	0	0	0	0	0	0	-1	0	0	2	0	0.0156994	0.0000010	-0.0000168
14	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0.0063110	0.0000063	0.0000027
15	0	0	0	0	0	0	0	0	0	-1	0	0	0	1	-0.0057976	-0.0000063	-0.0000189
16	0	0	0	0	0	0	0	0	0	-1	0	2	2	2	-0.0059641	-0.0000011	0.0000149
17	0	0	0	0	0	0	0	0	0	1	0	2	0	1	-0.0051613	-0.0000042	0.0000129
18	0	0	0	0	0	0	0	0	0	-2	0	2	0	1	0.0045893	0.0000050	0.0000031
19	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0.0063384	0.0000011	-0.0000150
20	0	0	0	0	0	0	0	0	0	0	0	2	2	2	-0.0038571	-0.0000001	0.0000158
679	0	0	8	-16	4	5	0	0	0	0	0	0	0	0	0.0001440	0.	0.0000000
680	0	0	-8	16	-4	-5	0	0	2	0	0	0	0	0	0.0000056	0.	-0.0000117
681	0	0	8	-16	4	5	0	0	2	0	0	0	0	0	0.0000125	0.	-0.0000043
682	0	0	0	0	0	0	-1	2	2	0	0	0	0	0	0.0000000	0.	0.0000005
683	0	0	-4	8	-1	-5	0	0	2	0	0	0	0	0	0.0000003	0.	-0.0000007
684	0	0	4	-8	3	0	0	0	1	0	0	0	0	0	0.0000003	0.	0.0000000
685	0	0	3	-8	3	0	0	0	0	0	0	1	-1	1	-0.0000114	0.	0.0000000
686	0	10	-3	0	0	0	0	0	0	-1	0	0	0	0	-0.0000219	0.	0.0000089
687	0	0	0	0	-2	6	-3	0	2	0	0	0	0	0	-0.0000003	0.	0.0000000
688	0	0	4	-8	3	0	0	0	0	0	0	0	0	0	-0.0000462	0.	0.0001604
689	0	0	-5	8	-3	0	0	0	0	0	0	1	-1	1	0.0000099	0.	0.0000000
690	0	0	-4	8	-3	0	0	0	1	0	0	0	0	0	-0.0000003	0.	0.0000000
691	0	0	4	-8	1	5	0	0	2	0	0	0	0	0	0.0000000	0.	0.0000006
692	0	-5	6	4	0	0	0	0	2	0	0	0	0	0	0.0000003	0.	0.0000000
693	0	0	0	0	2	-5	0	0	2	0	0	0	0	0	-0.0000012	0.	0.0000000
694	0	0	0	0	2	-5	0	0	1	0	0	0	0	0	0.0000014	0.	-0.0000218
695	0	0	-1	0	2	-5	0	0	0	0	0	1	-1	1	0.0000031	0.	-0.0000481
696	0	0	0	0	2	-5	0	0	0	0	0	0	0	0	-0.0000491	0.	0.0000128
697	0	0	-1	0	-2	5	0	0	0	0	0	1	-1	1	-0.0003084	0.	0.0005123
698	0	0	0	0	-2	5	0	0	1	0	0	0	0	0	-0.0001444	0.	0.0002409

9.7 Modeling the Earth's Rotation

IAU Resolution 1.8 of 2000 establishes two new reference points in the plane of the moving (instantaneous) equator for the measurement of Earth rotation: the point on the geocentric celestial sphere is called the Celestial Intermediate Origin (CIO) and the point on the surface of the Earth is called the Terrestrial Intermediate Origin (TIO). The CIO and TIO are specific examples of a concept called a *non-rotating origin* that was first described by Guinot [1979, 1981].

The Earth Rotation Angle, θ, is the geocentric angle between the directions of the CIO and TIO, and provides a new way to represent the rotation of the Earth in the transformation from terrestrial to celestial systems or vice versa. Traditionally, Greenwich sidereal time, which is the hour angle of the equinox with respect to the Greenwich meridian, has served this purpose. The CIO and TIO are defined in such a way that θ is a linear function of Universal Time (UT1) and independent of the Earth's precession and nutation; it is a direct measure of the rotational angle of the Earth around the Celestial Intermediate Pole (the Celestial Intermediate Pole is described in section 9.6). Since none of these statements holds for sidereal time, the scheme based on the CIO, TIO, and θ represents a simplification of the way the rotation of the Earth is treated. In particular, the transformation between Earth-fixed and space-fixed reference systems can now be specified by three rotation matrices that are independent of each other: one for polar motion, one for "pure" rotation (i.e., θ), and one for precession-nutation.

The recent IAU resolutions do not eliminate sidereal time or the use of the equinox as a fundamental reference point. Instead, the resolutions establish an alternative way of dealing with Earth rotation. The comparison between the two schemes can be illuminating. For example, the CIO helps to clarify the relationship between sidereal time and the Earth's rotation, since θ is now the "fast term" in the formula for sidereal time as a function of UT1. The remaining terms comprise the *equation of the origins* and represent the accumulated amount of precession and nutation along the equator as a function of time. The equation of the origins is the length of the arc between the equinox and the CIO.

9.7.1 A Messy Business

In the computation of the positions of celestial objects with respect to an Earth-fixed system – or, equivalently, in the transformation between terrestrial and celestial coordinate systems – sidereal time has conventionally represented the Earth's rotation about its axis. For example, the hour angle of an object is simply the local apparent sidereal time minus the object's apparent right ascension with respect to the true equator and equinox of date (see section 9.3.6.2). Once its hour angle and declination are available, the object's zenith distance and azimuth, or its coordinates with respect to some ground-based instrumental system, can be easily obtained. The same result can be accomplished by a direct transformation between the celestial and terrestrial coordinate systems, conventionally represented by a series of rotation matrices, one each for precession, nutation, sidereal time, and polar motion.

Yet there is something untidy about these procedures. The computation of apparent sidereal time mixes quantities related to Earth rotation, precession, and nutation (see equations (9.11)-(9.16)). Because sidereal time is defined as the hour angle of the equinox, the precession of the equinox in right ascension must be a part of the expression for sidereal time (the terms in parentheses in equation (9.13)), and the mean sidereal day is thereby shorter than the rotation period of the Earth by about $0\overset{s}{.}008$. Nutation also appears, in the equation of the equinoxes (equations (9.15) & (9.16)). The result is that in the computation of hour angle, precession and nutation enter twice: once in the sidereal time formula and again in the computation of the star's apparent right ascension; the two contributions cancel for stars on the equator. Similarly, in the transformation between the celestial and terrestrial coordinate systems, precession and nutation each enter into two of the rotation matrices, and none of the matrices represents Earth rotation alone.

A consequence of this way of doing things is that whenever improvements are made to the theory of precession, the numerical coefficients in the expression for sidereal time must also change. This was not an issue for most of the twentieth century, since no adjustments were made to the standard precession algorithm, and the expression for mean sidereal time derived from Newcomb's developments was used without much thought given to the matter. It was the change to this expression, necessitated by the adjustment of the precession constant in the IAU (1976) System of Astronomical Constants, that first motivated the search for a fundamental change of procedure. At about the same time, new high-precision observing techniques, such as VLBI and lunar laser ranging, were being applied to the study of all components of the Earth's rotation, and a review of the basic algorithms seemed appropriate. In particular, there was interest in constructing a new geometrical picture and set of expressions for the orientation of the Earth as a function of time that would cleanly separate the effects of rotation, precession and nutation, and polar motion. Furthermore, since VLBI is not sensitive to the equinox, a procedure that used a different reference point seemed desirable.

To bring the Earth's rotation period explicitly into the terrestrial-celestial transformation, one must define an angle of rotation about the Earth's axis. As described in section 9.6, what one specifically means by "the Earth's axis" is the line through the geocenter in the direction of the Celestial Intermediate Pole (CIP). The angle of rotation about this axis must be measured with respect to some agreed-upon direction in space. Since the CIP moves small amounts during one rotation of the Earth (~0.1 arcsecond with respect to the stars and ~0.005 arcsecond with respect to the Earth's crust), the reference direction cannot be simply a fixed vector or plane in inertial space. What one needs is an appropriate azimuthal[27)] origin – a point in the moving equatorial plane, which is orthogonal to the CIP.

27) The word "azimuthal" is used in its general sense, referring to an angle measured about the z-axis of a coordinate system.

9.7.2
Non-Rotating Origins

The reference point that one defines must be such that the rate of change of the Earth's rotation angle, measured with respect to this point, is the angular velocity of the Earth about the CIP. As the CIP moves, the point must move to remain in the equatorial plane; but the point's motion must be such that the measured rotation angle is not contaminated by some component of the motion of the CIP itself.

The concept of a "non-rotating origin" (NRO) on the equator can be applied to any rotating body. The NRO idea was first described by Bernard Guinot [Guinot, 1979, 1981] and further developed by Nicole Capitaine and collaborators [Capitaine, 1990, 2000; Capitaine and Chollet, 1991; Capitaine *et al.*, 2000, 1986]. The condition on the motion of such a point is simple: as the equator moves, the point's instantaneous motion must always be orthogonal to the equator. That is, the point's motion at some time t must be directly toward or away from the position of the pole of rotation at t. Any other motion of the point would have a component *around* the axis/pole and would thus introduce a spurious rate into the measurement of the rotation angle of the body as a function of time. The point is not unique; any arbitrary point on the moving equator could be made to move in the prescribed manner. For the Earth, the difference between the motion of a non-rotating origin and that of the equinox on the geocentric celestial sphere is illustrated in figure 9.7.

As illustrated in the figure, the motion of the non-rotating origin, σ, is always orthogonal to the equator, whereas the equinox has a motion along the equator (the precession in right ascension). How does one specify the location of a non-rotating origin? There are three possibilities, outlined in the Formulas subsection below. In the most straightforward scheme, one simply uses the GCRS right ascension of σ obtained from a numerical integration (the GCRS is the "geocentric ICRS"). Alternatively, the position of σ can be defined by a quantity, s, that is the difference between the lengths of two arcs on the celestial sphere. Finally, one can specify the location of σ with respect to the equinox, Υ: the equatorial arc $\overline{\Upsilon\sigma}$ is called the *equation of the origins*. Whatever geometry is used, the position of σ ultimately depends on an integral over time, because the defining property of σ is its *motion* – not a static geometrical relationship with other points or planes. The integral involved is fairly simple and depends only on the coordinates of the pole and their derivatives with respect to time. The initial point for the integration can be any point on the moving equator at any time t_0.

So far one has discussed a non-rotating origin only on the celestial sphere, required because of the movement of the CIP in a space-fixed reference system. But there is a corresponding situation on the surface of the Earth. The CIP has motions in both the celestial and terrestrial reference systems. Its motion in the celestial system is precession-nutation and its motion in the terrestrial system is polar motion, or wobble. From the point of view of a conventional geodetic coordinate system "attached" to the surface of the Earth (i.e., defined by the specified coordinates of a group of stations), the CIP wanders around near the geodetic pole in a quasi-circular motion with an amplitude of about 10 meters (0.3 arcsec) and two primary periods, 12 and

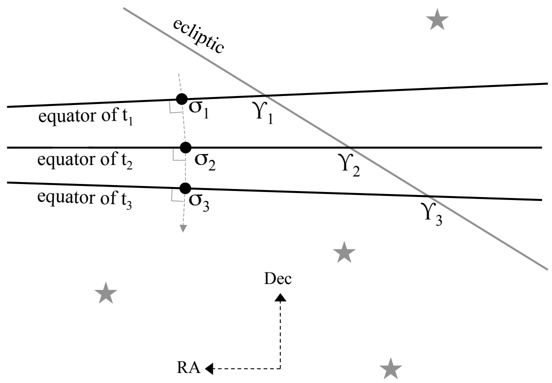

Figure 9.7 Motion of a non-rotating origin, σ, compared with that of the true equinox, Υ. "Snapshots" of the positions of the points are shown at three successive times, t_1, t_2, and t_3. The positions are shown with respect to a geocentric reference system that has no systematic rotation with respect to a set of extragalactic objects. The ecliptic is shown in the figure as fixed, although it, too, has a small motion in inertial space.

14 months. Thus the equator of the CIP has a slight quasi-annual wobble around the geodetic equator. Actually, it is better thought of in the opposite sense: the geodetic equator has a slight wobble with respect to the equator of the CIP. That point of view makes it is a little clearer why a simple "stake in the ground" at the geodetic equator would not be suitable for measuring the Earth rotation angle around the CIP. The situation is orders of magnitude less troublesome than that on the celestial sphere, but for completeness (and very precise applications) it is appropriate to define a terrestrial non-rotating origin, designated ϖ. It stays on the CIP equator, and assuming that the current amplitude of polar motion remains approximately constant, ϖ will bob north and south by about 10 m in geodetic latitude every year or so and will have a secular eastward motion in longitude of about 1.5 mm/cen. The exact motion of ϖ depends, of course, on what polar motion, which is unpredictable, actually turns out to be.

The two non-rotating origins, σ and ϖ, are called the Celestial Intermediate Origin (CIO) and the Terrestrial Intermediate Origin (TIO). Both lie in the same plane – the equator of the CIP. The Earth Rotation Angle, θ, is defined as the geocentric angle between these two points. The angle θ is a linear function of Universal Time (UT1). The formula, given in the note to resolution B1.8 of 2000, is simply $\theta = 2\pi\,(0.7790572732640 + 1.00273781191135448\ D_U)$, where D_U is the number of UT1 days from JD 2451545.0 UT1. The formula assumes a constant angular velocity of the Earth: no attempt is made to model its secular decrease due to tidal friction, monthly tidal variations, changes due to the exchange of angular momentum between the atmosphere and the solid Earth, and other phenomena. These effects will be reflected in the time series of UT1-UTC or ΔT values (see section 9.3) derived from precise observations.

The expression given above for θ is now the "fast term" in the formula for mean sidereal time; see equation (9.13). It accounts for the rotation of the Earth, while the other terms account for the motion of the equinox along the equator due to precession.

The plane defined by the geocenter, the CIP, and TIO is called the *TIO meridian*. For most ordinary astronomical purposes the TIO meridian can be considered to be identical to what is often referred to as the Greenwich meridian. The movement of this meridian with respect to a conventional geodetic system is important only for the most precise astrometric/geodetic applications. It is worth noting that the TIO meridian, and the zero-longitude meridians of modern geodetic systems, are about 100 m from the old transit circle at Greenwich [Dillon *et al.*, 1977; Gebel and Matthews, 1971]. The term "Greenwich meridian" has ceased to have a technical meaning in the context of precise geodesy – despite the nice line in the sidewalk at the old Greenwich observatory. This has become obvious to tourists carrying GPS receivers!

9.7.3
The Path of the CIO on the Sky

If one takes the epoch J2000.0 as the starting epoch for evaluating the integral that provides the position of the CIO, the only mathematical requirement for the initial point is that it lie on the instantaneous (true) equator of that date – its position along the equator is arbitrary. By convention, however, the initial position of the CIO on the instantaneous equator of J2000.0 is set so that equinox-based and CIO-based computations of Earth rotation yield the same answers; one wants the hour angle of a given celestial object to be the same, as a function of UT1 (or UTC), no matter how the calculation is done. For this to happen, the position of the CIO of J2000.0 must be at GCRS right ascension $0° \, 0' \, 00.''002012$. This is about 12.8 arcseconds west of the true equinox of that date.

Since the CIO rides on the instantaneous equator, its primary motion over the next few millenia is southward at the rate of precession in declination, initially 2004 arcseconds per century. Its rate of southward motion is modulated (but never reversed) by the nutation periodicities. Its motion in GCRS right ascension is orders of magnitude less rapid; remember that the CIO has no motion along the instantaneous equator, and the instantaneous equator of J2000.0 is nearly co-planar with the GCRS equator (xy-plane). The motion of the CIO in GCRS right ascension over the next few millenia is dominated by a term proportional to t^3; the GCRS right ascension of the CIO at the beginning of year 2100 is only $0.''068$; at the beginning of 2200 it is $0.''573$; and at the beginning of 2300 it is $1.''941$. Nutation does produce a very slight wobble in the CIO's right ascension, but the influence of the nutation terms is suppressed by several orders of magnitude relative to their effect on the position of the pole. One can say, therefore, that to within a few arcseconds error, the path of the CIO on the celestial sphere over the next few centuries is nearly a straight line southward along the GCRS α=0 hour circle.

The solid line on the left side of Figure 9.8 indicates the locus of the CIO in the GCRS over 50,000 years – about two precession cycles. The ecliptic is shown as a dashed line. The initial nearly straight southward motion from the starting point at J2000.0 is clearly shown. There are occasional "cusps" in the CIO's motion, where its secular motion comes to a temporary halt before reversing. The first of these stationary points occurs in just over a quarter of a precession cycle, as the section of the moving equator that is farthest south in ecliptic coordinates precesses to near the GCRS α=0 hour circle. At that time, the CIO will exhibit only nutational oscillations around a point that remains fixed on the celestial sphere to within 10 mas for almost a decade. Then its motion resumes, this time northward and westward.[28)] The motion of the equinox over the same 50,000-year time period begins at nearly the same point as the CIO (on the plot scale used, the points overlap), but smoothly follows the ecliptic westward (to the right on the plot), wrapping around twice and ending

28) There is nothing profound about the stationary points or the dates on which they occur. If one had started the CIO at GCRS right ascension 6^{h} or 18^{h} at J2000.0, it would have *started* at a stationary point.

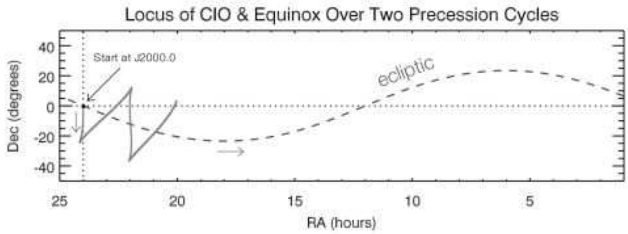

Figure 9.8 Locus of the CIO (solid line) and equinox (dashed line) on the celestial sphere over $5{\times}10^4$ years, with respect to space-fixed coordinates. During this time the equinox wraps around the figure twice and ends up approximately at the starting point.

up essentially at the starting point.

9.7.4
Transforming Vectors Between Reference Systems

The reference points described above allow us to define three geocentric reference systems that share, as a common reference plane, the instantaneous, or true, equator of date. The instantaneous equator is now defined as the plane through the geocenter orthogonal to the direction of the CIP at a given time, t. The three reference systems are:

1) True equator and equinox of t – azimuthal origin at the true equinox (Υ) of t
2) Celestial Intermediate Reference System (CIRS) – azimuthal origin at the Celestial Intermediate Origin (CIO or σ) of t
3) Terrestrial Intermediate Reference System (TIRS) – azimuthal origin at the Terrestrial Intermediate Origin (TIO or ϖ) of t

This chapter will often refer to these reference systems by the symbols E_Υ, E_σ, and E_ϖ, respectively; E denotes an equatorial system, and the subscript indicates the azimuthal origin. E_ϖ rotates with the Earth whereas the orientations of E_Υ and E_σ change slowly with respect to local inertial space. The transformation between E_Υ and E_ϖ is just a rotation about the z-axis (which points toward the CIP) by GAST, the angular equivalent of Greenwich apparent sidereal time. The transformation between E_σ and E_ϖ is a similar rotation, but by θ, the Earth Rotation Angle. These two transformations reflect different ways – old and new – of representing the rotation of the Earth.

A short digression into terrestrial, i.e., geodetic, reference systems is in order here. These systems all have their origin at the geocenter and rotate with the crust of the Earth. Terrestrial latitude, longitude, and height are now most commonly given with respect to a *reference ellipsoid*, an oblate spheroid that approximates the Earth's overall shape (actually, that best fits the *geoid*, a gravitational equipotential surface). The current standard reference elliposid for most purposes is that of the World Geodetic System 1984 (WGS 84), which forms the basis for the coordinates obtained from GPS. The WGS 84 ellipsoid has an equatorial radius of 6,378,137 meters and a polar flattening of 1/298.257223563. For the precise measurement of Earth rotation, however, the International Terrestrial Reference System (ITRS) is used, which was defined by the International Union of Geodesy and Geophysics (IUGG) in 1991. The ITRS is realized for practical purposes by the adopted coordinates and velocities[29)] of a large group of observing stations. These coordinates are expressed as geocentric rectangular 3-vectors and thus are not dependent on a reference ellipsoid. The list of stations and their coordinates is referred to as the International Terrestrial Reference Frame (ITRF). The fundamental terrestrial coordinate system is therefore defined in exactly the same way as the fundamental celestial coordinate system (see section 9.4): a prescription is given for an idealized coordinate system (the ITRS or

29) The velocities are quite small and are due to plate tectonics and post-glacial rebound.

the ICRS), which is realized in practice by the adopted coordinates of a group of reference points (the ITRF stations or the ICRF quasars). The coordinates may be refined as time goes on but the overall system is preserved. It is important to know, however, that the ITRS/ITRF is consistent with WGS 84 to within a few centimeters; thus for all astronomical purposes the GPS-obtained coordinates of instruments can be used with the algorithms presented here.

Our goal is to be able to transform an arbitrary vector (representing for example, an instrumental position, axis, boresight, or baseline) from the ITRS (≈WGS 84≈GPS) to the GCRS. The three equatorial reference systems described above – E_Υ, E_σ, and E_ϖ – are waypoints, or intermediate stops, in that process. The complete transformations are

Equinox-Based Transformation

ITRS or WGS 84

‖

polar motion

⇓

E_ϖ – Terrestrial Intermediate Ref. System

‖

Greenwich apparent sidereal time

⇓

E_Υ – true equator & equinox

‖

equinox-based rotation for
nutation + precession + frame bias

⇓

GCRS

CIO-Based Transformation

ITRS or WGS 84

‖

polar motion

⇓

E_ϖ – Terrestrial Intermediate Ref. System

‖

Earth Rotation Angle

⇓

E_σ – Celestial Intermediate Ref. System

‖

CIO-based rotation for
nutation + precession + frame bias

⇓

GCRS

which are equivalent. That is, given the same input vector, the same output vector will result from the two procedures. In the CIO-based transformation, the three sub-transformations (for polar motion, Earth Rotation Angle, and nutation/precession/frame bias) are independent. That is not true for the equinox-based method, because apparent sidereal time incorporates precession and nutation. Each of the two methods could be made into a single matrix, and the two matrices must be numerically identical. That means that the use of the CIO in the second method does not increase the precision of the result but simply allows for a mathematical redescription of the overall transformation – basically, a re-sorting of the effects to be taken into account. This redescription of the transformation provides a clean separation of the three main aspects of Earth rotation, and recognizes that the observations defining modern reference systems are not sensitive to the equinox. It thus yields a more straightforward conceptual view and facilitates a simpler observational analysis for Earth-rotation measurements and Earth-based astrometry.

These transformations are all rotations that pivot around a common point, the geocenter. Although developed for observations referred to the geocenter, the same set of rotations can be applied to observations made from a specific point on the surface of the Earth. (This follows from the assumption that all points in or on the Earth are rigidly attached to all other points. The actual non-rigidity – e.g., Earth tides – is handled as a separate correction.) In such a case, the computations for parallax, light-bending, aberration, etc., must take into account the non-geocentric position and velocity of the observer. Then the final computed coordinates are referred, not to the GCRS, but rather to a proper reference system (in the terminology of relativity) of the observer.

Let us return to the more familiar problem mentioned at the beginning of the section: the computation of local hour angle. In the usual equinox-based scheme, the apparent place[30)] of the star or planet is expressed with respect to the true equator and equinox of date (E_Υ). The local hour angle is just h = GAST - α_Υ + λ, where GAST is Greenwich apparent sidereal time, α_Υ is the apparent right ascension of the object, measured with respect to the true equinox, and λ is the longitude of the observer (corrected, where necessary, for polar motion). Obviously these quantities must all be given in the same units. In the CIO-based scheme, the apparent place would be expressed in the Celestial Intermediate Reference System (E_σ), and h = θ - α_σ + λ, where θ is the Earth Rotation Angle and α_σ is the apparent right ascension of the object, measured with respect to the CIO. (The recommended terminology is *intermediate place* for the object position in the E_σ and *intermediate right ascension* for the quantity α_σ, although some people hold that the term right ascension should refer only to an origin at the equinox.) In the CIO-based formula, precession and nutation come into play only once, in expressing the object's right ascension in the E_σ system. See section 9.7.5.7 for more details.

9.7.5
Formulas

The formulas below draw heavily on the developments presented previously. In particular, the 3×3 matrices **P**, **N**, and **B** represent the transformations for precession, nutation, and frame bias, respectively, and are taken directly from sections 9.6 and

30) See definition on page 750.

9.4. The matrices **P** and **N** are functions of time, t. The time measured in Julian centuries of TDB (or TT) from J2000.0 is denoted T and is given by T = (JD(TDB) - 2451545.0)/36525. The elementary rotation matrices $\mathbf{R}_1$, $\mathbf{R}_2$, and $\mathbf{R}_3$ are defined in "Abbreviations and Symbols Frequently Used". Formulas from section 9.3 for sidereal time and the Earth Rotation Angle are used. Explanations of, and formulas for the time scales UT1, TT, and TDB are also found in section 9.3.

The ultimate objective is to express "local" Earth-fixed vectors, representing geographic positions, baselines, instrumental axes and boresights, etc., in the GCRS, where they can be related to the proper coordinates[31)] of celestial objects. As mentioned above, the "GCRS" is a reference system that can be thought of as the "geocentric ICRS". Celestial coordinates in the GCRS are obtained from basic ICRS data (which are barycentric) by applying the usual algorithms for proper place; see section 9.2.3. This section will be working out entirely in a geocentric system and the GCRS will be obtained from a series of rotations that start with an ordinary Earth-fixed geodetic system.

9.7.5.1 *Location of Cardinal Points*

Let us start by establishing the positions of three cardinal points within the GCRS: the Celestial Intermediate Pole (CIP), the true equinox of date (♈), and the Celestial Intermediate Origin (CIO). The unit vectors toward these points will be designated $\mathbf{n}_{\text{GCRS}}$, $\boldsymbol{\Upsilon}_{\text{GCRS}}$, and $\boldsymbol{\sigma}_{\text{GCRS}}$, respectively. As the Earth precesses and nutates in local inertial space, these points are in continual motion.

The CIP and the equinox can easily be located in the GCRS at any time t simply by recognizing that they are, respectively, the z- and x-axes of the true equator and equinox of date system (E_{Υ}) at t. The unit vectors therefore are

$$\begin{aligned} \text{CIP} \quad : \qquad \mathbf{n}_{\text{GCRS}}(t) &= \mathbf{B}^{\mathrm{T}}\,\mathbf{P}^{\mathrm{T}}(t)\,\mathbf{N}^{\mathrm{T}}(t)\begin{pmatrix}0\\0\\1\end{pmatrix} \\ \text{Equinox:} \qquad \boldsymbol{\Upsilon}_{\text{GCRS}}(t) &= \mathbf{B}^{\mathrm{T}}\,\mathbf{P}^{\mathrm{T}}(t)\,\mathbf{N}^{\mathrm{T}}(t)\begin{pmatrix}1\\0\\0\end{pmatrix} \end{aligned} \tag{9.60}$$

where the matrix **B** accounts for the GCRS frame bias (same as for the ICRS) and the matrices $\mathbf{P}(t)$ and $\mathbf{N}(t)$ provide the transformations for precession and nutation, respectively, at time t. These matrices were developed in sections 9.4.9 and 9.6.4; the superscript T's above indicate that the transpose of each of these matrices as previously developed is used (i.e., one is using the "reverse" transformations here, from the true equator and equinox of t to the GCRS). The first equation above is simply equation (9.29) rewritten. Note that $\boldsymbol{\Upsilon}_{\text{GCRS}}$ is orthogonal to $\mathbf{n}_{\text{GCRS}}$ at each time t.

The components of the unit vector in the direction of the pole, $\mathbf{n}_{\text{GCRS}}$, are denoted X, Y, and Z, and another approach to determining $\mathbf{n}_{\text{GCRS}}$ is to use the series expansions

31) See footnote 1 on page 750.

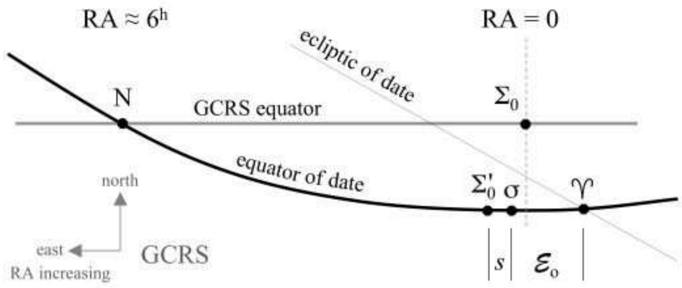

Figure 9.9 Relationship between various points involved in locating the CIO. The figure approximates the relative positions at 2020.0, although the spacings are not to scale. The point labeled σ is the CIO; Υ is the true equinox; Σ_0 is the GCRS right ascension origin; N is the ascending node of the instantaneous (true) equator of date on the GCRS equator; and Σ_0' is the point on the instantaneous equator that is the same distance from N as Σ_0. As shown, the quantities s and $\mathcal{E}_o$ are respectively positive and negative. The motion of the instantaneous equator (which is orthogonal to the CIP) is generally southward (down in the figure) near RA=0, which tends to move the equinox westward (right) and the CIO very slightly eastward (left) with respect to the GCRS.

for X and Y given in the McCarthy and Petit [2004]. There is a table of daily values of X and Y in section B of *The Astronomical Almanac* [2010] (there labeled $\mathcal{X}$ and $\mathcal{Y}$). Once X and Y are converted to dimensionless values, $Z = \sqrt{1 - X^2 - Y^2}$. (The IERS series for X and Y are part of a data analysis approach adopted by the IERS that avoids any explicit reference to the ecliptic or the equinox, although the underlying theories are those described in section 9.6.)

There are three possible procedures for obtaining the location of the Celestial Intermediate Origin on the celestial sphere at a given time: (1) following the arc on the instantaneous equator from the equinox to the CIO; (2) directly computing the position vector of the CIO in the GCRS by numerical integration; or (3) using the quantity s, representing the difference in two arcs on the celestial sphere, one of which ends at the CIO. These procedures will be described in the three subsections below. Figure 9.9 indicates the geometric relationships among the points mentioned.

9.7.5.2 *CIO Location Relative to the Equinox*

The arc on the instantaneous (true) equator of date t from the CIO to the equinox is called the *equation of the origins* and is the right ascension of the true equinox relative to the CIO (or, minus the true right ascension of the CIO). The equation of the origins is also the difference θ − GAST. It therefore equals the accumulated precession-nutation of the equinox in right ascension, given by the sum of the terms in parentheses from equation (9.13) and the equation of the equinoxes given in equation (9.16) (all times -1). The equation of the origins in arcseconds therefore is

$$\begin{aligned}\mathcal{E}_o = \; & - \; 0.014506 - 4612.156534\,T - 1.3915817\,T^2 + 0.00000044\,T^3 \\ & + \; 0.000029956\,T^4 + 0.0000000368\,T^5 \quad - \Delta\psi\,\cos\epsilon \\ & - \; 0.00264096\;\sin(\Omega) \\ & - \; 0.00006352\;\sin(2\Omega) \\ & - \; 0.00001175\;\sin(2F - 2D + 3\Omega) \\ & - \; 0.00001121\;\sin(2F - 2D + \Omega) \\ & + \; 0.00000455\;\sin(2F - 2D + 2\Omega) \\ & - \; 0.00000202\;\sin(2F + 3\Omega) \\ & - \; 0.00000198\;\sin(2F + \Omega) \\ & + \; 0.00000172\;\sin(3\Omega) \\ & + \; 0.00000087\,T\;\sin(\Omega) + \cdots \end{aligned} \tag{9.61}$$

where T is the number of centuries of TDB (or TT) from J2000.0; $\Delta\psi$ is the nutation in longitude, in arcseconds; ϵ is the mean obliquity of the ecliptic; and F, D, and Ω are fundamental luni-solar arguments. All of the angles are functions of time; see section 9.6 for expressions (especially equations (9.42), (9.46), & (9.50)). There is a table of daily values of $\mathcal{E}_o$ in section B of *The Astronomical Almanac* [2010].

To transform an object's celestial coordinates from the true equator and equinox of t to the Celestial Intermediate System (i.e., from E_Υ to E_σ), simply add $\mathcal{E}_o$ to the object's true right ascension. To similarly transform the components of a position vector, apply the rotation $\mathbf{R}_3(-\mathcal{E}_o)$. Since many existing software systems are set up to produce positions with respect to the equator and equinox of date, this is a relatively easy way to convert those positions to the Celestial Intermediate Reference System if desired. Note that in such a case there is no computational difference in using either the equinox-based or CIO-based methods for computing hour angle: $\mathcal{E}_o$ is computed in both methods and is just applied to different quantities. In the equinox-based method, $\mathcal{E}_o$ is subtracted from θ to form sidereal time; in the CIO-based method, $\mathcal{E}_o$ is added to the object's true right ascension so that θ can be used in place of sidereal time.

The position of the CIO in the GCRS, $\boldsymbol{\sigma}_{\text{GCRS}}$, can be established by taking the position vector of the equinox in the GCRS, $\boldsymbol{\Upsilon}_{\text{GCRS}}$, and rotating it counterclockwise by the angle $-\mathcal{E}_o$ (i.e., clockwise by $\mathcal{E}_o$) about the axis $\mathbf{n}_{\text{GCRS}}$. Equivalently, establish the orthonormal basis triad of the equator-and-equinox system within the GCRS:

$\boldsymbol{\Upsilon}_{\text{GCRS}}$, $(\mathbf{n}_{\text{GCRS}} \times \boldsymbol{\Upsilon}_{\text{GCRS}})$, and $\mathbf{n}_{\text{GCRS}}$. Then

$$\boldsymbol{\sigma}_{\text{GCRS}} = \boldsymbol{\Upsilon}_{\text{GCRS}} \cos\mathcal{E}_o - (\mathbf{n}_{\text{GCRS}} \times \boldsymbol{\Upsilon}_{\text{GCRS}}) \sin\mathcal{E}_o \tag{9.62}$$

9.7.5.3 *CIO Location from Numerical Integration*

As described above, a non-rotating origin can be described as a point on the moving equator whose instantaneous motion is always orthogonal to the equator. A simple geometric construction based on this definition yields the following differential equation for the motion of a non-rotating origin

$$\dot{\boldsymbol{\sigma}}(t) = -\Big(\boldsymbol{\sigma}(t) \cdot \dot{\mathbf{n}}(t)\Big)\,\mathbf{n}(t) \tag{9.63}$$

That is, if one has a model for the motion of the pole, $\mathbf{n}(t)$, the path of the non-rotating origin is described by $\boldsymbol{\sigma}(t)$, once an initial point on the equator, $\boldsymbol{\sigma}(t_0)$, is chosen. Conceptually and practically, it is simple to integrate this equation, using, for example, a standard 4th-order Runge-Kutta integrator. For the motions of the real Earth, fixed step sizes of order 0.5 day work quite well, and the integration is quite robust. This is actually a one-dimensional problem carried out in three dimensions, since one knows the non-rotating origin remains on the equator; one really needs only to know where along the equator it is. Therefore, two constraints can be applied at each step: $|\boldsymbol{\sigma}| = 1$ and $\boldsymbol{\sigma} \cdot \mathbf{n} = 0$. See McCarthy [2005].

The above equation is quite general, and to get the specific motion of the CIO, each of the vectors in the above equation is expressed with respect to the GCRS, i.e., $\boldsymbol{\sigma}(t) \to \boldsymbol{\sigma}_{\text{GCRS}}(t)$, $\mathbf{n}(t) \to \mathbf{n}_{\text{GCRS}}(t)$, and $\dot{\mathbf{n}}(t) \to \dot{\mathbf{n}}_{\text{GCRS}}(t)$. The pole's position, $\mathbf{n}_{\text{GCRS}}(t)$, is given by the first expression in equation (9.60). The pole's motion, $\dot{\mathbf{n}}_{\text{GCRS}}(t)$, can be obtained by numerical differentiation of the pole's position. By numerically integrating the above equation, one obtains a time series of unit vectors, $\boldsymbol{\sigma}_{\text{GCRS}}(t_i)$, where each i is an integration step. The fact that this is actually just a one-dimensional problem means that it is sufficient to store as output the CIO right ascensions (with respect to the GCRS), using equation (9.25) to decompose the $\boldsymbol{\sigma}_{\text{GCRS}}(t_i)$ vectors. In this way, the integration results in a tabulation of CIO right ascensions at discrete times. For example, see Kaplan [2005a], where the times are expressed as TDB Julian dates and the right ascensions are in arcseconds. This file runs from years 1700 to 2300 at 1.2-day intervals.

When one needs to obtain the position of the CIO for some specific time, the file of CIO right ascensions can be interpolated to that time. The CIO's unit vector, if required, can be readily computed: generally, given a fixed coordinate system within which a pole and its equator move, a point of interest on the equator has a position vector in the fixed system given by

$$\mathbf{r} = \begin{pmatrix} Z\cos\alpha \\ Z\sin\alpha \\ -X\cos\alpha - Y\sin\alpha \end{pmatrix} \tag{9.64}$$

where α is the right ascension of the point, and X, Y, and Z are the components of the pole's instantaneous unit vector, and all of these quantities are measured relative to the fixed coordinate system. The vector $\mathbf{r}$ is not in general of unit length but it can be readily normalized. This formula allows us to reconstruct the unit vector toward the CIO from just its GCRS right ascension value at the time of interest, since one already knows how to obtain the pole's position vector in the GCRS for that time.

Equation (9.63) can be also made to yield the locus of the Terrestrial Intermediate Origin (TIO), simply by referring all the vectors to the ITRS – a rotating geodetic system – rather than the GCRS. In this case, therefore, $\boldsymbol{\sigma}(t) \rightarrow \boldsymbol{\varpi}_{\text{ITRS}}(t)$, $\mathbf{n}(t) \rightarrow \mathbf{n}_{\text{ITRS}}(t)$, and $\dot{\mathbf{n}}(t) \rightarrow \dot{\mathbf{n}}_{\text{ITRS}}(t)$. The path of the CIP within the ITRS ($\mathbf{n}_{\text{ITRS}}(t)$) is what is called polar motion (usually specified by the parameters x_p and y_p), and is fundamentally unpredictable. The integration can therefore only be accurately done for past times, using observed pole positions. A computed future path of the TIO on the surface of the Earth depends on the assumption that the two major periodicities observed in polar motion will continue at approximately the current amplitude.

9.7.5.4 *CIO Location from the Arc-Difference s*

On the celestial sphere, the Earth's instantaneous (moving) equator intersects the GCRS equator at two nodes. Let N be the ascending node of the instantaneous equator on the GCRS equator. One can define a scalar quantity $s(t)$ that represents the *difference* between the length of the arc from N westward to the CIO (on the instantaneous equator) and the length of the arc from N westward to the GCRS origin of right ascension (on the GCRS equator). The quantity s is called the *CIO locator*. If σ represents the CIO and Σ_0 represents the right ascension origin of the GCRS (the direction of the GCRS x-axis), then

$$s = \overline{\sigma \mathrm{N}} - \overline{\Sigma_0 \mathrm{N}} \tag{9.65}$$

See figure 9.9, where the points Σ_0 and Σ_0' are equidistant from the node N. The quantity s is seen to be the "extra" length of the arc on the instantaneous equator from N to σ, the position of the CIO. The value of s is fundamentally obtained from an integral,

$$s(t) = -\int_{t_0}^{t} \frac{X(t)\dot{Y}(t) - Y(t)\dot{X}(t)}{1 + Z(t)}\, dt \;+ s_0 \tag{9.66}$$

where $X(t)$, $Y(t)$, and $Z(t)$ are the three components of the unit vector, $\mathbf{n}_{\text{GCRS}}(t)$, toward the celestial pole (CIP). See, e.g., Capitaine *et al.* [2000] or the McCarthy and Petit [2004]. The constant of integration, s_0, has been set to ensure that the equinox-based and CIO-based computations of Earth rotation yield the same answers: $s_0 = 94\ \mu$as [McCarthy and Petit, 2004]. Effectively, the constant adjusts the position of the CIO on the equator and is thus part of the arc $\overline{\sigma \mathrm{N}}$. For practical purposes, the value of s at any given time is provided by a series expansion, given in Table 5.2c of the McCarthy and Petit [2004]. Software to evaluate this series is available at the IERS Conventions web site and is also part of the SOFA package. There is a table of daily values of s in section B of *The Astronomical Almanac* [2010].

At any time, t, the the unit vector toward the node N is simply $\mathbf{N}_{\text{GCRS}} = (-Y, X, 0)/\sqrt{X^2 + Y^2}$ (where one is no longer explicitly indicating the time dependence of X and Y). To locate the CIO, one rearranges equation (9.65) to yield the arc length $\overline{\sigma\mathrm{N}}$

$$\overline{\sigma\mathrm{N}} = s + \overline{\Sigma_0\mathrm{N}} = s + \arctan(X/(-Y)) \tag{9.67}$$

The location of the CIO is then obtained by starting at the node N and moving along the instantaneous equator of t through the arc $\overline{\sigma\mathrm{N}}$. That is, $\boldsymbol{\sigma}_{\text{GCRS}}$ can be constructed by taking the position vector of the node N in the GCRS, $\mathbf{N}_{\text{GCRS}}$, and rotating it counterclockwise by the angle $-\overline{\sigma\mathrm{N}}$ (i.e., clockwise by $\overline{\sigma\mathrm{N}}$) about the axis $\mathbf{n}_{\text{GCRS}}$. Equivalently,

$$\boldsymbol{\sigma}_{\text{GCRS}} = \mathbf{N}_{\text{GCRS}} \cos(\overline{\sigma\mathrm{N}}) - (\mathbf{n}_{\text{GCRS}} \times \mathbf{N}_{\text{GCRS}}) \sin(\overline{\sigma\mathrm{N}}) \tag{9.68}$$

The three methods for determining the position of the CIO in the GCRS are numerically the same to within several microarcseconds (μas) over six centuries centered on J2000.0. One now has formulas in hand for obtaining the positions of the three cardinal points on the sky – the CIP, the CIO, and the equinox – that are involved in the ITRS-to-GCRS (terrestrial-to-celestial) transformations. In the following, it is assumed that $\mathbf{n}_{\text{GCRS}}$, $\boldsymbol{\sigma}_{\text{GCRS}}$, and $\boldsymbol{\Upsilon}_{\text{GCRS}}$ are known vectors for some time t of interest.

9.7.5.5 *Geodetic Position Vectors and Polar Motion*

Vectors representing the geocentric positions of points on or near the surface of the Earth are of the general form

$$\mathbf{r} = \begin{pmatrix} (aC + h)\cos\phi_G \cos\lambda_G \\ (aC + h)\cos\phi_G \sin\lambda_G \\ (aS + h)\sin\phi_G \end{pmatrix} \tag{9.69}$$

where λ_G is the geodetic longitude, ϕ_G is the geodetic latitude, and h is the height. These coordinates are measured with respect to a reference ellipsoid, fit to the equipotential surface that effectively defines mean sea level. The ellipsoid has an equatorial radius of a and a flattening factor f. The quantities C and S depend on the flattening

$$C = 1/\sqrt{\cos^2\phi_G + (1-f)^2 \sin^2\phi_G} \qquad S = (1-f)^2\, C \tag{9.70}$$

A complete description of geodetic concepts, reference ellipsoids, and computations is beyond the scope of this Chapter, but a brief summary can be found in section K of *The Astronomical Almanac* [2010] and a more thorough account is given in Chapter 4 of the Seidelmann and Urban [2010]. More information can be found in any introductory textbook on geodesy. Suffice it here to say that the reference ellipsoid for GPS is WGS 84, with $a = 6378137$ m and $f = 1/298.257223563$. For astronomical purposes it can be assumed that WGS 84 is a good approximation to the International Terrestrial Reference System (ITRS) described previously in this section. That is, GPS provides a realization of the ITRS.

It is worth noting that modern space techniques often measure geocentric positions in rectangular coordinates directly, without using a reference ellipsoid. Also, not all vectors of interest represent geographic locations. Vectors representing instrumental axes, baselines, and boresights are often of more interest to astronomers and these can usually be easily expressed in the same geodetic system as the instrument location. All Earth-fixed vectors, regardless of what they represent, are subject to the same transformations described below. The ITRS is the assumed starting point for these transformations, even though in most cases astronomers will be using vectors in some system that approximates the ITRS.

For astronomical applications, one must correct ITRS vectors for polar motion (also called wobble). In current terminology, this is a transformation from the ITRS to the Terrestrial Intermediate Reference System (TIRS; in this Chapter it is designated E_{ϖ}), and is the first transformation shown in the flowcharts on page 771. Polar motion is the small quasi-periodic excursion of the geodetic pole from the pole of rotation, or, more precisely stated, the excursion of the ITRS z-axis from the CIP. It is described by the parameters x_p and y_p, which are the coordinates of the CIP in the ITRS, and which generally amount to a few tenths of an arcsecond. Daily values of x and y, the observed pole coordinates, are published by the IERS (see, e.g., Central Bureau of the IERS [2010]). These published values should, for the most precise applications (<1 mas), be augmented by very small predictable components to polar motion, with periods ≤1.2 days. These extra components are evaluated and added after interpolation of the published x and y values – see section 5.4.2 of the McCarthy and Petit [2004]. The sum (total polar motion) is x_p and y_p. However, most users will be able to assume $x_p = x$ and $y_p = y$.

The transformation one seesk not only must reorient the pole from the ITRS z-axis to the CIP, it also must move the origin of longitude very slightly from the ITRS x-axis to the Terrestrial Intermediate Origin (TIO). The latter shift is so tiny that its magnitude can be given by an approximate formula, linear in time, based on the two main circular components of polar motion as observed over the last few decades. That shift is $s' = -47\ \mu\mathrm{as}\ T$, where T is the time (either TT or TDB) in centuries from J2000.0 [Lambert and Bizouard, 2002]. Since 47 μas amounts to 1.5 mm on the surface of the Earth, this correction is entirely negligible for most purposes. Nevertheless, it is included here for completeness.

The ITRS to E_{ϖ} transformation then is accomplished using the polar motion (wobble) matrix $\mathbf{W}(t)$

$$\mathbf{r}_{\mathrm{E}_{\varpi}} = \mathbf{W}(t)\ \mathbf{r}_{\mathrm{ITRS}} \tag{9.71}$$

where

$$\mathbf{W}(t) = \mathbf{R}_3(-s')\ \mathbf{R}_2(x_p)\ \mathbf{R}_1(y_p) \tag{9.72}$$

If one let

$$\begin{aligned} S_x &= \sin(x_p) \\ S_y &= \sin(y_p) \\ S_s &= \sin(-s') \end{aligned}$$

$$
\begin{aligned}
C_x &= \cos(x_p) \\
C_y &= \cos(y_p) \\
C_s &= \cos(-s')
\end{aligned} \tag{9.73}
$$

then the wobble matrix can also be written

$$
\mathbf{W}(t) = \begin{pmatrix} C_x C_s & S_x S_y C_s + C_y S_s & -S_x C_y C_s + S_y S_s \\ -C_x S_s & -S_x S_y S_s + C_y C_s & S_x C_y S_s + S_y C_s \\ S_x & -C_x S_y & C_x C_y \end{pmatrix} \approx \begin{pmatrix} 1 & -s' & -x_p \\ s' & 1 & y_p \\ x_p & -y_p & 1 \end{pmatrix} \tag{9.74}
$$

where the form on the right is a first-order approximation. Due to the smallness of the angles involved, the first-order matrix is quite adequate for most applications – further, s' can be set to zero.

9.7.5.6 *Complete Terrestrial to Celestial Transformation*

The transformations corresponding to the two flowcharts on page 771 are

$$
\begin{aligned}
&\text{Equinox-based transformation:} \quad & \mathbf{r}_{\text{GCRS}} &= \mathbf{B}^{\mathrm{T}}\,\mathbf{P}^{\mathrm{T}}\,\mathbf{N}^{\mathrm{T}}\,\mathbf{R}_3(-\text{GAST})\,\mathbf{W}\;\mathbf{r}_{\text{ITRS}} \\
&\text{CIO-based transformation:} \quad & \mathbf{r}_{\text{GCRS}} &= \mathbf{C}^{\mathrm{T}}\,\mathbf{R}_3(-\theta)\,\mathbf{W}\;\mathbf{r}_{\text{ITRS}}
\end{aligned} \tag{9.75}
$$

where all of the matrices except **B** are time-dependent. GAST is Greenwich apparent sidereal time and θ is the Earth Rotation Angle; formulas are given in section 9.3.6.2 (equations (9.11)-(9.16)). The matrices, working from right to left, perform the following sub-transformations

$\mathbf{W}$	ITRS to E_{ϖ}
$\mathbf{R}_3(-\text{GAST})$	E_{ϖ} to E_{Υ}
$\mathbf{B}^{\mathrm{T}}\,\mathbf{P}^{\mathrm{T}}\,\mathbf{N}^{\mathrm{T}}$	E_{Υ} to GCRS
$\mathbf{R}_3(-\theta)$	E_{ϖ} to E_{σ}
$\mathbf{C}^{\mathrm{T}}$	E_{σ} to GCRS

The three "E" reference systems were described on page 770.

The matrix **C** has not yet been developed in this section. One form of the matrix **C** was introduced in section 9.6.7 for the transformation from the GCRS to E_{σ}. The transpose is used here because one is interested in the opposite transformation. **C** or $\mathbf{C}^{\mathrm{T}}$ is easy to construct because one already has the three basis vectors of E_{σ} expressed in the GCRS: the z-axis is toward $\mathbf{n}_{\text{GCRS}}$, the CIP; the x-axis is toward $\boldsymbol{\sigma}_{\text{GCRS}}$, the CIO; and the y-axis is toward $\mathbf{n}_{\text{GCRS}} \times \boldsymbol{\sigma}_{\text{GCRS}}$. Call the latter vector $\mathbf{y}_{\text{GCRS}}$. Then,

$$
\mathbf{C}^{\mathrm{T}} = \begin{pmatrix} \boldsymbol{\sigma}_{\text{GCRS}} & \mathbf{y}_{\text{GCRS}} & \mathbf{n}_{\text{GCRS}} \end{pmatrix} = \begin{pmatrix} \sigma_1 & y_1 & n_1 \\ \sigma_2 & y_2 & n_2 \\ \sigma_3 & y_3 & n_3 \end{pmatrix} = \begin{pmatrix} \sigma_1 & y_1 & X \\ \sigma_2 & y_2 & Y \\ \sigma_3 & y_3 & Z \end{pmatrix} \tag{9.76}
$$

where X, Y, and Z are the CIP coordinates, expressed as dimensionless quantities. As in section 9.6.7, the matrix can also be constructed using only X and Y, together with the CIO locator, s

$$\mathbf{C}^{\mathrm{T}} = \begin{pmatrix} 1 - bX^2 & -bXY & X \\ -bXY & 1 - bY^2 & Y \\ -X & -Y & 1 - b(X^2 + Y^2) \end{pmatrix} \mathbf{R}_3(s) \tag{9.77}$$

where $b = 1/(1 + Z)$ and $Z = \sqrt{1 - X^2 - Y^2}$. The latter form is taken from the McCarthy and Petit [2004], Chapter 5, where $\mathbf{C}^{\mathrm{T}}$ is called $Q(t)$. The two constructions of $\mathbf{C}^{\mathrm{T}}$ are numerically the same.

9.7.5.7 *Hour Angle*

The local hour angle of a celestial object is given by

$$\begin{aligned} \text{Equinox-based formula:} \quad & \mathrm{h} = \mathrm{GAST} - \alpha_{\Upsilon} + \lambda \\ \text{CIO-based formula:} \quad & \mathrm{h} = \theta - \alpha_{\sigma} + \lambda \end{aligned} \tag{9.78}$$

where GAST is Greenwich apparent sidereal time, θ is the Earth Rotation Angle, and λ is the longitude of the observer. The quantities involved can be expressed in either angle or time units as long as they are consistent. The right ascension in the two cases is expressed with respect to different origins: α_{Υ} is the apparent right ascension of the object, measured with respect to the true equinox, and α_{σ} is the apparent right ascension of the object, measured with respect to the CIO. That is, the coordinates of the object are expressed in system E_{Υ} in the equinox-based formula and in system E_{σ} in the CIO-based formula. Since both systems share the same equator – the instantaneous equator of date, orthogonal to the CIP – the apparent declination of the object is the same in the two cases.

The two formulas in (9.78) are equivalent, which can be seen by substituting, in the equinox-based formula, $\mathrm{GAST} = \theta - \mathcal{E}_o$ and $\alpha_{\Upsilon} = \alpha_{\sigma} - \mathcal{E}_o$, where $\mathcal{E}_o$ is the equation of the origins.

The longitude of the observer, λ, is expressed in the E_{ϖ} system, that is, it is corrected for polar motion. Using the first-order form of the matrix $\mathbf{W}$, given in equation (9.74) (and assuming s'=0), it is straightforward to derive equation (9.19) for λ. Using notation consistent with that used in this section, this equation is

$$\lambda \equiv \lambda_{\mathrm{E}_{\varpi}} = \lambda_{\mathrm{ITRS}} + \left(x_p \sin\lambda_{\mathrm{ITRS}} + y_p \cos\lambda_{\mathrm{ITRS}}\right) \tan\phi_{\mathrm{ITRS}} \,/3600 \tag{9.79}$$

where λ_{ITRS} and ϕ_{ITRS} are the ITRS (geodetic) longitude and latitude of the observer, with λ_{ITRS} in degrees; and x_p and y_p are the coordinates of the pole (CIP), in arcseconds. This formula is approximate and should not be used for places at polar latitudes.

The corresponding equation for the latitude, ϕ, corrected for polar motion is

$$\phi \equiv \phi_{\mathrm{E}_{\varpi}} = \phi_{\mathrm{ITRS}} + \left(x_p \cos\lambda_{\mathrm{ITRS}} - y_p \sin\lambda_{\mathrm{ITRS}}\right)/3600 \tag{9.80}$$

although this equation is not needed in hour angle computations; it is given here only for completeness.

The common notion of hour angle can be expressed more precisely using concepts introduced in sections 9.6 and 9.7. The local hour angle of an object is the angle between two planes: the plane containing the geocenter, the CIP, and the observer; and the plane containing the geocenter, the CIP, and the object. Hour angle increases with time and is positive when the object is west of the observer as viewed from the geocenter. The two planes define meridians on the celestial sphere that meet at the CIP. From the point of view of the observer, the CIP is not, in general, exactly at the geodetic north point, which is the direction toward the ITRS z-axis. The azimuths of the two directions differ by as much as $\sqrt{x_p^2 + y_p^2}/\cos\phi_{\text{ITRS}}$, depending on time of day. This difference is small (usually <1 arcsecond) and often negligible for practical applications. The plane defining the astronomical Greenwich meridian (from which Greenwich hour angles are measured) can be understood to contain the geocenter, the CIP, and TIO; there, $\lambda \equiv \lambda_{\text{E}_\varpi} = 0$. This plane is now called the TIO meridian.

The CIO-based formula for hour angle is quite simple to use (since θ is linear with time) if one has the coordinates of the object expressed in system E_σ. Fortunately, this is straightforward if one has the object's coordinates expressed in the GCRS, because one also has the basis vectors of E_σ expressed in the GCRS. If the object's vector in the GCRS is $\mathbf{r}_{\text{GCRS}}$, then the object's vector in E_σ is simply

$$\mathbf{r}_{\text{E}_\sigma} = \mathbf{C}\,\mathbf{r}_{\text{GCRS}} = \begin{pmatrix} \mathbf{r}_{\text{GCRS}} \cdot \boldsymbol{\sigma}_{\text{GCRS}} \\ \mathbf{r}_{\text{GCRS}} \cdot \mathbf{y}_{\text{GCRS}} \\ \mathbf{r}_{\text{GCRS}} \cdot \mathbf{n}_{\text{GCRS}} \end{pmatrix} \qquad \text{where} \quad \mathbf{y}_{\text{GCRS}} = \mathbf{n}_{\text{GCRS}} \times \boldsymbol{\sigma}_{\text{GCRS}} \tag{9.81}$$

Then

$$\alpha_\sigma = \arctan\left(\frac{\mathbf{r}_{\text{GCRS}} \cdot \mathbf{y}_{\text{GCRS}}}{\mathbf{r}_{\text{GCRS}} \cdot \boldsymbol{\sigma}_{\text{GCRS}}}\right) \tag{9.82}$$

As a specific case, one knows the position vector of the equinox, $\boldsymbol{\Upsilon}_{\text{GCRS}}$. Applying equation (9.82) to $\boldsymbol{\Upsilon}_{\text{GCRS}}$ and using it in the second formula of (9.78), with λ=0, one obtains the hour angle of the equinox at the Greenwich (or TIO) meridian. But this is the definition of Greenwich Apparent Sidereal Time. Therefore,

$$\text{GAST} = \theta - \arctan\left(\frac{\boldsymbol{\Upsilon}_{\text{GCRS}} \cdot \mathbf{y}_{\text{GCRS}}}{\boldsymbol{\Upsilon}_{\text{GCRS}} \cdot \boldsymbol{\sigma}_{\text{GCRS}}}\right) \tag{9.83}$$

Evidently, then,

$$\mathcal{E}_o = \arctan\left(\frac{\boldsymbol{\Upsilon}_{\text{GCRS}} \cdot \mathbf{y}_{\text{GCRS}}}{\boldsymbol{\Upsilon}_{\text{GCRS}} \cdot \boldsymbol{\sigma}_{\text{GCRS}}}\right) \tag{9.84}$$

which merely restates the definition of $\mathcal{E}_o$ – the equatorial angle from the CIO to the equinox, i.e., the right ascension of the equinox in system E_σ.

A
Fundamental Solution of the Laplace Equation

This appendix gives solutions of the Laplace equations for scalar $F(t, \boldsymbol{x})$, vector $F_i(t, \boldsymbol{x})$, and tensor $F_{ij}(t, \boldsymbol{x})$ fields in three dimensional space with coordinates $\boldsymbol{r} = (x, y, z)$. These equations are

$$\triangle F(t, \boldsymbol{x}) = 0 \,, \tag{A.1a}$$

$$\triangle F_i(t, \boldsymbol{x}) = 0 \,, \tag{A.1b}$$

$$\triangle F_{ij}(t, \boldsymbol{x}) = 0 \,, \tag{A.1c}$$

where $\triangle = \delta^{ij}\partial_i\partial_j$ is the Laplace operator, and $\partial_i = \partial/\partial x^j$ is a partial derivative along i-th coordinate axis. The procedure of finding solutions of equations (A.1) is based on the approach developed by Gelfand *et al.* [1958] (see also [Blanchet and Damour, 1986; Thorne, 1980] and references therein).

Basic spherical harmonics are [Gelfand *et al.*, 1958, page 42, equation 14]

$$Y^{lm}(\theta, \phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi} P_l^m(cos\theta) \,, \qquad (-l \le m \le l), \tag{A.2}$$

where $P_l^m(\cos\theta)$ are the normalized associated Legendre functions

$$P_l^m(z) = \sqrt{\frac{(l+m)!}{(l-m)!}} \sqrt{\frac{2l+1}{2}} \frac{(1-z^2)^{-m/2}}{2^l l!} \frac{d^{l-m}(z^2-1)^l}{dz^{l-m}} \,. \tag{A.3}$$

Spherical harmonics $Y^{lm}(\theta, \phi)$ form an orthogonal system of elements of a canonical basis on the surface of a unit sphere.

The general theorem concerning the resolution of a unitary representation of the group of rotation into irreducible representations means that every function, $F(\theta, \phi)$, such that its squared modulus is integrable over the surface of the sphere, may be expanded in a convergent series with respect to the spherical harmonics

$$F(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} F_{lm} Y^{lm}(\theta, \phi) \,, \tag{A.4}$$

where F_{lm} are complex-numbered coefficients. Making use of transformation from spherical to Cartesian coordinates, one can obtain a one-to-one mapping between the spherical harmonics and the symmetric trace-free (STF) tensors with rank l [Thorne, 1980]:

$$Y^{lm}(\theta, \phi) = \mathcal{Y}^{lm}_{<K_l>} N^{K_l} \,, \tag{A.5}$$

where here and nowhere else, the repeated indices do not assume the Einstein summation rule, the object

$$N^{K_l} = n^{k_1} n^{k_2} ... n^{k_l} \,, \tag{A.6}$$

is a tensor of l-th rank made of the product of l components of the unit radial vector, $\boldsymbol{n} = (n^i) = (n^x, n^y, n^z)$. Tensors $\mathcal{Y}^{lm}_{<K_l>} \equiv \mathcal{Y}^{lm}_{<k_1 k_2 ... k_l>}$ with index m taking values $-l \le m \le l$, form a basis in $(2l+1)$-dimensional space of symmetric and trace-free tensors with l indices. Their exact form is rather long and is not important in the current discussion. It can be found in [Thorne, 1980, equation 2.12].

Any STF tensor of rank l can be represented in form of a finite decomposition

$$\mathcal{F}_{<K_l>} = \sum_{m=-l}^{l} F^{lm} \mathcal{Y}^{lm}_{<K_l>} \,, \tag{A.7}$$

where the multi-index $K_l \equiv k_1 k_2 ... k_l$. Hence, equation (A.4) can be recast to the following form

$$F(\theta, \phi) = \sum_{l=0}^{\infty} \mathcal{F}_{<K_l>} N^{K_l} \,, \tag{A.8}$$

or in equivalent form

$$F(\theta, \phi) = \sum_{l=0}^{\infty} \mathcal{F}_{<L>} N^{L} \,, \tag{A.9}$$

that replaces the multi-index $K_l \to L$ which can be understood as a set of any l indices which enter the summation as the dummy indices.

Spherical harmonics are the eigenfunctions of the orbital angular momentum operator

$$\mathbf{L}^2 Y^{lm} \equiv \left[\partial_r (r^2 \partial_r) - r^2 \triangle\right] Y^{lm} = l(l+1) Y^{lm} \,, \tag{A.10}$$

that is a consequence of definition of the canonical basis in the sub-space in which the irreducible representation with weight l is realized [Gelfand *et al.*, 1958]. Equations (A.8) and (A.10) reveal that any scalar function $F(t, \boldsymbol{x})$ that is a solution of the Laplace equation (A.1a) is given by

$$F(t, \boldsymbol{x}) = \sum_{l=0}^{\infty} \left[A_{<L>} \partial_L r^{-1} + B_{<L>} x^L\right] \,, \tag{A.11}$$

where $A_{<L>}$ and $B_{<L>}$ are STF multipole moments depending on time t only, and $\partial_L = \partial_{k_1} \partial_{k_2} ... \partial_{k_l}$ is the operator of l partial derivatives. Two terms in the right side of equation (A.11) behave differently as the radial distance goes to zero and to infinity. These two solutions are singled out by the boundary conditions imposed on the solution. If the solution is supposed to be convergent at the coordinate origin, the coefficients $A_{<L>} = 0$. On the other hand, if the solution is supposed to behave well at infinity, the coefficients $B_{<L>} = 0$. Notice that the solution with coefficients $A_{<L>}$ consists of a set of partial derivatives taken from the fundamental solution $1/r$ of the Laplace equation.

Vector and tensor spherical harmonics are obtained from the direct product of two irreducible representations of the rotation group with weights n and k which can be expanded into irreducible representations with weights $|n-k| \le l \le |n+k|$. Canonical orthonormal basis in the sub-space of the vector spherical harmonics in which the irreducible representation has weight l, is the set of $3(2l+1)$ functions

$$\mathbf{Y}^{l-1,\,lm}(\theta,\phi) = \sum_{p=-l-1}^{l-1}\sum_{q=-1}^{1} B^{lm}_{1,q;l-1,p}\boldsymbol{\xi}_q Y^{l-1\,p}(\theta,\phi)\,, \tag{A.12a}$$

$$\mathbf{Y}^{l,\,lm}(\theta,\phi) = \sum_{p=-l}^{l}\sum_{q=-1}^{1} B^{lm}_{1,q;l,p}\boldsymbol{\xi}_q Y^{lp}(\theta,\phi)\,, \tag{A.12b}$$

$$\mathbf{Y}^{l+1,\,lm}(\theta,\phi) = \sum_{p=-l+1}^{l+1}\sum_{q=-1}^{1} B^{lm}_{1,q;l+1,p}\boldsymbol{\xi}_q Y^{l+1\,p}(\theta,\phi)\,, \tag{A.12c}$$

(A.12d)

where notation $B^{lm}_{k,q;n,p}$ stands for the Clebsch-Gordan coefficients [Gelfand *et al.*, 1958, page 148, equation 16], and three unit vectors

$$\boldsymbol{\xi}_{-1} = \frac{\mathbf{e}_x - i\mathbf{e}_y}{\sqrt{2}}\quad,\quad \boldsymbol{\xi}_0 = \mathbf{e}_z\quad,\quad \boldsymbol{\xi}_1 = -\frac{\mathbf{e}_x + i\mathbf{e}_y}{\sqrt{2}} \tag{A.13}$$

represent a canonical basis of the main matrix representation of order 1.

Canonical basis in the three-dimensional space of constant tensors of second rank is made of 9 tensors such that five of them,

$$\mathbf{t}^m = \sum_{p=-1}^{1}\sum_{q=-1}^{1} B^{2m}_{1,q;1,p}\boldsymbol{\xi}_p \otimes \boldsymbol{\xi}_q\,,\qquad (-2 \le m \le 2) \tag{A.14}$$

are symmetric trace-free tensors; three tensors,

$$\mathbf{P}^m = \sum_{p=-1}^{1}\sum_{q=-1}^{1} B^{1m}_{1,q;1,p}\boldsymbol{\xi}_p \otimes \boldsymbol{\xi}_q\,,\qquad (-1 \le m \le 1) \tag{A.15}$$

are fully antisymmetric, and one,

$$\boldsymbol{\delta} = \sum_{p=-1}^{1}\sum_{q=-1}^{1} B^{00}_{1,q;1,p}\boldsymbol{\xi}_p \otimes \boldsymbol{\xi}_q\,, \tag{A.16}$$

is the unit tensor. The antisymmetric part of the basis will be abandoned as one is interested only in the symmetric tensors. Hence, the canonical basis in the sub-space of such tensors of the second rank with irreducible representation with weight l is formed from $6(2l+1)$ tensor harmonics

$$\mathbf{T}^{2n,\,lm} = \sum_{p=-n}^{n}\sum_{q=-2}^{2} B^{lm}_{2,q;n,p} Y^{np}\mathbf{t}^{q}\,, \tag{A.17}$$

where each n takes one of the values $l-2 \le n \le l+2$, along with $2l+1$ scalar spherical harmonics

$$\mathbf{T}^{0l,\,lm} = Y^{lm}\boldsymbol{\delta}\,. \tag{A.18}$$

Finally, solutions of the Laplace equations for vector, $F_i = F_i(t,\boldsymbol{x})$, and tensor, $F_{ij} = F_{ij}(t,\boldsymbol{x})$, functions (A.1a) are given as follows

$$\begin{aligned} F_i(t,\boldsymbol{x}) &= \sum_{l=0}^{\infty}\left[C_{<L>}\partial_{iL}r^{-1} + D_{<L>}x^{iL}\right] \\ &+ \sum_{l=1}^{\infty}\left[G_{<iL-1>}\partial_{L-1}r^{-1} + H_{<iL-1>}x^{L-1}\right] \\ &+ \sum_{l=1}^{\infty}\varepsilon_{ipq}\left[E_{<qL-1>}\partial_{pL-1}r^{-1} + F_{<qL-1>}x^{pL-1}\right]\,, \end{aligned} \tag{A.19}$$

$$\begin{aligned} F_{ij}(t,\boldsymbol{x}) &= \sum_{l=0}^{\infty}\left[I_{<L>}\partial_{L}r^{-1} + J_{<L>}x^{L}\right]\delta_{ij} \\ &+ \sum_{l=0}^{\infty}\left[K_{<L>}\partial_{ijL}r^{-1} + M_{<L>}x^{<ijL>}\right] \\ &+ \sum_{l=2}^{\infty}\left[V_{<ijL-2>}\partial_{L-2}r^{-1} + W_{<ijL-2>}x^{L-2}\right] \\ &+ \sum_{l=1}^{\infty}\left[N_{<iL-1>}\partial_{jL-1}r^{-1} + P_{<iL-1>}x^{<jL-1>}\right]^{\mathrm{Sym}(ij)} \\ &+ \sum_{l=2}^{\infty}\left[\varepsilon_{ipq}\left(S_{<qjL-2>}\partial_{pL-2}r^{-1} + T_{<qjL-1>}x^{<pL-2>}\right)\right]^{\mathrm{Sym}(ij)} \\ &+ \sum_{l=1}^{\infty}\left[\varepsilon_{ipq}\left(Q_{<qL-1>}\partial_{jpL-1}r^{-1} + R_{<qL-1>}x^{<jpL-1>}\right)\right]^{\mathrm{Sym}(ij)}\,, \end{aligned} \tag{A.20}$$

where the symbol $[...]^{\mathrm{Sym}}(ij)$ denotes symmetrization of the terms enclosed to the brackets with respect to indices i and j; coefficients $C_L, D_L, ..., T_L$ represent STF multipole moments depending on time t only, and multi-indices $L-1 \equiv k_1k_2...k_{l-1}$, $L-2 \equiv k_1k_2...k_{l-2}$.

B
Astronomical Constants

This appendix does not contain a list of *adopted* fundamental astronomical constants, because the IAU is no longer maintaining such a list. The last set of officially adopted constant values was the IAU (1976) System of Astronomical Constants. That list is almost entirely obsolete. For a while, an IAU working group maintained a list of "best estimates" of various constant values, but the IAU General Assembly of 2003 did not renew that mandate. It can be argued that a list of fundamental astronomical constants is no longer possible, given the complexity of the models now used and the many free parameters that must be adjusted in each model to fit observations. That is, there are more constants now to consider, and their values are theory dependent. In many cases, it would be incorrect to attempt to use a constant value, obtained from the fit of one theory to observations, with another theory.

One is left with three *defining constants* with IAU-sanctioned values that are intended to be fixed:

1) The Gaussian gravitational constant: $k = 0.01720209895$. The dimensions of k^2 are $\mathrm{AU}^3\mathrm{M}_\odot^{-1}\mathrm{d}^{-2}$ where AU is the astronomical unit, $\mathrm{M}_\odot$ is the solar mass, and d is the day of 86400 seconds.
2) The speed of light: $c = 299\,792\,458\ \mathrm{m\,s^{-1}}$.
3) The fractional difference in rate between the time scales TT and TCG: $\mathrm{L}_G = 6.969290134\times10^{-10}$. Specifically, the derivative $d\mathrm{TT}/d\mathrm{TCG} = 1-\mathrm{L}_G$. (See Chapter 9.4.)

The [McCarthy and Petit, 2004] includes a list of constants as its Table 1.1. Several useful ones from this list that are not highly theory dependent (for astronomical use, at least) are:

1) Equatorial radius of the Earth: $a_E = 6\,378\,136.6$ m.
2) Flattening factor of the Earth: $f = 1/298.25642$.
3) Dynamical form factor of the Earth: $J_2 = 1.0826359\times10^{-3}$.
4) Nominal mean angular velocity of Earth rotation: $\omega = 7.292115\times10^{-5}\ \mathrm{rad\,s^{-1}}$.
5) Constant of gravitation: $G = 6.673\times10^{-11}\ \mathrm{m^3kg^{-1}s^{-2}}$.
(CODATA 2002 recommended value: $6.6742\times10^{-11}\ \mathrm{m^3kg^{-1}s^{-2}}$).

The first four values above were recommended by Special Commission 3 of the International Association of Geodesy; the first three are so-called "zero tide" values. The need to introduce the concept of "zero tide" values indicates how theory creeps into even such basic constants as the radius of the Earth as the precision of measurement increases. See section 1.1 of the [McCarthy and Petit, 2004]. Planetary masses, the length of the astronomical unit, and related constants used in or obtained from the Jet Propulsion Laboratory DE405/LE405 ephemeris are listed with its description in Chapter 9.5.1. The rate of general precession in longitude (the "constant of precession") is given in Chapter 9.6 on the precession and nutation theories.

The World Geodetic System 1984 (WGS 84), which is the basis for coordinates obtained from GPS, uses an Earth ellipsoid with $a_E = 6378137$m and $f = 1/298.257223563$.

Some astronomical "constants" (along with reference data such as star positions) actually represent quantities that slowly vary, and the values given must therefore be associated with a specific epoch. That epoch is now almost always 2000 January 1, 12^{h} (JD 2451545.0), which can be expressed in any of the usual time scales. If, however, that epoch is considered an event at the geocenter and given in the TT time scale, the epoch is designated J2000.0. See Chapter 9.3.

The table below yields a list of the best estimates of the astronomical constants compiled by Dr. B. Luzum from the United States Naval Observatory on the basis of references cited at the end of the table.

Defining Constants

Constant	Description	Value	Reference
		Natural Defining Constant	
c	Speed of light	$2.997\,924\,58 \times 10^{8}$ ms^{-1}	[7]
		Auxiliary Defining Constants	
k[1]	Gaussian gravitational constant	$1.720\,209\,895 \times 10^{-2}$	[14, 11]
L_G	1–d(TT)/d(TCG)	$6.969\,290\,134 \times 10^{-10}$	[15, 25]
L_B	1–d(TDB)/d(TCB)	$1.550\,519\,768 \times 10^{-8}$	[16]
TDB_0[2]	TDB –TCB at T_0	-6.55×10^{-5} s	[16]
θ_0[3]	Earth Rotation Angle at J2000.0	0.779 057 273 264 0 revolutions	[15, 4]
$d\theta/dt$[3]	Rate of advance of Earth Rotation Angle	1.002 737 811 911 354 48 revolutions UT1-day^{-1}	[15, 4]

Current Best Estimates

Constant	Description	Value	Uncertainty	Reference
		Natural Measurable Constants		
G	Constant of gravitation	$6.674\,28 \times 10^{-11}$ m^{3}kg^{-1}s^{-2}	6.7×10^{-15} m^{3}kg^{-1}s^{-2}	[7]
		Derived Constants		
au[4]	Astronomical unit	$1.495\,978\,707\,00 \times 10^{11}$ m	3 m	[26]
L_C	Average value of 1–d(TCG)/d(TCB)	$1.480\,826\,867\,41 \times 10^{-8}$	2×10^{-17}	[18]
		Body Constants[5]		
M_M/M_E	Ratio of the mass of the Moon to the Earth	$1.230\,003\,71 \times 10^{-2}$	4×10^{-10}	[26]
M_S/M_{Me}	Ratio of the mass of the Sun to Mercury	$6.023\,6 \times 10^{6}$	3×10^{2}	[1]
M_S/M_{Ve}	Ratio of the mass of the Sun to Venus	$4.085\,237\,19 \times 10^{5}$	8×10^{-3}	[23]
M_S/M_{Ma}	Ratio of the mass of the Sun to Mars	$3.098\,703\,59 \times 10^{6}$	2×10^{-2}	[24]
M_S/M_J	Ratio of the mass of the Sun to Jupiter	$1.047\,348\,644 \times 10^{3}$	1.7×10^{-5}	[20]
M_S/M_{Sa}	Ratio of the mass of the Sun to Saturn	$3.497\,901\,8 \times 10^{3}$	1×10^{-4}	[21]
M_S/M_U	Ratio of the mass of the Sun to Uranus	$2.290\,298 \times 10^{4}$	3×10^{-2}	[19]

M_S/M_N	Ratio of the mass of the Sun to Neptune	$1.941\,226 \times 10^4$	3×10^{-2}	[22]
M_S/M_P	Ratio of the mass of the Sun to Pluto	$1.365\,66 \times 10^8$	2.8×10^4	[29]
M_S/M_{Eris}	Ratio of the mass of the Sun to Eris	1.191×10^8	1.4×10^6	[2]
M_{Ceres}/M_S	Ratio of the mass of Ceres to the Sun	4.72×10^{-10}	3×10^{-12}	[26]
M_{Pallas}/M_S	Ratio of the mass of Pallas to the Sun	1.03×10^{-10}	3×10^{-12}	[26]
M_{Vesta}/M_S	Ratio of the mass of Vesta to the Sun	1.35×10^{-10}	3×10^{-12}	[26]
a_E[6]	Equatorial radius of the Earth	$6.378\,136\,6 \times 10^6$ m	1×10^{-1} m	[12, 3]
J_2[6]	Dynamical form factor	$1.082\,635\,9 \times 10^{-3}$	1×10^{-10}	[12]
dJ_2/dt	Long-term variation in J_2	-3.001×10^{-9} cy^{-1}	6×10^{-10} cy^{-1}	[16]
GM_S	Heliocentric gravitational constant	$1.327\,124\,420\,99 \times 10^{20}$ m^{3}s^{-2} (TCB-compatible) $1.327\,124\,400\,41 \times 10^{20}$ m^{3}s^{-2} (TDB-compatible)	1.0×10^{10} m^{3}s^{-2} (TCB-compatible) 1.0×10^{10} m^{3}s^{-2} (TDB-compatible)	[8]
GM_E	Geocentric gravitational constant	$3.986\,004\,418 \times 10^{14}$ m^{3}s^{-2} (TCB-compatible) $3.986\,004\,415 \times 10^{14}$ m^{3}s^{-2} (TT-compatible) $3.986\,004\,356 \times 10^{14}$ m^{3}s^{-2} (TDB-compatible)	8×10^5 m^{3}s^{-2} (TCB-compatible) 8×10^5 m^{3}s^{-2} (TT-compatible) 8×10^5 m^{3}s^{-2} (TDB-compatible)	[27]
W_0	Potential of the geoid	$6.263\,685\,60 \times 10^7$ m^{2}s^{-2}	5×10^{-1} m^{2}s^{-2}	[12]
ω[7]	Nominal mean angular velocity of the Earth	$7.292\,115 \times 10^{-5}$ rad s^{-1}		[12]
Initial Values at J2000.0				
ε_{J2000}[8]	Obliquity of the ecliptic at J2000.0	$8.438\,140\,6 \times 10^4$ ″	1×10^{-3} ″	[16, 13, 6]

Notes:

1. The Gaussian gravitational constant, k, defines au.
2. This constant comes from the expression TDB = TCB – L_B × (JD_{TCB} – T_0) × 86400 + TDB_0, where T_0 = 2443144.5003725.
3. This constant comes from the expression θ(UT1) = 2π(0.7790572732640 + 1.00273781191135448 × (Julian UT1 date – 2451545.0).
4. The value for au is TDB-compatible. An accepted definition for the TCB-compatible value of au is still under discussion.
5. All values of the masses from Mars through Eris are the sum of the masses of the celestial body and its satellites.

6. The values for a_E and J_2 are "zero tide" values (see IERS Conventions for an explanation of the terminology). Values according to other conventions can be found in Groten *et al*. (2000).
7. ω is a nominal value and was chosen to have the number of significant digits limited to those for which the value can be considered constant.
8. ε_{J2000} is a component of the IAU 2006 precession model, which includes expressions that are time dependent.
9. The rate of precession appearing in previous lists of constants is no longer appropriate given the IAU 2006 precession model [16].

References

1. Anderson, J.D., Colombo, G., Esposito, P.B., Lau, E.L., and Trager, G.B., 1987, "The Mass Gravity Field and Ephemeris of Mercury," *Icarus*, **71**, pp. 337-349.
2. Brown, M.E. and Schaller, E.L., 2007, "The mass of Dwarf Planet Eris," *Science*, **316**, p. 1585, doi: 10.1126/science.1139415.
3. Bursa, M., Kouba, J., Radej, K., True, S.A., Vatrt, V., Vojtiskova, M., 1998, "Mean Earth's Equipotential Surface from Topex/Poseidon Altimetry," *Studia Geoph. Et Geod.*, **42**, pp. 459-466, doi: 10.1023/A:1023356803773.
4. Capitaine, N., Guinot, B., and McCarthy, D.D., 2000, "Definition of the Celestial Ephemeris Origin and of UT1 in the International Celestial Reference Frame," *Astron. Astrophys.*, **355**, pp. 398-405.
5. Capitaine, N., Wallace, P., and Chapront, J., 2003, "Expressions for IAU 2000 precession quantities," *Astron. Astrophys.*, **412**, pp. 567-586.
6. Chapront, J., Chapront-Touzé, M., and Francou, G., 2002, "A new determination of lunar orbital parameters, precession constant and tidal acceleration from LLR measurements," *Astron. Astrophys.*, **387**, pp. 700-709, doi 10.1051/0004-6361:20020420.
7. CODATA, 2006, http://physics.nist.gov/cuu/Constants/index.html.
8. Folkner, W.M., Williams, J.G., and Boggs, D.H., 2008, "The Planetary and Lunar Ephemeris DE 421," Memorandum IOM 343R-08-003, 31 pp.
9. Fukushima, T., 2000, "Report on Astronomical Constants," *Proc. IAU Colloquium 180*, Johnston, K.J., McCarthy, D.D., Luzum, B.J., and Kaplan, G.H. (eds.), pp. 417-427.
10. Fukushima, T., 2003, "Report on Astronomical Constants," *Highlights of Astronomy*, **Vol. 13**, International Astronomical Union, 2002, Rickman, H. (ed.), pp. 107-112.
11. Gauss, C.F., 1857, *Theory of the Motion of the Heavenly Bodies Moving About the Sun in Conic Sections*, Boston: Little, Brown, and Company, p. 2.
12. Groten, E., 2000, Geodesists Handbook 2000, part 4, http://www.gfy.ku.dk/~iag/HB2000/part4/groten.htm. See also "Parameters of Common Relevance of Astronomy, Geodesy, and Geodynamics," *J. Geod.*, **74**, pp. 134-140.
13. Hilton, J.L., Capitaine, N., Chapront, J., Ferrandiz, J.M., Fienga, A., Fukushima, T., Getino, J., Mathews, P., Simon, J.-L., Soffel, M., Vondrak, J., Wallace, P., and Williams, J., 2006, "Report of the International Astronomical Union Division I

Working Group on Precession and the Ecliptic," *Celest. Mech. Dyn. Astr.*, **94**, pp. 351-367, doi 10.1007/s10569-006-0001-2.
14. International Astronomical Union (IAU), 1976, "Proceedings of the Sixteenth General Assembly," Transactions of the IAU, XVIB, p. 31, pp. 52-66.
15. International Astronomical Union (IAU), 2000, "Proceedings of the Twenty-Fourth General Assembly," Transactions of the IAU, XXIVB, pp. 34-57.
16. International Astronomical Union (IAU), 2006, "Proceedings of the Twenty-Sixth General Assembly," Transactions of the IAU, XXVIB.
17. IERS Conventions, McCarthy, D.D. and Petit, G., 2003, IERS Technical Note 32, Verlag des Bundesamts für Kartographie und Geodäsie, Frankfurt am Main, 127 pp.
18. Irwin, A. and Fukushima, T., 1999, "A numerical time ephemeris of the Earth," *Astron. Astrophys.*, **348**, pp. 642-652.
19. Jacobson, R.A., Campbell, J.K., and Taylor, A.H., 1992, "The Masses of Uranus and its Major Satellites from Voyager Tracking Data and Earth-based Uranian Satellite Data," *Astron. J.*, **103(6)**, pp. 2068-2078.
20. Jacobson, R.A., Haw, R.J., McElrath, T.P., and Antreasian, P.G., 2000, "A Comprehensive Orbit Reconstruction for the Galileo Prime Mission in the J2000 System," *J. Astronaut. Sci.*, **48(4)**, pp. 495-516.
21. Jacobson, R.A., Antreasian, P.G., Bordi, J.J., Criddle, K.E., Ionasescu, R., Jones, J.B., Mackenzie, R.A., Pelletier, F.J., Owen, Jr., W.M., Roth, D.C., and Stauch, J.R., 2006, "The gravity field of the Saturnian system from satellite observations and spacecraft tracking data," *Astron. J.*, **132(6)**, pp. 2520-2526.
22. Jacobson, R.A., 2009, "The Orbits of the Neptunian Satellites and the Orientation of the Pole of Neptune," *Astron. J.*, **137**, pp. 4322-4329, doi:10.1088/0004-6256/137/4322.
23. Konopliv, A.S., Banerdt, W.B., and Sjogren, W.L., 1999, "Venus Gravity: 180th Degree and Order Model," *Icarus*, **139**, pp. 3-18.
24. Konopliv, A.S., Yoder, C.F., Standish, E.M., Yuan, D.N., and Sjogren, W.L., 2006, "A global solution for the Mars static and seasonal gravity, Mars orientation, and Phobos and Deimos masses," *Icarus*, **182(1)**, pp. 23-50.
25. Petit, G., 2000, "Report of the BIPM/IAU Joint Committee on relativity for space-time reference systems and metrology," in *Proc. IAU Colloquium 180*, Johnston, K.J., McCarthy, D.D., Luzum, B.J., and Kaplan, G.H. (eds.), U.S. Naval Observatory, Washington, D.C. pp. 275-282.
26. Pitjeva, E. and Standish, E.M., 2009, "Proposals for the masses of the three largest asteroids, the Moon-Earth mass ratio and the astronomical unit," *Celest. Mech. Dyn. Astr.*, **103**, pp. 365-372, doi:10.1007/s10569-009-9203-8.
27. Ries, J.C., Eanes, R.J., Shum, C.K., and Watkins, M.M., 1992, "Progress in the Determination of the Gravitational Coefficient of the Earth," *Geophys. Res. Lett.*, **19(6)**, pp. 529-531.
28. Standish, E.M., 1995, "Report of the IAU WGAS Sub-group on Numerical Standards," *Highlights in Astronomy*, Vol. 12, International Astronomical Union, 1994, Appenzeller, I. (ed.), pp. 180-184.
29. Tholen, D.J., Buie, M.W., and Grundy, W., 2008, "Masses of Nix and Hydra" *Astron. J.*, **135(3)**, pp. 777-784.

C
Text of IAU Resolutions

The full text of recent resolutions of the International Astronomical Union (IAU) regarding fundamental reference systems, time scales, and Earth orientation models is given on the following pages. The resolutions are grouped by date of adoption (1997, 2000, 2006, and 2009), not subject matter.

Some of the later resolutions clarify or even supersede previous ones. For example, the terminology of Celestial (or Terrestrial) Ephemeris Origin used in resolution B1.8 of 2000 was replaced by Celestial (or Terrestrial) Intermediate Origin in resolution B2 (recommendation 1) of 2006. Resolution B1.6 of 2000 "encourages...the development of new expressions for precession consistent with the IAU 2000A model" but the final precession model was not adopted until resolution B1 of 2006. Similarly, the resolutions that established the BCRS in 2000 left its spatial orientation undefined; the orientation was specified in 2006 as that of the ICRS. It could be argued that the BCRS and ICRS should have been specified at the same time, since they are so closely linked (the BCRS is the relativistic basis for the ICRS; the ICRS is the spatial orientation of the BCRS), and possibly given a single name that encompasses the complete definition.

Legislation of this kind is rarely a complete or perfect representation of the underlying science, and compromises, poor wording, and unintended consequences sometimes creep into the process. A particular concern is specifications that rely on references to papers not yet published at the time the resolution was adopted or to web sites with content that changes.

C.1

Text of IAU Resolutions of 1997 Adopted at the XXIII-rd General Assembly, Kyoto

Resolution B2 On the International Celestial Reference System (ICRS)

The XXIIIrd International Astronomical Union General Assembly

Considering

(a) That Recommendation VII of Resolution A4 of the 21st General Assembly specifies the coordinate system for the new celestial reference frame and, in particular, its continuity with the FK5 system at J2000.0;

(b) That Resolution B5 of the 22nd General Assembly specifies a list of extragalactic sources for consideration as candidates for the realization of the new celestial reference frame;

(c) That the IAU Working Group on Reference Frames has in 1995 finalized the positions of these candidate extragalactic sources in a coordinate frame aligned to that of the FK5 to within the tolerance of the errors in the latter (see note 1);

(d) That the Hipparcos Catalogue was finalized in 1996 and that its coordinate frame is aligned to that of the frame of the extragalactic sources in (c) with one sigma uncertainties of ± 0.6 milliarcseconds (mas) at epoch J1991.25 and ± 0.25 mas per year in rotation rate;

Noting

That all the conditions in the IAU Resolutions have now been met;

Resolves

(a) That, as from 1 January 1998, the IAU celestial reference system shall be the International Celestial Reference System (ICRS) as specified in the 1991 IAU Resolution on reference frames and as defined by the International Earth Rotation Service (IERS) (see note 2);

(b) That the corresponding fundamental reference frame shall be the International Celestial Reference Frame (ICRF) constructed by the IAU Working Group on Reference Frames;

(c) That the Hipparcos Catalogue shall be the primary realization of the ICRS at optical wavelengths;

(d) That IERS should take appropriate measures, in conjunction with the IAU Working Group on reference frames, to maintain the ICRF and its ties to the reference frames at other wavelengths.

Note 1: IERS 1995 Report, Observatoire de Paris, p. II-19 (1996).

Note 2: "The extragalactic reference system of the International Earth Rotation Service (ICRS)", Arias, E.F. et al. A&A 303, 604 (1995).

Resolution B4 On Non-Rigid Earth Nutation Theory

The XXIIIrd International Astronomical Union General Assembly

Recognizing

that the International Astronomical Union and the International Union of Geodesy and Geophysics Working Group (IAU-IUGG WG) on Non-rigid Earth Nutation Theory has met its goal by identifying the remaining geophysical and astronomical phenomena that must be modeled before an accurate theory of nutation for a non-rigid Earth can be adopted, and

that, as instructed by IAU Recommendation C1 in 1994, the International Earth Rotation Service (IERS) has published in the IERS Conventions (1996) an interim precession-nutation model that matches the observations with an uncertainty of ± 1 milliarcsecond (mas),

endorses

the conclusions of the IAU-IUGG WG on Non-rigid Earth Nutation Theory given in the appendix,

requests

the IAU-IUGG WG on Non-rigid Earth Nutation Theory to present a detailed report to the next IUGG General Assembly (August 1999), at which time the WG will be discontinued,

and urges

the scientific community to address the following questions in the future:

- completion of a new rigid Earth nutation series with the additional terms necessary for the theory to be complete to within ± 5 microarcseconds, and

- completion of a new non-rigid Earth transfer function for an Earth initially in non-hydrostatic equilibrium, incorporating mantle inelasticity and a Free Core Nutation period in agreement with the observations, and taking into account better modeling of the fluid parts of the planet, including dissipation.

APPENDIX

The WG on Non-rigid Earth Nutation Theory has quantified the problems in the nutation series adopted by the IAU in 1980 by noting:

(1) that there is a difference in the precession rate of about −3.0milliarcseconds per year (mas/year) between the value observed by Very Long Baseline Interferometry (VLBI) and Lunar Laser Ranging (LLR) and the adopted value,

(2) that the obliquity has been observed (by VLBI and LLR) to change at a rate of about −0.24mas/year, although there is no such change implied by the 1980

precession-nutation theory,

(3) that, in addition to these trends, there are observable peak-to-peak differences of up to 20 milliarcseconds (mas) between the nutation observed by VLBI and LLR and the nutation adopted by the IAU in 1980,

(4) that these differences correspond to spectral amplitudes of up to several mas, and

(5) that the differences between observation and theory are well beyond the present observational accuracy.

The WG has recognized the improvements made in the modeling of these quantities, and recommends, in order to derive a more precise nutation model, at the mas level in spectral amplitudes and at a few mas level in the peak to peak analysis, the use of models:

(1) based on a new non-rigid Earth transfer function for an Earth initially in non-hydrostatic equilibrium, incorporating mantle inelasticity, a core-mantle-boundary flattening giving a Free Core Nutation (FCN) period in agreement with the observed value, and a global Earth dynamical flattening in agreement with the observed precession, and

(2) based on a new rigid Earth nutation series which takes into account the following perturbing effects:

1. in lunisolar ephemerides: indirect planetary effects, lunar inequality, J2-tilt, planetary-tilt, secular variations of the amplitudes, effects of precession and nutation,
2. in the perturbing bodies to be considered: in addition to the Moon and the Sun, the direct planetary effects of Venus, Jupiter, Mars, and Saturn, should be included,
3. in the order of the external potential to be considered: J3 and J4 effects for the Moon, and
4. in the theory itself: effects of the tri-axiality of the Earth, relativistic effects and second order effects.

The WG recognizes that this new generation of models still has some imperfections, the principal one being poor modeling of the dissipation in the core and of certain effects of the ocean and the atmosphere, and urges the scientific community to address these questions in the future.

The WG recognizes that, due to the remaining imperfections of the present theoretical nutation models, the nutation series published in the IERS Conventions (1996), following 1994 IAU recommendation C1, still provides the users with the best nutation series. This IERS model being based on observations of the celestial pole offset, the WG supports the recommendation that the scientific community continue VLBI and LLR observations to provide accurate estimations of nutation, precession and rate of change in obliquity.

C.2

Text of IAU Resolutions of 2000 Adopted at the XXIV-th General Assembly, Manchester

Resolution B1.1 Maintenance and Establishment of Reference Frames and Systems

The XXIVth International Astronomical Union

Noting

1. that Resolution B2 of the XXIIIrd General Assembly (1997) specifies that "the fundamental reference frame shall be the International Celestial Reference Frame (ICRF) constructed by the IAU Working Group on Reference Frames,"

2. that Resolution B2 of the XXIIIrd General Assembly (1997) specifies "That the Hipparcos Catalogue shall be the primary realization of the ICRS at optical wavelengths", and

3. the need for accurate definition of reference systems brought about by unprecedented precision, and

Recognizing

1. the importance of continuing operational observations made with Very Long Baseline Interferometry (VLBI) to maintain the ICRF,

2. the importance of VLBI observations to the operational determination of the parameters needed to specify the time-variable transformation between the International Celestial and Terrestrial Reference Frames,

3. the progressive shift between the Hipparcos frame and the ICRF, and

4. the need to maintain the optical realization as close as possible to the ICRF

Recommends

1. that IAU Division I maintain the Working Group on Celestial Reference Systems formed from Division I members to consult with the International Earth Rotation Service (IERS) regarding the maintenance of the ICRS,

2. that the IAU recognize the International VLBI service (IVS) for Geodesy and

Astrometry as an IAU Service Organization,

3. that an official representative of the IVS be invited to participate in the IAU Working Group on Celestial Reference Systems,

4. that the IAU continue to provide an official representative to the IVS Directing Board,

5. that the astrometric and geodetic VLBI observing programs consider the requirements for maintenance of the ICRF and linking to the Hipparcos optical frame in the selection of sources to be observed (with emphasis on the Southern Hemisphere), design of observing networks, and the distribution of data, and

6. that the scientific community continue with high priority ground- and space-based observations (a) for the maintenance of the optical Hipparcos frame and frames at other wavelengths and (b) for the links of the frames to the ICRF.

Resolution B1.2 Hipparcos Celestial Reference Frame

The XXIVth International Astronomical Union

Noting

1. that Resolution B2 of the XXIIIrd General Assembly (1997) specifies, "That the Hipparcos Catalogue shall be the primary realization of the International Celestial Reference System (ICRS) at optical wavelengths,"

2. the need for this realization to be of the highest precision,

3. that the proper motions of many of the Hipparcos stars known, or suspected, to be multiple are adversely affected by uncorrected orbital motion,

4. the extensive use of the Hipparcos Catalogue as reference for the ICRS in extension to fainter stars,

5. the need to avoid confusion between the International Celestial Reference Frame (ICRF) and the Hipparcos frame, and

6. the progressive shift between the Hipparcos frame and the ICRF,

Recommends

1. that Resolution B2 of the XXIIIrd IAU General Assembly (1997) be amended by excluding from the optical realization of the ICRS all stars flagged C, G, O, V and X in the Hipparcos Catalogue, and

2. that this modified Hipparcos frame be labeled the Hipparcos Celestial Reference Frame (HCRF).

Resolution B1.3 Definition of Barycentric Celestial Reference System and Geocentric Celestial Reference System

The XXIVth International Astronomical Union

Considering

1. that the Resolution A4 of the XXIst General Assembly (1991) has defined a system of spacetime coordinates for (a) the solar system (now called the Barycentric Celestial Reference System, (BCRS)) and (b) the Earth (now called the Geocentric Celestial Reference System (GCRS)), within the framework of General Relativity,

2. the desire to write the metric tensors both in the BCRS and in the GCRS in a compact and self-consistent form, and

3. the fact that considerable work in General Relativity has been done using the harmonic gauge that was found to be a useful and simplifying gauge for many kinds of applications,

Recommends

1. the choice of harmonic coordinates both for the barycentric and for the geocentric reference systems.

2. writing the time-time component and the space-space component of the barycentric metric $g_{\mu\nu}$ with barycentric coordinates $(t, \boldsymbol{x})$ (t = Barycentric Coordinate Time (TCB)) with a single scalar potential $w(t, \boldsymbol{x})$ that generalizes the Newtonian potential, and the spacetime component with a vector potential $w^i(t, \boldsymbol{x})$; as a boundary condition it is assumed that these two potentials vanish far from the solar system, explicitly,

$$
\begin{aligned}
g_{00} &= -1 + \frac{2w}{c^2} - \frac{2w^2}{c^4} \,, \\
g_{0i} &= -\frac{4}{c^3} w^i \,, \\
g_{ij} &= \delta_{ij}\left(1 + \frac{2}{c^2} w\right) ,
\end{aligned}
$$

with

$$
\begin{aligned}
w(t, \boldsymbol{x}) &= G \int d^3x' \frac{\sigma(t, \boldsymbol{x}')}{|\boldsymbol{x} - \boldsymbol{x}'|} + \frac{G}{2c^2} \frac{\partial^2}{\partial t^2} \int d^3x' \sigma(t, \boldsymbol{x}') |\boldsymbol{x} - \boldsymbol{x}'| \,, \\
w^i(t, \boldsymbol{x}) &= G \int d^3x' \frac{\sigma^i(t, \boldsymbol{x}')}{|\boldsymbol{x} - \boldsymbol{x}'|} \,.
\end{aligned}
$$

Here, σ and σ^i are the gravitational mass and current densities, respectively.

3. writing the geocentric metric tensor $G_{\alpha\beta}$ with geocentric coordinates $(T, \boldsymbol{X})$ (T= Geocentric Coordinate Time (TCG)) in the same form as the barycentric one but with potentials $W(T, \boldsymbol{X})$ and $W^a(T, \boldsymbol{X})$; these geocentric potentials should be split into two parts – potentials W and W^a arising from the gravitational action of the Earth and external parts W_{ext} and W^a_{ext} due to tidal and inertial effects; the external parts of the metric potentials are assumed to vanish at the geocenter and admit an expansion into positive powers of $\boldsymbol{X}$,

explicitly,

$$
\begin{aligned}
G_{00} &= -1 + \frac{2W}{c^2} - \frac{2W^2}{c^4} , \\
G_{0a} &= -\frac{4}{c^3} W^a , \\
G_{ab} &= \delta_{ab}\left(1 + \frac{2}{c^2} W\right) .
\end{aligned}
$$

The potentials W and W^a should be split according to

$$
\begin{aligned}
W(T, \boldsymbol{X}) &= W_{\text{E}}(T, \boldsymbol{X}) + W_{\text{ext}}(T, \boldsymbol{X}) , \\
W^a(T, \boldsymbol{X}) &= W^a_{\text{E}}(T, \boldsymbol{X}) + W^a_{\text{ext}}(T, \boldsymbol{X}) .
\end{aligned}
$$

The Earth's potentials W_{E} and W^a_{E} are defined in the same way as w and w^i but with quantities calculated in the GCRS with integrals taken over the whole Earth.

4. using, if accuracy requires, the full post-Newtonian coordinate transformation between the BCRS and the GCRS as induced by the form of the corresponding metric tensors,

explicitly, for the kinematically non-rotating GCRS (T=TCG, t=TCB, $r^i_{\text{E}} \equiv x^i - x^i_{\text{E}}(t)$ and a summation from 1 to 3 over equal indices is implied),

$$
\begin{aligned}
T &= t - \frac{1}{c^2}\left[A(t) + v^i_{\text{E}} r^i_{\text{E}}\right] + \frac{1}{c^4}\left[B(t) + B^i(t) r^i_{\text{E}} + B^{ij}(t) r^i_{\text{E}} r^j_{\text{E}} + C(t, \boldsymbol{x})\right] + O(c^{-5}) , \\
X^a &= \delta_{ai}\left[r^i_{\text{E}} + \frac{1}{c^2}\left(\frac{1}{2} v^i_{\text{E}} v^j_{\text{E}} r^j_{\text{E}} + w_{\text{ext}}(\boldsymbol{x}_{\text{E}}) r^i_{\text{E}} + r^i_{\text{E}} a^j_{\text{E}} r^j_{\text{E}} - \frac{1}{2} a^i_{\text{E}} r^2_{\text{E}}\right)\right] + O\left(c^{-4}\right) ,
\end{aligned}
$$

where

$$
\begin{aligned}
\frac{dA(t)}{dt} &= \frac{1}{2}v_{\rm E}^2 + w_{\rm ext}(\boldsymbol{x}_{\rm E}) \ , \\
\frac{dB(t)}{dt} &= -\frac{1}{8}v_{\rm E}^4 - \frac{3}{2}v_{\rm E}^2 w_{\rm ext}(\boldsymbol{x}_{\rm E}) + 4v_{\rm E}^i w_{\rm ext}^i(\boldsymbol{x}_{\rm E}) + \frac{1}{2}w_{\rm ext}^2(\boldsymbol{x}_{\rm E}) \ , \\
B^i(t) &= -\frac{1}{2}v_{\rm E}^2 v_{\rm E}^i + 4w_{\rm ext}^i(\boldsymbol{x}_{\rm E}) - 3v_{\rm E}^i w_{\rm ext}(\boldsymbol{x}_{\rm E}) \ , \\
B^{ij}(t) &= -v_{\rm E}^i \delta_{aj} Q^a + 2\frac{\partial}{\partial x^j} w_{\rm ext}^i(\boldsymbol{x}_{\rm E}) - v_{\rm E}^i \frac{\partial}{\partial x^j} w_{\rm ext}(\boldsymbol{x}_{\rm E}) + \frac{1}{2}\delta^{ij}\dot{w}_{ext}(\boldsymbol{x}_{\rm E}) \ , \\
C(t, \boldsymbol{x}) &= -\frac{1}{10} r_{\rm E}^2 (\dot{a}_{\rm E}^i r_{\rm E}^i) \ .
\end{aligned}
$$

Here $x_{\rm E}^i$, $v_{\rm E}^i$, and $a_{\rm E}^i$ are the barycentric position, velocity and acceleration vectors of the Earth, the dot stands for the total derivative with respect to t, and

$$
Q^a = \delta_{ai}\left[\frac{\partial}{\partial x_i} w_{\rm ext}(\boldsymbol{x}_{\rm E}) - a_{\rm E}^i\right] \ .
$$

The external potentials, $w_{\rm ext}$ and $w_{\rm ext}^i$, are given by

$$
w_{\rm ext} = \sum_{\rm A \neq E} w_{\rm A}, \quad w_{\rm ext}^i = \sum_{\rm A \neq E} w_{\rm A}^i
$$

where E stands for the Earth and $w_{\rm A}$ and $w_{\rm A}^i$ are determined by the expressions for w and w^i with integrals taken over body A only.

Notes.

It is to be understood that these expressions for w and w^i give g_{00} correct up to $O(c^{-5})$, g_{0i} up to $O(c^{-5})$, and g_{ij} up to $O\left(c^{-4}\right)$. The densities σ and σ^i are determined by the components of the energy momentum tensor of the matter composing the solar system bodies as given in the references. Accuracies for $G_{\alpha\beta}$ in terms of c^{-n} correspond to those of $g_{\mu\nu}$.

The external potentials $W_{\rm ext}$ and $W_{\rm ext}^a$ can be written in the form

$$
\begin{aligned}
W_{\rm ext} &= W_{\rm tidal} + W_{\rm iner} \ , \\
W_{\rm ext}^a &= W_{\rm tidal}^a + W_{\rm iner}^a \ .
\end{aligned}
$$

$W_{\rm tidal}$ generalizes the Newtonian expression for the tidal potential. Post-Newtonian expressions for $W_{\rm tidal}$ and $W^a_{\rm tidal}$ can be found in the references. The potentials $W_{\rm iner}$, $W^a_{\rm iner}$ are inertial contributions that are linear in X^a. The former is determined mainly by the coupling of the Earth's non-sphericity to the external potential. In the kinematically non-rotating Geocentric Celestial Reference System, $W^a_{\rm iner}$ describes the Coriolis force induced mainly by geodetic precession.

Finally, the local gravitational potentials $W_{\rm E}$ and $W^a_{\rm E}$ of the Earth are related to the barycentric gravitational potentials $w_{\rm E}$ and $w^i_{\rm E}$ by

$$
\begin{aligned}
W_{\rm E}(T,\boldsymbol{X}) &= w_{\rm E}(t,\boldsymbol{x})\left(1+\frac{2}{c^2}v^2_{\rm E}\right)-\frac{4}{c^2}v^i_{\rm E}w^i_{\rm E}(t,\boldsymbol{x})+O\left(c^{-4}\right)\ ,\\
W^a_{\rm E}(T,\boldsymbol{X}) &= \delta_{ai}\left(w^i_{\rm E}(t,\boldsymbol{x})-v^i_{\rm E}w_{\rm E}(t,\boldsymbol{x})\right)+O\left(c^{-2}\right)\ .
\end{aligned}
$$

References:

Brumberg, V. A., Kopeikin, S. M., 1989, *Nuovo Cimento*, **B103**, 63.

Brumberg, V. A., 1991, *Essential Relativistic Celestial Mechanics*, Hilger, Bristol.

Damour, T., Soffel, M., Xu, C., *Phys. Rev. D*, **43**, 3273 (1991); **45**, 1017 (1992); **47**, 3124 (1993); **49**, 618 (1994).

Klioner, S. A., Voinov, A. V., 1993, *Phys Rev. D*, **48**, 1451.

Kopeikin, S. M., 1988, *Celest. Mech.*, **44**, 87.

Resolution B1.4 Post-Newtonian Potential Coefficients

The XXIVth International Astronomical Union

Considering

1. that for many applications in the fields of celestial mechanics and astrometry a suitable parametrization of the metric potentials (or multipole moments) outside the massive solar-system bodies in the form of expansions in terms of potential coefficients are extremely useful, and

2. that physically meaningful post-Newtonian potential coefficients can be derived from the literature,

Recommends

1. expansion of the post-Newtonian potential of the Earth in the Geocentric Celestial Reference System (GCRS) outside the Earth in the form

$$W_{\mathrm{E}}(T, \boldsymbol{X}) = \frac{GM_{\mathrm{E}}}{R}\left\{1 + \sum_{l=2}^{\infty}\sum_{m=0}^{+l}\left(\frac{R_{\mathrm{E}}}{R}\right)^{l} P_{lm}(\cos\theta)\left[C^{E}_{lm}(T)\cos m\phi + S^{E}_{lm}(T)\sin m\phi\right]\right\} ,$$

where C^{E}_{lm} and S^{E}_{lm} are, to sufficient accuracy, equivalent to the post-Newtonian multipole moments introduced in (Damour *et al.*, *Phys. Rev. D*, **43**, 3273, 1991), θ and ϕ are the polar angles corresponding to the spatial coordinates X^a of the GCRS and $R = |X|$, and

2. expression of the vector potential outside the Earth, leading to the well-known Lense-Thirring effect, in terms of the Earth's total angular momentum vector $\boldsymbol{S}_{\mathrm{E}}$ in the form

$$W^{a}_{\mathrm{E}}(T, \boldsymbol{X}) = -\frac{G}{2}\frac{(\boldsymbol{X}\times\boldsymbol{S}_{\mathrm{E}})^{a}}{R^{3}} .$$

Resolution B1.5 Extended relativistic framework for time transformations and realization of coordinate times in the solar system

The XXIVth International Astronomical Union

Considering

1. that the Resolution A4 of the XXIst General Assembly (1991) has defined systems of spacetime coordinates for the solar system (Barycentric Reference System) and for the Earth (Geocentric Reference System), within the framework of General Relativity,

2. that Resolution B1.3 entitled "Definition of Barycentric Celestial Reference System and Geocentric Celestial Reference System" has renamed these systems the Barycentric Celestial Reference System (BCRS) and the Geocentric Celestial Reference System (GCRS), respectively, and has specified a general framework for expressing their metric tensor and defining coordinate transformations at the first post-Newtonian level,

3. that, based on the anticipated performance of atomic clocks, future time and frequency measurements will require practical application of this framework in the BCRS, and

4. that theoretical work requiring such expansions has already been performed,

Recommends

that for applications that concern time transformations and realization of coordinate times within the solar system, Resolution B1.3 be applied as follows:

1. the metric tensor be expressed as

$$
\begin{aligned}
g_{00} &= -1 + \frac{2}{c^2}\left[w(t,\boldsymbol{x}) + w_L(t,\boldsymbol{x})\right] - \frac{2}{c^4}\left[w^2(t,\boldsymbol{x}) + \Delta(t,\boldsymbol{x})\right] , \\
g_{0i} &= -\frac{4}{c^3} w^i(t,\boldsymbol{x}) , \\
g_{ij} &= \left(1 + \frac{2w(t,\boldsymbol{x})}{c^2}\right)\delta_{ij} ,
\end{aligned}
$$

where ($t \equiv$ Barycentric Coordinate Time (TCB), $\boldsymbol{x}$) are the barycentric coordinates, $w = G\sum_{\mathrm{A}} M_{\mathrm{A}}/r_{\mathrm{A}}$ with the summation carried out over all solar system bodies A, $\boldsymbol{r}_{\mathrm{A}} = \boldsymbol{x} - \boldsymbol{x}_{\mathrm{A}}, \boldsymbol{x}_{\mathrm{A}}$ are the coordinates of the center of mass of body A, $r_{\mathrm{A}} = |\boldsymbol{r}_{\mathrm{A}}|$, and where w_L contains the expansion in terms of multipole moments [see their definition in the Resolution B1.4 entitled “Post-Newtonian Potential Coefficients"] required for each body. The vector potential $w^i(t,\boldsymbol{x}) = \sum_{\mathrm{A}} w^i_{\mathrm{A}}(t,\boldsymbol{x})$, and the function $\Delta(t,\boldsymbol{x}) = \sum_{\mathrm{A}} \Delta_{\mathrm{A}}(t,\boldsymbol{x})$ are given in note 2.

2. the relation between TCB and Geocentric Coordinate Time (TCG) can be expressed to sufficient accuracy by

$$
\begin{aligned}
\mathrm{TCB} - \mathrm{TCG} &= \frac{1}{c^2}\left\{\int_{t_0}^{t}\left[\frac{v_{\mathrm{E}}^2}{2} + w_{\mathrm{ext}}(\boldsymbol{x}_{\mathrm{E}})\right]dt + v^i_{\mathrm{E}} r^i_{\mathrm{E}}\right\} \\
&+ \frac{1}{c^4}\int_{t_0}^{t}\left[\frac{1}{8}v_{\mathrm{E}}^4 + \frac{3}{2}v_{\mathrm{E}}^2 w_{\mathrm{ext}}(\boldsymbol{x}_{\mathrm{E}}) - 4v^i_{\mathrm{E}} w^i_{\mathrm{ext}}(\boldsymbol{x}_{\mathrm{E}}) - \frac{1}{2}w^2_{\mathrm{ext}}(\boldsymbol{x}_{\mathrm{E}})\right]dt \\
&+ \frac{1}{c^4}\left[3w_{\mathrm{ext}}(\boldsymbol{x}_{\mathrm{E}}) + \frac{v_{\mathrm{E}}^2}{2}\right]v^i_{\mathrm{E}} r^i_{\mathrm{E}} ,
\end{aligned}
$$

where v_{E} is the barycentric velocity of the Earth and where the index “ext" refers to summation over all bodies except the Earth.

Notes

1. This formulation will provide an uncertainty not larger than 5×10^{-18} in rate and, for quasi-periodic terms, not larger than 5×10^{-18} in rate amplitude and 0.2 ps in phase amplitude, for locations farther than a few solar radii from the Sun. The same

uncertainty also applies to the transformation between TCB and TCG for locations within 50000 km of the Earth. Uncertainties in the values of astronomical quantities may induce larger errors in the formulas.

2. Within the above mentioned uncertainties, it is sufficient to express the vector potential $w^i_{\mathrm{A}}(t, \boldsymbol{x})$ of body A as

$$w^i_{\mathrm{A}}(t, \boldsymbol{x}) = G\left[\frac{M_{\mathrm{A}} v^i_{\mathrm{A}}}{r_{\mathrm{A}}} - \frac{(\boldsymbol{r}_{\mathrm{A}} \times \boldsymbol{S}_{\mathrm{A}})^i}{2r^3_{\mathrm{A}}}\right] ,$$

where $\boldsymbol{S}_{\mathrm{A}}$ is the total angular momentum of body A and v^i_{A} is the barycentric coordinate velocity of body A. As for the function $\Delta_{\mathrm{A}}(t, \boldsymbol{x})$ it is sufficient to express it as

$$\Delta_{\mathrm{A}}(t, \boldsymbol{x}) = \frac{GM_{\mathrm{A}}}{r_{\mathrm{A}}}\left[-2v^2_a + \sum_{B \neq A} \frac{GM_B}{r_{BA}} + \frac{1}{2}\left(\frac{(r^k_{\mathrm{A}} v^k_{\mathrm{A}})^2}{r^2_{\mathrm{A}}} + r^k_{\mathrm{A}} a^k_{\mathrm{A}}\right)\right] + \frac{2Gv^k_{\mathrm{A}}(\boldsymbol{r}_{\mathrm{A}} \times \boldsymbol{S}_{\mathrm{A}})^k}{r^3_{\mathrm{A}}} ,$$

where $r_{BA} = |\boldsymbol{x}_B - \boldsymbol{x}_{\mathrm{A}}|$ and a^k_{A} is the barycentric coordinate acceleration of body A. In these formulas, the terms in $\boldsymbol{S}_{\mathrm{A}}$ are needed only for Jupiter ($|\boldsymbol{S}_{\mathrm{J}}| \approx 6.9 \times 10^{38}\mathrm{m}^2\mathrm{s}^{-1}\mathrm{kg}$) and Saturn ($|\boldsymbol{S}_{\mathrm{S}}| \approx 1.4 \times 10^{38}\mathrm{m}^2\mathrm{s}^{-1}\mathrm{kg}$), in the immediate vicinity of these planets.

3. Because the present recommendation provides an extension of the IAU 1991 recommendations valid at the full first post-Newtonian level, the constants L_{C} and L_{B} that were introduced in the IAU 1991 recommendations should be defined as <TCG/TCB>= $1 - L_{\mathrm{C}}$ and <TT/TCB>= $1 - L_{\mathrm{B}}B$, where TT refers to Terrestrial Time and < ... > refers to a sufficiently long average taken at the geocenter. The most recent estimate of L_{C} is (Irwin, A. and Fukushima, T., *Astron. Astroph.*, **348**, 642-652, 1999)

$$L_{\mathrm{C}} = 1.48082686741 \times 10^{-8} \pm 2 \times 10^{-17} .$$

From Resolution B1.9 on "Redefinition of Terrestrial Time TT", one infers $L_{\mathrm{B}} = 1.55051976772 \times 10^{-8} \pm 2 \times 10^{-17}$ by using the relation $1 - L_{\mathrm{B}} = (1 - L_{\mathrm{C}})(1 - L_{\mathrm{G}})$. L_{G} is defined in Resolution B1.9.

Because no unambiguous definition may be provided for L_{B} and L_{C}, these constants should not be used in formulating time transformations when it would require knowing their value with an uncertainty of order 1×10^{-16} or less.

4. If TCB−TCG is computed using planetary ephemerides which are expressed in terms of a time argument (noted $\mathrm{T_{eph}}$) which is close to Barycentric Dynamical Time

(TDB), rather than in terms of TCB, the first integral in Recommendation 2 above may be computed as

$$\int_{t_0}^{t}\left(\frac{v_{\mathrm{E}}^2}{2}+w_{\mathrm{ext}}(\boldsymbol{x}_{\mathrm{E}})\right)dt=\frac{1}{1-L_{\mathrm{B}}}\int_{T_{\mathrm{eph}_0}}^{T_{\mathrm{eph}}}\left(\frac{v_{\mathrm{E}}^2}{2}+w_{\mathrm{ext}}(\boldsymbol{x}_{\mathrm{E}})\right)dt\,.$$

Resolution B1.6 IAU 2000 Precession-Nutation Model

The XXIVth International Astronomical Union

Recognizing

1. that the International Astronomical Union and the International Union of Geodesy and Geophysics Working Group (IAU-IUGG WG) on 'Non-rigid Earth Nutation Theory' has met its goals by

a. establishing new high precision rigid Earth nutation series, such as (1) SMART97 of Bretagnon *et al.*, 1998, *Astron. Astroph.*, **329**, 329-338; (2) REN2000 of Souchay *et al.*, 1999, *Astron. Astroph. Supl. Ser.*, **135**, 111-131; (3) RDAN97 of Roosbeek and Dehant 1999, *Celest. Mech.*, **70**, 215-253;
b. completing the comparison of new non-rigid Earth transfer functions for an Earth initially in non-hydrostatic equilibrium, incorporating mantle anelasticity and a Free Core Nutation period in agreement with observations,
c. noting that numerical integration models are not yet ready to incorporate dissipation in the core,
d. noting the effects of other geophysical and astronomical phenomena that must be modelled, such as ocean and atmospheric tides, that need further development;

2. that, as instructed by IAU Recommendation C1 in 1994, the International Earth Rotation Service (IERS) will publish in the IERS Conventions (2000) a precession-nutation model that matches the observations with a weighted rms of 0.2 milliarcsecond (mas);

3. that semi-analytical geophysical theories of forced nutation are available which incorporate some or all of the following – anelasticity and electromagnetic couplings at the core-mantle and inner core-outer core boundaries, annual atmospheric tide, geodesic nutation, and ocean tide effects;

4. that ocean tide corrections are necessary at all nutation frequencies; and

5. that empirical models based on a resonance formula without further corrections do also exist;

Accepts

the conclusions of the IAU-IUGG WG on Non-rigid Earth Nutation Theory published by Dehant *et al.*, 1999, *Celest. Mech.* **72(4)**, 245-310 and the recent comparisons between the various possibilities, and

Recommends

that, beginning on 1 January 2003, the IAU 1976 Precession Model and IAU 1980 Theory of Nutation, be replaced by the precession-nutation model IAU 2000A (MHB2000, based on the transfer functions of Mathews, Herring and Buffett, 2000 – submitted to the *Journal of Geophysical Research*) for those who need a model at the 0.2 mas level, or its shorter version IAU 2000B for those who need a model only at the 1 mas level, together with their associated precession and obliquity rates, and their associated celestial pole offsets at J2000.0, to be published in the IERS Conventions 2000, and

Encourages

1. the continuation of theoretical developments of non-rigid Earth nutation series,

2. the continuation of VLBI observations to increase the accuracy of the nutation series and the nutation model, and to monitor the unpredictable free core nutation, and

3. the development of new expressions for precession consistent with the IAU 2000A model.

Resolution B1.7 Definition of Celestial Intermediate Pole

The XXIVth International Astronomical Union

Noting

the need for accurate definition of reference systems brought about by unprecedented observational precision, and

Recognizing

1. the need to specify an axis with respect to which the Earth's angle of rotation is defined,

2. that the Celestial Ephemeris Pole (CEP) does not take account of diurnal and higher frequency variations in the Earth's orientation,

Recommends

1. that the Celestial Intermediate Pole (CIP) be the pole, the motion of which is

specified in the Geocentric Celestial Reference System (GCRS, see Resolution B1.3) by motion of the Tisserand mean axis of the Earth with periods greater than two days,

2. that the direction of the CIP at J2000.0 be offset from the direction of the pole of the GCRS in a manner consistent with the IAU 2000A (see Resolution B1.6) precession-nutation model,

3. that the motion of the CIP in the GCRS be realized by the IAU 2000A model for precession and forced nutation for periods greater than two days plus additional time-dependent corrections provided by the International Earth Rotation Service (IERS) through appropriate astro-geodetic observations,

4. that the motion of the CIP in the International Terrestrial Reference System (ITRS) be provided by the IERS through appropriate astro-geodetic observations and models including high-frequency variations,

5. that for highest precision, corrections to the models for the motion of the CIP in the ITRS may be estimated using procedures specified by the IERS, and

6. that implementation of the CIP be on 1 January 2003.

Notes

1. The forced nutations with periods less than two days are included in the model for the motion of the CIP in the ITRS.

2. The Tisserand mean axis of the Earth corresponds to the mean surface geographic axis, quoted B axis, in Seidelmann, 1982, *Celest. Mech.*, **27**, 79-106.

3. As a consequence of this resolution, the Celestial Ephemeris Pole is no longer necessary.

Resolution B1.8 Definition and use of Celestial and Terrestrial Ephemeris Origin

The XXIVth International Astronomical Union

Recognizing

1. the need for reference system definitions suitable for modern realizations of the conventional reference systems and consistent with observational precision,

2. the need for a rigorous definition of sidereal rotation of the Earth,

3. the desirability of describing the rotation of the Earth independently from its orbital motion, and

Noting

that the use of the "non-rotating origin" (Guinot, 1979) on the moving equator fulfills the above conditions and allows for a definition of UT1 which is insensitive to changes in models for precession and nutation at the microarcsecond level,

Recommends

1. the use of the "non-rotating origin" in the Geocentric Celestial Reference System (GCRS) and that this point be designated as the Celestial Ephemeris Origin (CEO) on the equator of the Celestial Intermediate Pole (CIP),

2. the use of the "non-rotating origin" in the International Terrestrial Reference System (ITRS) and that this point be designated as the Terrestrial Ephemeris Origin (TEO) on the equator of the CIP,

3. that UT1 be linearly proportional to the Earth Rotation Angle defined as the angle measured along the equator of the CIP between the unit vectors directed toward the CEO and the TEO,

4. that the transformation between the ITRS and GCRS be specified by the position of the CIP in the GCRS, the position of the CIP in the ITRS, and the Earth Rotation Angle,

5. that the International Earth Rotation Service (IERS) take steps to implement this by 1 January 2003, and

6. that the IERS will continue to provide users with data and algorithms for the conventional transformations.

Note

1. The position of the CEO can be computed from the IAU 2000A model for precession and nutation of the CIP and from the current values of the offset of the CIP from the pole of the ICRF at J2000.0 using the development provided by Capitaine *et al.* (2000).

2. The position of the TEO is only slightly dependent on polar motion and can be extrapolated as done by Capitaine *et al.* (2000) using the IERS data.

3. The linear relationship between the Earth's rotation angle θ and UT1 should ensure the continuity in phase and rate of UT1 with the value obtained by the conventional relationship between Greenwich Mean Sidereal Time (GMST) and UT1. This is accomplished by the following relationship:

$\theta(\mathrm{UT1})=2\pi(0.7790572732640+1.00273781191135448\times(\text{Julian UT1 date} - 2451545.0))$.

References:

Guinot, B., 1979, in D.D. McCarthy and J.D. Pilkington (eds.), *Time and the Earth's*

Rotation, D. Reidel Publ., 7-18.

Capitaine, N., Guinot, B., McCarthy, D.D., 2000, “Definition of the Celestial Ephemeris Origin and of UT1 in the International Celestial Reference Frame", *Astron. Astrophys.*, **355**, 398-405.

Resolution B1.9 Re-definition of Terrestrial Time TT

The XXIVth International Astronomical Union

Considering

1. that IAU Resolution A4 (1991) has defined Terrestrial Time (TT) in its Recommendation 4, and

2. that the intricacy and temporal changes inherent to the definition and realization of the geoid are a source of uncertainty in the definition and realization of TT, which may become, in the near future, the dominant source of uncertainty in realizing TT from atomic clocks,

Recommends

that TT be a time scale differing from TCG by a constant rate: $\mathrm{dTT/dTCG} = 1 - L_\mathrm{G}$, where $L_\mathrm{G} = 6.969290134 \times 10^{-10}$ is a defining constant,

Note

L_G was defined by the IAU Resolution A4 (1991) in its Recommendation 4 as equal to U_G/c^2 where U_G is the geopotential at the geoid. L_G is now used as a defining constant.

Resolution B2 Coordinated Universal Time

The XXIVth International Astronomical Union

Recognizing

1. that the definition of Coordinated Universal Time (UTC) relies on the astronomical observation of the UT1 time scale in order to introduce leap seconds,

2. that the unpredictable leap seconds affects modern communication and navigation systems,

3. that astronomical observations provide an accurate estimate of the secular deceleration of the Earth’s rate of rotation

Recommends

1. that the IAU establish a working group reporting to Division I at the General Assembly in 2003 to consider the redefinition of UTC,

2. that this study discuss whether there is a requirement for leap seconds, the possibility of inserting leap seconds at pre-determined intervals, and the tolerance limits for UT1−UTC, and

3. that this study be undertaken in cooperation with the appropriate groups of the International Union of Radio Science (URSI), the International Telecommunications Union (ITU-R), the International Bureau for Weights and Measures (BIPM), the International Earth Rotation Service (IERS) and relevant navigational agencies.

C.3

Text of IAU Resolutions of 2006 Adopted at the XXVI-th General Assembly, Prague

Resolution B1 Adoption of the P03 Precession Theory and Definition of the Ecliptic

The XXVIth International Astronomical Union General Assembly,

Noting

1. the need for a precession theory consistent with dynamical theory,

2. that, while the precession portion of the IAU 2000A precession-nutation model, recom- mended for use beginning on 1 January 2003 by resolution B1.6 of the XXIVth IAU General Assembly, is based on improved precession rates with respect to the IAU 1976 precession, it is not consistent with dynamical theory, and

3. that resolution B1.6 of the XXIVth General Assembly also encourages the development of new expressions for precession consistent with the IAU 2000A precession-nutation model, and

Recognizing

1. that the gravitational attraction of the planets make a significant contribution to the motion of the Earth's equator, making the terms *lunisolar precession* and *planetary precession* misleading,

2. the need for a definition of the ecliptic for both astronomical and civil purposes, and

3. that in the past, the ecliptic has been defined both with respect to an observer situated in inertial space (inertial definition) and an observer comoving with the ecliptic

(rotating definition),

Accepts

the conclusions of the IAU Division I Working Group on Precession and the Ecliptic published in Hilton et al. (2006, Celest. Mech. 94, 351), and

Recommends

1. that the terms *lunisolar precession* and *planetary precession* be replaced by *precession of the equator* and *precession of the ecliptic*, respectively,

2. that, beginning on 1 January 2009, the precession component of the IAU 2000A precession-nutation model be replaced by the P03 precession theory, of Capitaine et al. (2003, A&A, 412, 567-586) for the precession of the equator (Eqs. 37) and the precession of the ecliptic (Eqs. 38); the same paper provides the polynomial developments for the P03 primary angles and a number of derived quantities for use in both the equinox based and CIO based paradigms,

3. that the choice of precession parameters be left to the user, and

4. that the ecliptic pole should be explicitly defined by the mean orbital angular momentum vector of the Earth-Moon barycenter in the Barycentric Celestial Reference System (BCRS), and this definition should be explicitly stated to avoid confusion with other, older definitions.

Notes

1. Formulas for constructing the precession matrix using various parameterizations are given in Eqs. 1, 6, 7, 11, 12 and 22 of Hilton et al. (2006). The recommended polynomial developments for the various parameters are given in Table 1 of the same paper, including the P03 expressions set out in expressions (37) to (41) of Capitaine et al. (2003) and Tables 3-5 of Capitaine et al. (2005).

2. The time rate of change in the dynamical form factor in P03 is
$dJ_2/dt = -3.001 \times 10^{-9}\text{century}^{-1}$.

References

Capitaine, N., Wallace, P.T., & Chapront, J. 2003, A&A, 412, 567
Capitaine, N., Wallace, P.T., & Chapront, J. 2005, A&A, 432, 355
Hilton, J.L., Capitaine, N., Chapront, J., Ferrandiz, J.M., Fienga, A., Fukushima, T., Getino, J., Mathews, P., Simon, J.-L., Soffel, M., Vondrak, J., Wallace, P., & Williams, J. 2006, Celest. Mech., 94, 351

Actions to be taken by the General Secretary upon adoption of the Resolution *Adoption of the P03 Precession Theory and Definition of the Ecliptic.*
The following institutions should receive formal notification of the action:

Her Majesty's Nautical Almanac Office, Institute de mécanique céleste et de calcul des éphémérides, Institute of Applied Astronomy of the Russian Academy of Sciences, International Association of Geodesy, (IAG), International Earth Rotation and Reference Systems Service (IERS), International Union of Geodesy and Geophysics (IUGG), International VLBI Service for Geodesy and Astrometry (IVS), Japan Coast Guard (JCG), National Astronomical Observatory of Japan (NAOJ), Nautical Almanac Office of the United States Naval Observatory.

Resolution B2 Supplement to the IAU 2000 Resolutions on reference systems

RECOMMENDATION 1.*Harmonizing the name of the pole and origin to "intermediate"*

The XXVIth International Astronomical Union General Assembly,

Noting

1. the adoption of resolutions IAU B1.1 through B1.9 by the IAU General Assembly of 2000,

2. that the International Earth Rotation and Reference Systems Service (IERS) and the Standards Of Fundamental Astronomy (SOFA) activity have made available the models, procedures, data and software to implement these resolutions operationally, and that the Almanac Offices have begun to implement them beginning with their 2006 editions, and

3. the recommendations of the IAU Working Group on "Nomenclature for Fundamental Astronomy" (IAU Transactions XXVIA, 2005), and

Recognizing

1. that using the designation "intermediate" to refer to both the pole and the origin of the new systems linked to the Celestial Intermediate Pole and the Celestial or Terrestrial Ephemeris origins, defined in Resolutions B1.7 and B1.8, respectively would improve the consistency of the nomenclature, and

2. that the name "Conventional International Origin" with the potentially conflicting acronym CIO is no longer commonly used to refer to the reference pole for measuring polar motion as it was in the past by the International Latitude Service,

Recommends

1. that, the designation "intermediate" be used to describe the moving celestial and terrestrial reference systems defined in the 2000 IAU Resolutions and the various related entities, and

2. that the terminology "Celestial Intermediate Origin" (CIO) and "Terrestrial In-

termediate Origin" (TIO) be used in place of the previously introduced "Celestial Ephemeris Origin" (CEO) and "Terrestrial Ephemeris Origin" (TEO), and

3. that authors carefully define acronyms used to designate entities of astronomical reference systems to avoid possible confusion.

RECOMMENDATION 2. *Default orientation of the Barycentric Celestial Reference System (BCRS) and Geocentric Celestial Reference System (GCRS)*

The XXVIth International Astronomical Union General Assembly,

Noting

1. the adoption of resolutions IAU B1.1 through B1.9 by the IAU General Assembly of 2000,

2. that the International Earth Rotation and Reference Systems Service (IERS) and the Standards Of Fundamental Astronomy (SOFA) activity have made available the models, procedures, data and software to implement these resolutions operationally, and that the Almanac Offices have begun to implement them beginning with their 2006 editions,

3. that, in particular, the systems of spacetime coordinates defined by IAU 2000 Resolution B1.3 for (a) the solar system (called the Barycentric Celestial Reference System, BCRS) and (b) the Earth (called the Geocentric Celestial Reference System, GCRS) have begun to come into use,

4. the recommendations of the IAU Working Group on Nomenclature for Fundamental Astronomy (IAU Transactions XXVIA, 2005), and

5. a recommendation from the IAU Working Group on Relativity in Celestial Mechanics, Astrometry and Metrology,

Recognizing

1. that the BCRS definition does not determine the orientation of the spatial coordinates,

2. that the natural choice of orientation for typical applications is that of the ICRS, and

3. that the GCRS is defined such that its spatial coordinates are kinematically nonrotating with respect to those of the BCRS,

Recommends

that the BCRS definition is completed with the following: "For all practical applications, unless otherwise stated, the BCRS is assumed to be oriented according to the ICRS axes. The orientation of the GCRS is derived from the ICRS-oriented BCRS."

Resolution B3 Re-definition of Barycentric Dynamical Time, TDB

The XXVIth International Astronomical Union General Assembly,

Noting

1. that IAU Recommendation 5 of Commissions 4, 8 and 31 (1976) introduced, as a replacement for Ephemeris Time (ET), a family of dynamical time scales for barycentric ephemerides and a unique time scale for apparent geocentric ephemerides,

2. that IAU Resolution 5 of Commissions 4, 19 and 31 (1979) designated these time scales as Barycentric Dynamical Time (TDB) and Terrestrial Dynamical Time (TDT) respectively, the latter subsequently renamed Terrestrial Time (TT), in IAU Resolution A4, 1991,

3. that the difference between TDB and TDT was stipulated to comprise only periodic terms, and

4. that Recommendations III and V of IAU Resolution A4 (1991) (i) introduced the coordinate time scale Barycentric Coordinate Time (TCB) to supersede TDB, (ii) recognized that TDB was a linear transformation of TCB, and (iii) acknowledged that, where discontinuity with previous work was deemed to be undesirable, TDB could be used, and

Recognizing

1. that TCB is the coordinate time scale for use in the Barycentric Celestial Reference System,

2. the possibility of multiple realizations of TDB as defined currently,

3. the practical utility of an unambiguously defined coordinate time scale that has a linear relationship with TCB chosen so that at the geocenter the difference between this coordinate time scale and Terrestrial Time (TT) remains small for an extended time span,

4. the desirability for consistency with the Teph time scales used in the Jet Propulsion Laboratory (JPL) solar-system ephemerides and existing TDB implementations such as that of Fairhead & Bretagnon (A&A 229, 240, 1990), and

5. the 2006 recommendations of the IAU Working Group on "Nomenclature for Fundamental Astronomy" (IAU Transactions XXVIB, 2006),

Recommends

that, in situations calling for the use of a coordinate time scale that is linearly related to Barycentric Coordinate Time (TCB) and, at the geocenter, remains close to Terrestrial Time (TT) for an extended time span, TDB be defined as the following linear transformation of TCB:

$$\mathrm{TDB} = \mathrm{TCB} - L_\mathrm{B} \times (\mathrm{JD}_\mathrm{TCB} - T_0) \times 86400 + \mathrm{TDB}_0 \,,$$

where $T_0 = 2443144.5003725$, and $L_\mathrm{B} = 1.550519768 \times 10^{-8}$ and $\mathrm{TDB}_0 = -6.55 \times 10^{-5}$ s are defining constants.

Notes

1. JD_TCB is the TCB Julian date. Its value is $T_0 = 2443144.5003725$ for the event 1977 January 1 00h 00m 00s TAI at the geocenter, and it increases by one for each 86400s of TCB.

2. The fixed value that this definition assigns to L_B is a current estimate of $L_\mathrm{C} + L_\mathrm{G} - L_\mathrm{C} \times L_\mathrm{G}$, where L_G is given in IAU Resolution B1.9 (2000) and L_C has been determined (Irwin & Fukushima, 1999, A&A 348, 642) using the JPL ephemeris DE405. When using the JPL Planetary Ephemeris DE405, the defining L_B value effectively eliminates a linear drift between TDB and TT, evaluated at the geocenter. When realizing TCB using other ephemerides, the difference between TDB and TT, evaluated at the geocenter, may include some linear drift, not expected to exceed 1 ns per year.

3. The difference between TDB and TT, evaluated at the surface of the Earth, remains under 2 ms for several millennia around the present epoch.

4. The independent time argument of the JPL ephemeris DE405, which is called T_eph (Standish, A&A, 336, 381, 1998), is for practical purposes the same as TDB defined in this Resolution.

5. The constant term TDB_0 is chosen to provide reasonable consistency with the widely used TDB - TT formula of Fairhead & Bretagnon (1990).
n.b. The presence of TDB_0 means that TDB is not synchronized with TT, TCG and TCB at 1977 Jan 1.0 TAI at the geocenter.

6. For solar system ephemerides development the use of TCB is encouraged.

C.4

Text of IAU Resolutions of 2009 Adopted at the XXVII-th General Assembly, Rio de Janeiro

Resolution B2 IAU 2009 Astronomical Constants

The XXVII General Assembly of International Astronomical Union,

Considering

1. the need for a self-consistent set of accurate numerical standards for use in astronomy,

2. that improved values of astronomical constants have been derived from recent observations and published in refereed journals, and

3. that conventional values have been adopted by IAU GA 2000 and IAU GA 2006 resolutions for a number of astronomical quantities,

Recognizing

1. the continuing need for a set of Current Best Estimates (CBEs) of astronomical numerical constants, and

2. the need for an operational service to the astronomical community to maintain the CBEs

Recommends

1. that the list of previously published constants compiled in the report of the Working Group on Numerical Standards of Fundamental Astronomy (see http://maia.usno.navy.mil/NSFA/CBE.html) be adopted as the IAU (2009) System of Astronomical Constants.

2. that Current Best Estimates of Astronomical Constants be permanently maintained as an electronic document,

3. that, in order to ensure the integrity of the CBEs, IAU Division I develop a formal procedure to adopt new values and archive older versions of the CBEs, and

4. that the IAU establish within IAU Division I a permanent body to maintain the CBEs for fundamental astronomy.

Resolution B3 Second Realization of the International Celestial Reference Frame

The International Astronomical Union XXVII General Assembly,

noting

1. that Resolution B2 of the XXIII General Assembly (1997) resolved "That, as from 1 January 1998, the IAU celestial reference system shall be the International Celestial Reference System (ICRS)",

2. that Resolution B2 of the XXIII General Assembly (1997) resolved that the "fundamental reference frame shall be the International Celestial Reference Frame (ICRF) constructed by the IAU Working Group on Reference Frames",

3. that Resolution B2 of the XXIII General Assembly (1997) resolved "That IERS should take appropriate measures, in conjunction with the IAU Working Group on reference frames, to maintain the ICRF and its ties to the reference frames at other wavelengths",

4. that Resolution B7 of the XXIII General Assembly (1997) recommended "that high-precision astronomical observing programs be organized in such a way that astronomical reference systems can be maintained at the highest possible accuracy for both northern and southern hemispheres",

5. that Resolution B1.1 of the XXIV General Assembly (2000) recognized "the importance of continuing operational observations made with Very Long Baseline Interferometry (VLBI) to maintain the ICRF",

recognizing

1. that since the establishment of the ICRF, continued VLBI observations of ICRF sources have more than tripled the number of source observations,

2. that since the establishment of the ICRF, continued VLBI observations of extragalactic sources have significantly increased the number of sources whose positions are known with a high degree of accuracy,

3. that since the establishment of the ICRF, improved instrumentation, observation strategies, and application of state-of-the-art astrophysical and geophysical models have significantly improved both the data quality and analysis of the entire relevant astrometric and geodetic VLBI data set.,

4. that a working group on the ICRF formed by the International Earth Rotation and Reference Systems Service (IERS) and the International VLBI Service for Geodesy and Astrometry (IVS), in conjunction with the IAU Division I Working Group on the Second Realization of the International Celestial Reference Frame has finalized a prospective second realization of the ICRF in a coordinate frame aligned to that of the ICRF to within the tolerance of the errors in the latter (see note 1),

5. that the prospective second realization of the ICRF as presented by the IAU Work-

ing Group on the Second Realization of the International Celestial Reference Frame represents a significant improvement in terms of source selection, coordinate accuracy, and total number of sources, and thus represents a significant improvement in the fundamental reference frame realization of the ICRS beyond the ICRF adopted by the XXIII General Assembly (1997),

resolves

1. that from 01 January 2010 the fundamental astrometric realization of the International Celestial Reference System (ICRS) shall be the Second Realization of the International Celestial Reference Frame (ICRF2) as constructed by the IERS/IVS working group on the ICRF in conjunction with the IAU Division I Working Group on the Second Realization of the International Celestial Reference Frame (see note 1),

2. that the organizations responsible for astrometric and geodetic VLBI observing programs (e.g. IERS, IVS) take appropriate measures to continue existing and develop improved VLBI observing and analysis programs to both maintain and improve ICRF2,

3. that the IERS, together with other relevant organizations continue efforts to improve and densify high accuracy reference frames defined at other wavelengths and continue to improve ties between these reference frames and ICRF2.

Note 1: The Second Realization of the International Celestial Reference Frame by Very Long Baseline Interferometry, Presented on behalf of the IERS / IVS Working Group, Alan Fey and David Gordon (eds.). (IERS Technical Note 35) Frankfurt am Main: Verlag des Bundesamts für Kartographie und Geodäsie, 2009. See http://www.iers.org/nn_11216/IERS/EN/Publications/TechnicalNotes/tn35.html or http://hpiers.obspm.fr/icrs-pc/.

Bibliography

Abbott, B.P., Abbott, R., Acernese, F., Adhikari, R., Ajith, P., Allen, B., Allen, G., Alshourbagy, M., Amin, R.S., Anderson, S.B., Anderson, W.G., Antonucci, F., Aoudia, S., Arain, M.A., Araya, M., Armandula, H., Armor, P., Arun, K.G., Aso, Y., Aston, S., Astone, P., Aufmuth, P., Aulbert, C., Babak, S., Baker, P., Ballardin, G., Ballmer, S., Barker, C., Barker, D., Barone, F., Barr, B., Barriga, P., Barsotti, L., Barsuglia, M., Barton, M.A., Bartos, I., Bassiri, R., Bastarrika, M., Bauer, T.S., Behnke, B., Beker, M., Benacquista, M., Betzwieser, J., Beyersdorf, P.T., Bigotta, S., Bilenko, I.A., Billingsley, G., Birindelli, S., Biswas, R., Bizouard, M.A., Black, E., Blackburn, J.K., Blackburn, L., Blair, D., Bland, B., Boccara, C., Bodiya, T.P., Bogue, L., Bondu, F., Bonelli, L., and 619 additional authors not shown (2010) Searches for Gravitational Waves from Known Pulsars with Science Run 5 LIGO Data. *Astrophysical Journal*, **713**, 671–685, doi:10.1088/0004-637X/713/1/671. 3.10

Abdullah, K. and Albouy, A. (2001) On a strange resonance noticed by M. Herman. *Regul. Chaotic Dyn.*, **6** (4), 421–432. 1.5.6

Abiko, S. (2003) On Einstein's distrust of the electromagnetic theory: The origin of the lightvelocity postulate. *Historical Studies in the Physical and Biological Sciences*, **33**, 193–215. URL http://www.jstor.org/stable/pdfplus/10.1525/hsps.2003.33.2.193.pdf. 2.2.2

Adler, C.G. (1987) Does mass really depend on velocity, dad? *American Journal of Physics*, **55**, 739–743, doi:10.1119/1.15314. 2.7.1

Aghanim, N., Prunet, S., Forni, O., and Bouchet, F.R. (1998) Moving gravitational lenses: imprints on the cosmic microwave background. *Astronomy & Astrophysics*, **334**, 409–419. 7.6.2.2

Aguiar, O.D., Andrade, L.A., Barroso, J.J., Castro, P.J., Costa, C.A., de Souza, S.T., de Waard, A., Fauth, A.C., Frajuca, C., Frossati, G., Furtado, S.R., Gratens, X., Maffei, T.M.A., Magalhães, N.S., Marinho, Jr., R.M., Oliveira, Jr., N.F., Pimentel, G.L., Remy, M.A., Tobar, M.E., Abdalla, E., Alves, M.E.S., Bessada, D.F.A., Bortoli, F.S., Brandão, C.S.S., Costa, K.M.F., de Araújo, H.A.B., de Araujo, J.C.N., de Gouveia Dal Pino, E.M., de Paula, W., de Rey Neto, E.C., Evangelista, E.F.D., Lenzi, C.H., Marranghello, G.F., Miranda, O.D., Oliveira, S.R., Opher, R., Pereira, E.S., Stellati, C., and Weber, J. (2008) The Schenberg spherical gravitational wave detector: the first commissioning runs. *Classical and Quantum Gravity*, **25** (11), 114 042–+, doi:10.1088/0264-9381/25/11/114042. 3.10.5

Aharonov, Y. and Bohm, D. (1959) Significance of electromagnetic potentials in the quantum theory. *Physical Review D*, **115** (3), 485–491, doi:10.1103/PhysRev.115.485. 57

Aharonov, Y. and Bohm, D. (1961) Further considerations on electromagnetic potentials in the quantum theory. *Physical Review D*, **123** (4), 1511–1524, doi:10.1103/PhysRev.123.1511. 57

Alba, D. and Lusanna, L. (2007) Generalized Radar 4-COORDINATES and Equal-Time Cauchy Surfaces for Arbitrary Accelerated Observers. *International Journal of Modern Physics D*, **16**, 1149–1186, doi:10.1142/S021827180701064X. 2.6.4

Albertella, A., Migliaccio, F., and Sansó, F. (2002) GOCE: The Earth Gravity Field by Space Gradiometry. *Celestial Mechanics and Dynamical Astronomy*, **83**, 1–15. 8.6, 8.7

Aldrovandi, R. and Pereira, J.G. (2007) Gravitation: in search of the missing torsion. *Ann. Fond. Louis de Broglie*, **32** (3), 229–252. URL http://www.ensmp.fr/aflb/AFLB-322/aflb322m605.htm. 3.7.4

Alexander, M.E. (1973) The Weak Friction Approximation and Tidal Evolution in Close Binary Systems. *Astrophysics & Space Sciences*, **23**, 459–510, doi:10.1007/BF00645172. 6.3.1, 6.3.1

Altamimi, Z. and Collilieux, X. (2008) IGS contribution to the ITRF. *Journal of Geodesy*, **83**, 375–383, doi:10.1007/s00190-008-0294-x. 8.1

Altschul, B., Bailey, Q.G., and Kostelecký, V.A. (2010) Lorentz violation with an antisymmetric tensor. *Physical Review D*, **81** (6), 065 028–19, doi:10.1103/PhysRevD.81.065028. 3.8.7

Anderson, J.D., Lau, E.L., and Giampieri, G. (2004) Measurement of the PPN Parameter gamma with Radio Signals from the Cassini Spacecraft at X-bands and Ka-Bands, in *Proceedings of the 22nd Texas Symposium on Relativistic Astrophysics at Stanford University, Stanford, California, 13-17 Dec 2004*. URL http://www.slac.stanford.edu/econf/C041213/papers/0305.PDF. 7.7.1, 7.8

Anderson, J.D. and Nieto, M.M. (2010) Astrometric solar-system anomalies, in *IAU Symposium*, *IAU Symposium*, vol. 261 (ed. S. A. Klioner, P. K. Seidelmann, & M. H. Soffel), *IAU Symposium*, vol. 261, pp. 189–197, doi:10.1017/S1743921309990378. 1.6

Anderson, J.L. (1967) *Principles of Relativity Physics*, New York: Academic Press; 484 pp. 3.8.1

Anderson, J.L. and Decanio, T.C. (1975) Equations of hydrodynamics in general relativity in the slow motion approximation. *General Relativity and Gravitation*, **6**, 197–237, doi:10.1007/BF00769986. 3.10.1, 4.1, 7.7.1

Anderson, J.L. and Gautreau, R. (1969) Operational Formulation of the Principle of Equivalence. *Phys. Rev.*, **185** (5), 1656–1661, doi:10.1103/PhysRev.185.1656. 3.8.1

Anderson, J.L., Kates, R.E., Kegeles, L.S., and Madonna, R.G. (1982) Divergent integrals of post-Newtonian gravity: Nonanalytic terms in the near-zone expansion of a gravitationally radiating system found by matching. *Physical Review D*, **25**, 2038–2048, doi:10.1103/PhysRevD.25.2038. 4.1, 5.2.1

Anderson, J.L. and Madonna, R. (1983) Matching Zerilli to Chandrasekhar-Esposito through the first post-Newtonian order. *General Relativity and Gravitation*, **15**, 1121–1129, doi:10.1007/BF00761578. 5.2.1

Andrei, A.H., Souchay, J., Zacharias, N., Smart, R.L., Vieira Martins, R., da Silva Neto, D.N., Camargo, J.I.B., Assafin, M., Barache, C., Bouquillon, S., Penna, J.L., and Taris, F. (2009) The Large Quasar Reference Frame (LQRF) - an optical representation of the ICRS. *ArXiv e-prints*. 9.4.8

Arakida, H. (2004) Gravitomagnetic translational motion on gravitational time delay, in *A fool in space. Proceedings of the 36th Symposium on Celestial Mechanics* (eds K. Tanikawa, H. Arakida, T. Yamamoto, and M. Saito), pp. 259–266. 7.9.6.2

Arfken, G.B. and Weber, H.J. (2001) *Mathematical methods for physicists*, Harcourt/Academic Press, Burlington, MA, 5th edn.. 3.10.1, 7.7.2

Armstrong, J., Bertotti, B., Estabrook, F., Iess, L., and Wahlquist, H. (1997) Low-Frequency Gravitational Wave Experiments, in *General Relativity and Gravitational Physics* (eds M. Bassan, V. Ferrari, M. Francaviglia, F. Fucito, and I. Modena), pp. 439–+. 7.8

Armstrong, J.W. (2006) Low-Frequency Gravitational Wave Searches Using Spacecraft Doppler Tracking. *Living Reviews in Relativity*, **9**, 1–+. 7.8

Armstrong, J.W., Estabrook, F.B., Asmar, S.W., Iess, L., and Tortora, P. (2008) Reducing antenna mechanical noise in precision spacecraft tracking. *Radio Science*, **43**, 3010–+, doi:10.1029/2007RS003766. 7.8

Armstrong, J.W., Iess, L., Tortora, P., and Bertotti, B. (2003) Stochastic Gravitational Wave Background: Upper Limits in the 10^{-6} to 10^{-3} Hz Band. *Astrophysical Journal*, **599**, 806–813, doi:10.1086/379505. 7.8

Arnold, V.I. (1995) *Mathematical Methods of Classical Mechanics*, Berlin: Springer. 1.4.7, 1.4.8, 1.5.3.1, 1.5.7, 2.1.1, 2.1.3, 3.6.3, 3.9.2, 3.9.2.1, 7.9.1, 7.9.3.2

Arnowitt, R.L., Deser, S., and Misner, C.W. (1962) *Canonical analysis of general relativity*, Oxford: Pergamon Press, pp. 127–136. 3.8.5, 3.10.5, 7.9.3.1

Arnowitt, R.L., Deser, S., and Misner, C.W. (2008) Republication of: The dynamics of general relativity. *General Relativity and Gravitation*, **40**, 1997–2027, doi:10.1007/s10714-008-0661-1. 3.8.5, 7.9.3.1

Ashby, N. (1998) Relativistic Effects in the Global Positioning System, in *Gravitation and Relativity: At the Turn of the Millennium* (eds N. Dadhich and J. Narlikar), pp. 231–258. 4.5.2.3

Ashby, N. (2002) Relativity and the Global Positioning System. *Physics Today*, **55** (5), 050 000–47. 7.8

Ashby, N. (2003) Relativity in the Global Positioning System. *Living Reviews in Relativity*, **6**, 1–+. 7.8

Ashby, N. and Allison, T. (2007) Relativistic canonical planetary perturbation equations for momentum-dependent forces, and the Lense-Thirring precession, in *New Trends in Astrodynamics and Applications III*, *American Institute of Physics Conference Series*, vol. 886 (ed. E. Belbruno), *American Institute of Physics Conference Series*, vol. 886, pp. 253–267, doi:10.1063/1.2710061. 1.5.4

Ashby, N. and Bertotti, B. (1984) Relativistic Perturbations of an Earth Satellite. *Physical Review Letters*, **52**, 485–488. 4.1, 5.1.1, 5.2.5.1, 6.1.1, 8.2

Ashby, N. and Bertotti, B. (1986) Relativistic effects in local inertial frames. *Physical Review D*, **34**, 2246–2259. 4.1, 5.1.1, 5.2.6, 6.1.1, 8.2

Astone, P., Babusci, D., Baggio, L., Bassan, M., Bignotto, M., Bonaldi, M., Camarda, M., Carelli, P., Cavallari, G., Cerdonio, M., Chincarini, A., Coccia, E., Conti, L., D'Antonio, S., de Rosa, M., di Paolo Emilio, M., Drago, M., Dubath, F., Fafone, V., Falferi, P., Foffa, S., Fortini, P., Frasca, S., Gemme, G., Giordano, G., Giusfredi, G., Hamilton, W.O., Hanson, J., Inguscio, M., Johnson, W.W., Liguori, N., Longo, S., Maggiore, M., Marin, F., Marini, A., McHugh, M.P., Mezzena, R., Miller, P., Minenkov, Y., Mion, A., Modestino, G., Moleti, A., Nettles, D., Ortolan, A., Pallottino, G.V., Parodi, R., Piano Mortari, G., Poggi, S., Prodi, G.A., Quintieri, L., Re, V., Rocchi, A., Ronga, F., Salemi, F., Soranzo, G., Sturani, R., Taffarello, L., Terenzi, R., Torrioli, G., Vaccarone, R., Vandoni, G., Vedovato, G., Vinante, A., Visco, M., Vitale, S., Weaver, J., Zendri, J.P., and Zhang, P. (2007) Results of the IGEC-2 search for gravitational wave bursts during 2005. *Physical Review D*, **76** (10), 102 001–+, doi:10.1103/PhysRevD.76.102001. 3.10.5

Babak, S. (2008) Building a stochastic template bank for detecting massive black hole binaries. *Classical and Quantum Gravity*, **25** (19), 195 011–+, doi:10.1088/0264-9381/25/19/195011. 3.10.5

Babak, S., Gair, J.R., and Porter, E.K. (2009) An algorithm for the detection of extreme mass ratio inspirals in LISA data. *Classical and Quantum Gravity*, **26** (13), 135 004–+, doi:10.1088/0264-9381/26/13/135004. 3.10.5

Babak, S.V. and Grishchuk, L.P. (2000) The Energy-Momentum Tensor for the Gravitational Field. *Physical Review D*, **61**, 024 038. 3.9.1, 3.9.7, 3.9.7

Babak, S.V. and Grishchuk, L.P. (2003) Finite-Range Gravity and its Role in Gravitational Waves, Black Holes and Cosmology. *International Journal of Modern Physics D*, **12**, 1905–1959, doi:10.1142/S0218271803004250. 7.9.4.3

Bailey, Q.G. and Kostelecky, A. (2008) SME Gravitational Tests. *APS Meeting Abstracts, Vol. 53, No 5*, p. T10.00002. URL http://meetings.aps.org/Meeting/APR08/Event/83905. 3.8.7

Bailey, Q.G. and Kostelecký, V.A. (2006) Signals for Lorentz violation in post-Newtonian gravity. *Physical Review D*, **74** (4), 045 001–+, doi:10.1103/PhysRevD.74.045001. 7.9.4.2

Bak, D., Cangemi, D., and Jackiw, R. (1994) Energy-momentum conservation in gravity theories. *Physical Review D*, **49**, 5173–5181, doi:10.1103/PhysRevD.49.5173. 3.9.4.3

Barbieri, A. and Guadagnini, E. (2005)

Gravitational helicity interaction. *Nuclear Physics B*, **719**, 53–66, doi:10.1016/j.nuclphysb.2005.04.023. 7.1

Barbour, J. and Pfister, H. (eds) (1995) *Mach's Principle - From Newton's Bucket to Quantum Gravity*, Boston: Birkhauser. 3.1.5

Bardeen, J.M. (1980) Gauge-invariant cosmological perturbations. *Physical Review D*, **22**, 1882–1905, doi:10.1103/PhysRevD.22.1882. 3.10, 4.1

Barker, B.M., Byrd, G.G., and O'Connell, R.F. (1986) Relativistic Kepler's third law. *Astrophysical Journal*, **305**, 623–633, doi:10.1086/164276. 6.4.2, 6.4.2

Barker, B.M., O'Brien, G.M., and O'Connell, R.F. (1981) Relativistic quadrupole moment. *Physical Review D*, **24**, 2332–2335, doi:10.1103/PhysRevD.24.2332. 6.4.1

Barker, B.M. and O'Connell, R.F. (1974) Nongeodesic motion in general relativity. *General Relativity and Gravitation*, **5** (5), 539–554. 3.8.3

Barker, B.M. and O'Connell, R.F. (1976) Lagrangian-Hamiltonian formalism for the gravitational two-body problem with spin and parameterized post-Newtonian parameters gamma and beta. *Physical Review D*, **14**, 861–869, doi:10.1103/PhysRevD.14.861. 6.3.4, 6.4.1

Barker, B.M. and O'Connell, R.F. (1984) Time transformations in post-Newtonian Lagrangians. *Physical Review D*, **29**, 2721–2725, doi:10.1103/PhysRevD.29.2721. 6.4.2

Barker, B.M. and O'Connell, R.F. (1987) On the completion of the post-Newtonian gravitational two-body problem with spin. *Journal of Mathematical Physics*, **28**, 661–667, doi:10.1063/1.527600. 6.4.1, 6.4.1

Barriga, P., Blair, D.G., Coward, D., Davidson, J., Dumas, J., Howell, E., Ju, L., Wen, L., Zhao, C., McClelland, D.E., Scott, S.M., Slagmolen, B.J.J., Inta, R., Munch, J., Ottaway, D.J., Veitch, P., Hosken, D., Melatos, A., Chung, C., Sammut, L., Galloway, D.K., Marx, J., Whitcomb, S., Shoemaker, D., Hughes, S.A., Reitze, D.H., Iyer, B.R., Dhurandhar, S.V., Souradeep, T., Unnikrishnan, C.S., Rajalakshmi, G., Man, C.N., Heidmann, A., Cohadon, P., Briant, T., Grote, H., Danzmann, K., Lück, H., Willke, B., Strain, K.A., Sathyaprakash, B.S., Cao, J., Cheung, Y., and Zhang, Y. (2010) AIGO: a southern hemisphere detector for the worldwide array of ground-based interferometric gravitational wave detectors. *Classical and Quantum Gravity*, **27** (8), 084 005–+, doi:10.1088/0264-9381/27/8/084005. 3.10.5

Barrow, J.D. and Levin, J. (2001) Twin paradox in compact spaces. *Physical Review A*, **63** (4), 044 104–+, doi:10.1103/PhysRevA.63.044104. 3.8.7

Barrow, J.D. and Levin, J. (2003) The Copernican principle in compact spacetimes. *Monthly Notices of the Royal Astronomical Society*, **346**, 615–618, doi:10.1046/j.1365-2966.2003.07117.x. 3.8.7

Bartelmann, M. (1998) Cluster Potential Reconstruction, in *Eighteenth Texas Symposium on Relativistic Astrophysics* (eds A.V. Olinto, J.A. Frieman, and D.N. Schramm), pp. 533–+. 7.6.2.1

Barton, G. (1989) *Elements of Green's functions and propagation*, Oxford Science Publications, The Clarendon Press Oxford University Press, New York. Potentials, diffusion, and waves. 7.2

Baskaran, D. and Grishchuk, L.P. (2004) Components of the gravitational force in the field of a gravitational wave. *Classical and Quantum Gravity*, **21**, 4041–4061, doi:10.1088/0264-9381/21/17/003. 3.10.5

Battat, J., Murphy, T., Adelberger, E., Hoyle, C.D., McMillan, R., Michelsen, E., Nordtvedt, K., Orin, A., Stubbs, C., and Swanson, H.E. (2007) APOLLO: Testing Gravity with Millimeter-precision Lunar Laser Ranging. *APS Meeting Abstracts, Volume 52, Number 3*, p. K12.00003. 3.8.7, 7.7.1

Batygin, V.V. and Toptygin, I.N. (1978) *Problems in Electrodynamics: Second Edition*, Academic Press: London, England. 2.5.7, 6.3.1

Bauch, A. (2003) Caesium atomic clocks: function, performance and applications. *Measurement Science and Technology*, **14**, 1159–1173, doi:10.1088/0957-0233/14/8/301. 3.8.6.1

Beasley, A.J., Gordon, D., Peck, A.B., Petrov, L., MacMillan, D.S., Fomalont, E.B., and Ma, C. (2002) The VLBA Calibrator Survey-VCS1. *Astrophysical Journal Supplement Series*, **141**, 13–21, doi:10.1086/339806. 9.4.8

Beekman, G. (2006) I. O. Yarkovsky and the Discovery of 'his' Effect. *Journal for the History of Astronomy*, **37**, 71–86. 8.6

Behera, H. (2002) A comment on the Invariance of the Speed of Light. *Bulletin of Orissa Physical Society*, **X**, 153. URL http://arxiv.org/abs/physics/0304087. 2.2.2

Bekenstein, J.D. (2004) Relativistic gravitation theory for the modified Newtonian dynamics paradigm. *Physical Review D*, **70** (8), 083 509–+, doi:10.1103/PhysRevD.70.083509. 3.8.7

Bel, L. (1970) Dynamique des systèmes de *N* particules ponctuelles en relativité restreinte. *Ann. Inst. H. Poincaré Sect. A (N.S.)*, **12**, 307–321. 7.3

Bel, L., Damour, T., Deruelle, N., Ibanez, J., and Martin, J. (1981) Poincaré-invariant gravitational field and equations of motion of two pointlike objects: the postlinear approximation of general relativity. *General Relativity and Gravitation*, **13**, 963–1004, doi:10.1007/BF00756073. 7.2, 7.3, 7.4

Bel, L. and Fustero, X. (1976) Mécanique relativiste predictive des systèmes de n particules. *Ann. Inst. H. Poincaré Sect. A (N.S.)*, **25**, 411–436. 7.3

Bel, L., Salas, A., and Sànchez, J.M. (1973) Approximate Solutions of Predictive Relativistic Mechanics for the Electromagnetic Interaction. *Physical Review D*, **7**, 1099–1106, doi:10.1103/PhysRevD.7.1099. 7.3

Belinfante, F.J. (1939) On the spin angular momentum of mesons. *Physica*, **6**, 887–898. 3.9.4.3, 3.9.6.2

Belinfante, F.J. (1940) On the current and the density of the electric charge, the energy, the linear momentum and the angular momentum of arbitrary fields. *Physica*, **7** (5), 449 – 474, doi:DOI:10.1016/S0031-8914(40)90091-X. URL http://www.sciencedirect.com/science/article/B6X42-4CB7521-3D/2/99492acccdb3f22a27b0a73d7d0ff99e. 3.9.4.3, 3.9.6.2

Bergeron, J. (1992) Proceedings of the Twenty-first General Assembly. *Transactions of the International Astronomical Union, Series B*, **21**. 4

Bergmann, P. (1961) Observables in General Relativity. *Reviews of Modern Physics*, **33** (4), 510–514, doi:10.1103/RevModPhys.33.510. 3.8.5

Bernard, A. and Touboul, P. (1987) Gradio - An electrostatic spaceborne gravity gradiometer. *NASA STI/Recon Technical Report A*, **88**, 22 578–+. 8.1

Bertotti, B., Ashby, N., and Iess, L. (2008) The effect of the motion of the Sun on the light-time in interplanetary relativity experiments. *Classical and Quantum Gravity*, **25** (4), 045 013–+, doi:10.1088/0264-9381/25/4/045013. 1.3.5.1

Bertotti, B. and Giampieri, G. (1992) Relativistic effects for Doppler measurements near solar conjunction. *Classical and Quantum Gravity*, **9**, 777–793. 7.6.2.2, 7.7.1, 7.8, 7.8.3, 7.8.4

Bertotti, B. and Iess, L. (1985) Doppler search for a gravitational background radiation with two spacecraft. *General Relativity and Gravitation*, **17**, 1043–1058. 7.7.1, 7.8

Bertotti, B., Iess, L., and Tortora, P. (2003) A test of general relativity using radio links with the Cassini spacecraft. *Nature*, **425**, 374–376. 3.8.6.3, 7.7.1, 7.8, 7.9.6.1

Bertotti, B., Vecchio, A., and Iess, L. (1999) Gravitational waves from coalescing binaries and Doppler experiments. *Physical Review D*, **59** (8), 082 001–+. 3.10, 7.7.1, 7.8

Bevington, P.R. and Robinson, D.K. (1992) *Data reduction and error analysis for the physical sciences*, McGraw-Hill Inc., New York, 2nd edn.. With 1 IBM-PC floppy disk (5.25 inch; HD). 3.8.7

Bini, D., Geralico, A., Ruggiero, M.L., and Tartaglia, A. (2008) Emission versus Fermi coordinates: applications to relativistic positioning systems. *Classical and Quantum Gravity*, **25** (20), 205 011–+, doi:10.1088/0264-9381/25/20/205011. 2.3.3

Bini, D., Lusanna, L., and Mashhoon, B. (2005) Limitations of radar coordinates. *International Journal of Modern Physics D*, **14** (8), 1413 – 1429. URL http://www.worldscinet.com/ijmpd/14/1408/S0218271805006961.html. 2.6.4

Binney, J. and Tremaine, S. (2008) *Galactic Dynamics: Second Edition*, Princeton

University Press. 4.2.3

Birkinshaw, M. (1989) Moving Gravitational Lenses, in *Gravitational Lenses*, *Lecture Notes in Physics, Berlin Springer Verlag*, vol. 330 (eds J.M. Moran, J.N. Hewitt, and K.Y. Lo), *Lecture Notes in Physics, Berlin Springer Verlag*, vol. 330, pp. 59–+, doi:10.1007/3-540-51061-3. 7.6.2, 7.6.2.2

Birkinshaw, M. and Gull, S.F. (1983) A test for transverse motions of clusters of galaxies. *Nature*, **302**, 315–317. 7.6.2.2, 7.6.2.2, 7.9.5.2

Bjerhammar, A. (1985) On a relativistic geodesy. *Bulletin Geodesique*, **59**, 207–220. 8.1

Blanchet, L. (1987) Radiative Gravitation Fields in General Relativity II. Asymptotic Behaviour at Future Null Infinity . *Royal Society of London Proceedings Series A*, **409**, 383–389, doi:10.1098/rspa.1987.0022. 3.8.5, 4.4.1, 1

Blanchet, L. (1998) On the multipole expansion of the gravitational field. *Classical and Quantum Gravity*, **15**, 1971–1999, doi:10.1088/0264-9381/15/7/013. 3.10.1, 3.10.3, 4.3.1, 6.3.1, 7.1

Blanchet, L. (1999) Post-Newtonian Expansion of Gravitational Radiation. *Progress of Theoretical Physics Supplement*, **136**, 146–157, doi:10.1143/PTPS.136.146. 3.10.1

Blanchet, L. (2002a) Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries. *Living Reviews in Relativity*, **5** (3). 3.10.1, 3.10.3, 3.10.3, 3.10.5, 3.10.5, 3.10.5

Blanchet, L. (2002b) Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries. *Living Reviews in Relativity*, **5** (3), 1–109. URL http://www.livingreviews.org/lrr-2002-3. 5.2.2

Blanchet, L. and Damour, T. (1984a) The post-Minkowski iteration method and the structure of radiative gravitational fields. *Academie des Science Paris Comptes Rendus Serie B Sciences Physiques*, **298**, 431–434. 3.10.1

Blanchet, L. and Damour, T. (1984b) The post-Minkowski iteration method and the structure of radiative gravitational fields. *Academie des Science Paris Comptes Rendus Serie B Sciences Physiques*, **298**, 431–434. 4.1, 4.2.3

Blanchet, L. and Damour, T. (1986) Radiative gravitational fields in general relativity. I - General structure of the field outside the source. *Royal Society of London Philosophical Transactions Series A*, **320**, 379–430. 3.10.1, 3.10.1, 3.10.2, 3.10.3, 3.10.3, 3.10.4, 3.10.5, 4.2.3, 4.4.1, 5.2.5.3, 7.1, 7.9.3.1, 7.9.3.2, A

Blanchet, L. and Damour, T. (1988) Tail-transported temporal correlations in the dynamics of a gravitating system. *Physical Review D*, **37**, 1410–1435. 3.10.4, 7.2

Blanchet, L. and Damour, T. (1989) Post-Newtonian generation of gravitational waves. *Annales de L'Institut Henri Poincaré Section Physique Theorique*, **50**, 377–408. 1.2.2, 3.10.2, 3.10.2, 3.10.5, 4.1, 4.4.1, 4.4.2, 4.4.2, 4.4.2, 4.4.3, 4.4.4, 6.1.1, 6.3.1, 7.9.3.1, 7.9.3.2

Blanchet, L., Kopeikin, S., and Schäfer, G. (2001a) Gravitational Radiation Theory and Light Propagation, in *Gyros, Clocks, Interferometers ...: Testing Relativistic Gravity in Space*, *Lecture Notes in Physics, Berlin Springer Verlag*, vol. 562 (ed. C. Lämmerzahl, C. W. F. Everitt, & F. W. Hehl), *Lecture Notes in Physics, Berlin Springer Verlag*, vol. 562, pp. 141–166. 3.10.3, 3.10.5

Blanchet, L., Salomon, C., Teyssandier, P., and Wolf, P. (2001b) Relativistic theory for time and frequency transfer to order c^{-3}. *Astronomy & Astrophysics*, **370**, 320–329, doi:10.1051/0004-6361:20010233. 7.8.3

Blanchet, L. and Schäfer, G. (1993) Gravitational wave tails and binary star systems. *Classical and Quantum Gravity*, **10**, 2699–2721. 7.2

Blandford, R. and Teukolsky, S.A. (1976) Arrival-time analysis for a pulsar in a binary system. *Astrophysical Journal*, **205**, 580–591. 6.4.1, 7.6.1.1

Blome, H., Chicone, C., Hehl, F.W., and Mashhoon, B. (2010) Nonlocal modification of Newtonian gravity. *Physical Review D*, **81** (6), 065 020–+, doi:10.1103/PhysRevD.81.065020. 3.7.4, 3.8.7

Bolotovskiĭ, B.M. and Serov, A.V. (1994) METHODOLOGICAL NOTES: Details of the motion of charged nonrelativistic particles in a variable field. *Physics Uspekhi*, **37**, 515–516, doi:10.1070/PU1994v037n05ABEH000111. 84

Bondi, H. and Samuel, J. (1997) The Lense-Thirring Effect and Mach's Principle. *Physics Letters A*, **228**, 121–126. 3.1.5

Bonvin, C. (2008) Effect of peculiar motion in weak lensing. *Physical Review D*, **78** (12), 123 530–+, doi:10.1103/PhysRevD.78.123530. 7.6.2

Born, M. (1909) Die Theorie des starren Elektrons in der Kinematik des Relativitätsprinzips. *Annalen der Physik*, **30**, 1–56, doi:10.1002/andp.19093351102. 2.6.3

Boulliau, I. (1645) Astronomia philolaïca: Opus novum, in quo motus Planetarum per nouam ac veram hypothesim demonstrantur., *700 pp.*, Paris: Sumptibus S. Piget. 4

Braginskii, V.B. and Grishchuk, L.P. (1985) Kinematic resonance and the memory effect in free mass gravitational antennas. *Soviet Physics - JETP*, **62** (3), 427–430. 84

Braginskii, V.B., Polnarev, A.G., and Thorne, K.S. (1984) Foucault pendulum at the south pole - Proposal for an experiment to detect the earth's general relativistic gravitomagnetic field. *Physical Review Letters*, **53**, 863–866. 8.6

Braginskiĭ, V.B. and Polnarev, A.G. (1980) Relativistic spin quadrupole gravitational effect. *Soviet Journal of Experimental and Theoretical Physics Letters*, **31**, 415–418. 8.1, 8.7

Braginsky, V.B. (1994) Experimental gravitation (what is possible and what is interesting to measure). *Classical and Quantum Gravity*, **11**, A1–A7. 3.1.1

Braginsky, V.B. and Panov, V.I. (1972) Verification of the equivalence of inertial and gravitational mass. *Soviet Physics - JETP*, **34**, 463–466. 3.1.2

Brans, C. and Dicke, R.H. (1961) Mach's Principle and a Relativistic Theory of Gravitation. *Physical Review*, **124**, 925–935, doi:10.1103/PhysRev.124.925. 2.8.5, 2.8.5, 3.1.5, 19, 4.1, 6.1.5, 9.2.4

Breuer, R.A. and Rudolph, E. (1982) The force law for the dynamic two-body problem in the second post-Newtonian approximation of general relativity. *General Relativity and Gravitation*, **14**, 181–211. 3.1.4, 4.5.1

Broucke, R.A. and Cefola, P.J. (1972) On the Equinoctial Orbit Elements. *Celestial Mechanics*, **5**, 303–310, doi:10.1007/BF01228432. 1.4.6

Brouwer, D. and Clemence, G.M. (1961) *Methods of celestial mechanics*, New York: Academic Press. 1.4.3, 1.4.7

Brovar, V.V., Kopeikina, Z.S., and Pavlova, M.V. (2001) Solution of the Dirichlet and Stokes exterior boundary problems for the Earth's ellipsoid. *Journal of Geodesy*, **74**, 767–772, doi:10.1007/s001900000100. 8.1

Brown, E.W. (1896) *An introductory treatise on the lunar theory*, Cambridge: University Press. 4.4.1

Brumberg, V.A. (1972) *Relativistic Celestial Mechanics*, Moscow: Nauka, 1972. In Russian. 1.5.1, 1.5.7, 2.5.9, 2.8.2, 3.9.5.4, 4.1, 5.2.2, 6.1.1, 6.1.1, 6.3.2, 6.3.4, 6.3.4, 6.4.1, 6.4.2, 6.4.2, 6.4.2, 6.4.3.1, 6.4.3.4, 7.1, 7.5.3, 7.7.1

Brumberg, V.A. (1981) Relativistic reduction of astronomical measurements and reference frames, in *ASSL Vol. 86: IAU Colloq. 56: Reference Coordinate Systems for Earth Dynamics* (eds E.M. Gaposchkin and B. Kolaczek), pp. 283–294. 7.7.4

Brumberg, V.A. (1991) *Essential Relativistic Celestial Mechanics*, Bristol: Adam Hilger. 1.5.1, 1.5.7, 2.5.9, 2.8.2, 4.1, 5.2.2, 6.1.1, 6.3, 6.3.4, 6.4.1, 6.4.2, 6.4.2, 6.4.3.1, 6.4.3.2, 6.4.3.4, 7.1, 1.3.5.1, 7.5.2, 7.7.1, 7.9.3.1, 7.9.3.2, 7.9.4.1, 7.9.6.1, 8.6, 9.3.2

Brumberg, V.A., Bretagnon, P., Capitaine, N., Damour, T., Eubanks, T.M., Fukushima, T., Guinot, B., Klioner, S.A., Kopeikin, S.M., Krivov, A.V., Seidelmann, P.K., and Soffel, M.H. (1998) General Relativity and the IAU Resolutions. *Highlights of Astronomy*, **11**, 194–202. 4.1, 4.5.1

Brumberg, V.A., Evdokimova, L.S., and Kochina, N.G. (1971) Analytical methods for the orbits of artificial satellites of the moon. *Celestial Mechanics*, **3**, 197–221, doi:10.1007/BF01228033. 1.5.7

Brumberg, V.A., Klioner, S.A., and Kopeikin, S.M. (1990) Relativistic reduction of astrometric observations at POINTS level of accuracy, in *Inertial Coordinate System on the Sky*, *IAU Symposium*, vol. 141 (eds J.H. Lieske and V.K. Abalakin), *IAU Symposium*, vol. 141, pp. 229–239. 7.1, 7.5.2, 7.7.4, 7.9.1, 7.9.3.1, 7.9.3.2, 7.9.3.2, 7.9.5.2, 9.2.3

Brumberg, V.A. and Kopeikin, S.M. (1989) Relativistic equations of motion of the earth's satellite in the geocentric frame of reference. *Kinematika i Fizika Nebesnykh Tel*, **5**, 3–8. 4.1, 5.1.1, 5.2.1, 6.3.2, 8.2, 8.6

Brumberg, V.A. and Kopeikin, S.M. (1990) Relativistic time scales in the solar system. *Celestial Mechanics and Dynamical Astronomy*, **48**, 23–44. 4.5.2.3, 5.2.1, 5.2.5.1, 5.2.5.1, 5.2.5.2, 8.1, 9.3.2

Brumberg, V.A. and Kopejkin, S.M. (1989a) Relativistic Reference Systems and Motion of Test Bodies in the Vicinity of the Earth. *Nuovo Cimento B*, **103**, 63–98. 4.1, 4.5.1, 5.1.1, 5.2.1, 5.2.2, 5.2.3.1, 5.2.6, 6.1.1, 6.3.2, 7.7.4, 7.9.3.1, 8.2, 8.6, 9.3.2

Brumberg, V.A. and Kopejkin, S.M. (1989b) Relativistic theory of celestial reference frames, in *Reference Frames*, *Astrophysics and Space Science Library*, vol. 154 (eds J. Kovalevsky, I.I. Mueller, and B. Kolaczek), *Astrophysics and Space Science Library*, vol. 154, pp. 115–141. 4.1, 4.3.1, 5.1.1, 5.2.1, 5.2.2, 7.7.4, 7.9.3.1, 8.2, 8.2, 8.4, 8.4, 8.5

Buchdahl, H.A. (1959) On the trace of the energy-momentum tensor of fields associated with particles of zero rest mass. *Journal of the Australian Mathematical Society*, **1** (01), 99–105, doi:10.1017/S1446788700025118. 3.9.2.2

Buffett, B.A., Mathews, P.M., and Herring, T.A. (2002) Modeling of nutation and precession: Effects of electromagnetic coupling. *Journal of Geophysical Research (Solid Earth)*, **107**, 2070–+, doi:10.1029/2000JB000056. 9.6.3, 9.6.8

Burgay, M., D'Amico, N., Possenti, A., Manchester, R.N., Lyne, A.G., Joshi, B.C., McLaughlin, M.A., Kramer, M., Sarkissian, J.M., Camilo, F., Kalogera, V., Kim, C., and Lorimer, D.R. (2003) An increased estimate of the merger rate of double neutron stars from observations of a highly relativistic system. *Nature*, **426**, 531–533, doi:10.1038/nature02124. 7.6.1.1

Burke, W.L. (1971) Gravitational Radiation Damping of Slowly Moving Systems Calculated Using Matched Asymptotic Expansions. *Journal of Mathematical Physics*, **12**, 401–418. 5.2.1

Caldwell, R.R., Dave, R., and Steinhardt, P.J. (1998) Quintessential Cosmology Novel Models of Cosmological Structure Formation. *Astrophysics & Space Sciences*, **261**, 303–310, doi:10.1023/A:1002028230277. 4.1

Callandreau, O. (1902a) Proprietes d'une certaine anomalie pouvant remplacer les anomalies déjà connues dans le calcul des perturbations des petites planetes. *Comptes Rendus*, **134**, 1478–1482. 1.3.6

Callandreau, O. (1902b) Sur le calcul numérique des coefficients dans le developpements de la fonction perturbatrice. *J. de l'École Polytechnique*, **7** (2), 29–99. 1.3.6

Campbell, W.B., Macek, J., and Morgan, T.A. (1977) Relativistic time-dependent multipole analysis for scalar, electromagnetic, and gravitational fields. *Physical Review D*, **15**, 2156–2164, doi:10.1103/PhysRevD.15.2156. 3.10.2, 3.10.2

Capitaine, N. (1990) The celestial pole coordinates. *Celestial Mechanics and Dynamical Astronomy*, **48**, 127–143. 9.7.2

Capitaine, N. (2000) Definition of the Celestial Ephemeris Pole and the Celestial Ephemeris Origin, in *IAU Colloq. 180: Towards Models and Constants for Sub-Microarcsecond Astrometry* (eds K.J. Johnston, D.D. McCarthy, B.J. Luzum, and G.H. Kaplan), pp. 153–+. 9.6.2, 9.7.2

Capitaine, N., Andrei, A.H., Calabretta, M., Dehant, V., Fukushima, T., Guinot, B., Hohenkerk, C., Kaplan, G., Klioner, S., Kovalevsky, J., Kumkova, I., Ma, C., McCarthy, D.D., seidelmann, K., and Wallace, P. (2007) *IAU Division 1 Working Group: Nomenclature for Fundamental Astronomy (NFA)*. URL `http://syrte.obspm.fr/iauWGnfa/`. 5, 9.1.3

Capitaine, N. and Chollet, F. (1991) The Use of the Nonrotating Origin in the Computation of Apparent Places of Stars for Estimating Earth Rotation Parameters, in *IAU Colloq. 127: Reference Systems* (eds J.A. Hughes, C.A. Smith, and G.H. Kaplan), pp. 224–227. 9.7.2

Capitaine, N., Guinot, B., and McCarthy, D.D. (2000) Definition of the Celestial Ephemeris Origin and of UT1 in the International Celestial Reference Frame. *Astronomy & Astrophysics*, **355**, 398–405. 9.7.2, 9.7.5.4

Capitaine, N., Mathews, P.M., Dehant, V., Wallace, P.T., and Lambert, S.B. (2009) On the IAU 2000/2006 precession nutation and comparison with other models and VLBI observations. *Celestial Mechanics and Dynamical Astronomy*, **103**, 179–190, doi:10.1007/s10569-008-9179-9. 9.6

Capitaine, N., Souchay, J., and Guinot, B. (1986) A non-rotating origin on the instantaneous equator - Definition, properties and use. *Celestial Mechanics*, **39**, 283–307. 9.7.2

Capitaine, N., Wallace, P.T., and Chapront, J. (2003) Expressions for IAU 2000 precession quantities. *Astronomy & Astrophysics*, **412**, 567–586, doi:10.1051/0004-6361:20031539. 9.3.6.2, 9.6.3, 9.6.5

Capitaine, N., Wallace, P.T., and Chapront, J. (2005) Improvement of the IAU 2000 precession model. *Astronomy & Astrophysics*, **432**, 355–367, doi:10.1051/0004-6361:20041908. 9.6.3

Caporali, A. (1981) A reformulation of the post-Newtonian approximation to general relativity. I - The metric and the local equations of motion. II - Post-Newtonian equation of motion for extended bodies. *Nuovo Cimento B Serie*, **61**, 181–212. 5.2.2, 2, 6.1.1, 6.1.3

Capozziello, S., Lambiase, G., Papini, G., and Scarpetta, G. (1999) Higher order corrections in gravitational microlensing. *Physics Letters A*, **254**, 11–17, doi:10.1016/S0375-9601(99)00121-8. 7.9.5.2

Caputo, M. (2006) Research Note: On New Limits of the Coefficient of Gravitation Shielding. *Journal of Astrophysics and Astronomy*, **27**, 439–441. 4

Carilli, C.L. and Rawlings, S. (2004) Science with the Square Kilometre Array. *New Astronomy Review*, **48**, 979–984, doi:10.1016/j.newar.2004.09.001. 7.9.1

Carlip, S. (2000) Aberration and the speed of gravity. *Physics Letters A*, **267**, 81–87, doi:10.1016/S0375-9601(00)00101-8. 2.7.5

Carlip, S. (2004) Model-dependence of Shapiro time delay and the 'speed of gravity/speed of light' controversy. *Classical and Quantum Gravity*, **21**, 3803–3812, doi:10.1088/0264-9381/21/15/011. 7.9.4.3

Carrier, M. (1994) *The Completeness of Scientific Theories*, Kluewer Academic Publishers: Dordrecht. 62, 8.2, 7

Carroll, S.M., Duvvuri, V., Trodden, M., and Turner, M.S. (2004) Is cosmic speed-up due to new gravitational physics? *Phys. Rev. D*, **70** (4), 043 528, doi:10.1103/PhysRevD.70.043528. 3.9.3.4

Casse, R. (2006) *Projective geometry: an introduction*, Oxford University Press, Oxford. 1.3.5.2, 18

Castagnino, M. (1968) Some remarks on the Marzke-Wheeler method of measurement. *Nuovo Cimento B Serie*, **54**, 149–150, doi:10.1007/BF02711534. 7

Central Bureau of the IERS (2010) *IERS Bulletin A*. URL `http://www.iers.org/nn_11252/IERS/EN/DataProducts/EarthOrientationData/bulA__MD.html`. 9.3.2, 9.3.6.2, 9.7.5.5

Cerdonio, M. (2002) Progress in gravitational waves detection: "bar" detectors. *Nuclear Physics B Proceedings Supplements*, **110**, 197–201, doi:10.1016/S0920-5632(02)80125-6. 3.10

Chan, H.A., Moody, M.V., and Paik, H.J. (1987) Superconducting gravity gradiometer for sensitive gravity measurements. II. Experiment. *Physical Review D*, **35**, 3572–3597, doi:10.1103/PhysRevD.35.3572. 8.1, 8.7

Chan, H.A. and Paik, H.J. (1987) Superconducting gravity gradiometer for sensitive gravity measurements. I. Theory. *Physical Review D*, **35**, 3551–3571. 8.1, 8.7

Chandrasekhar, S. (1965) The Post-Newtonian Equations of Hydrodynamics in General Relativity. *Astrophysical Journal*, **142**, 1488–+, doi:10.1086/148432. 4.1, 6.3.3

Chandrasekhar, S. (1983) *The mathematical theory of black holes*, Oxford: Oxford University Press; International Series of Monographs on Physics. Volume 69, 663 pp. 3.5.4

Chandrasekhar, S. and Esposito, F.P. (1970) The $2\frac{1}{2}$-POST-NEWTONIAN Equations of Hydrodynamics and Radiation Reaction in General Relativity. *Astrophysical Journal*, **160**, 153–+, doi:10.1086/150414. 4.1, 7.7.1

Chandrasekhar, S. and Lebovitz, N.R. (1962) The Potentials and the Superpotentials of Homogeneous Ellipsoids. *Astrophysical Journal*, **136**, 1037–+, doi:10.1086/147456. 6.3.3

Chandrasekhar, S. and Nutku, Y. (1969) The Second Post-Newtonian Equations of Hydrodynamics in General Relativity.

Astrophysical Journal, **158**, 55–+, doi:10.1086/150171. 3.8.7, 3.10.5, 4.1

Chashchina, O., Iorio, L., and Silagadze, Z. (2009) Elementary Derivation of the Lense-Thirring Precession. *Acta Physica Polonica B*, **40**, 2363–+. 1.5.4

Cherepashchuk, A.M. (2003) Search for black holes. *Uspekhi Fizicheskikh Nauk*, **46**, 335–371. 5.2.5.1

Chicone, C., Kopeikin, S.M., Mashhoon, B., and Retzloff, D.G. (2001) Delay equations and radiation damping. *Physics Letters A*, **285**, 17–26, doi:10.1016/S0375-9601(01)00327-9. 1.5.4

Chupp, T.E., Hoare, R.J., Loveman, R.A., Oteiza, E.R., Richardson, J.M., Wagshul, M.E., and Thompson, A.K. (1989) Results of a new test of local Lorentz invariance: A search for mass anisotropy in ^{21}Ne. *Physical Review Letters*, **63**, 1541–1545, doi:10.1103/PhysRevLett.63.1541. 3.1.5

Ciufolini, I. (1986) Measurement of the lense-thirring drag on high-altitude, laser-ranged artificial satellites. *Physical Review Letters*, **56** (4), 278–281, doi:10.1103/PhysRevLett.56.278. 1.5.4, 8.6

Ciufolini, I. (2007) Dragging of inertial frames. *Nature*, **449**, 41–47, doi:10.1038/nature06071. 8.5

Ciufolini, I., Kopeikin, S., Mashhoon, B., and Ricci, F. (2003) On the gravitomagnetic time delay. *Physics Letters A*, **308**, 101–109. 7.6.2.1

Ciufolini, I. and Pavlis, E.C. (2004) A confirmation of the general relativistic prediction of the Lense-Thirring effect. *Nature*, **431**, 958–960, doi:10.1038/nature03007. 8.5, 8.6

Ciufolini, I., Pavlis, E.C., and Peron, R. (2006) Determination of frame-dragging using Earth gravity models from CHAMP and GRACE. *New Astronomy*, **11**, 527–550, doi:10.1016/j.newast.2006.02.001. 8.5

Ciufolini, I. and Wheeler, J.A. (1995) *Gravitation and Inertia*, Princeton: Princeton University Press. 3.1.5, 8.5

Cocconi, G. and Salpeter, E.E. (1958) A search for anisotropy of inertia. *Il Nuovo Cimento*, **10**, 646–651. 3.1.5

Cocconi, G. and Salpeter, E.E. (1960) Upper Limit for the Anisotropy of Inertia from the Mössbauer Effect. *Physical Review Letters*, **4**, 176 – 177. 3.1.5

Cochran, W.D. (1996) Precise Measurement of Stellar Radial Velocities, in *Proceedings for the Workshop on High Resolution Data Processing held in June 1996, at the Headquarter of the National Astronomical Observatory, Tokyo, Japan. SUBARA Telescope Technical Report No. 55. Edited by M. Iye, T. Takata, and E. J. Wampler, p.30*, pp. 30–+. 7.5.3

Cohen, I.B. and Smith, G.E. (eds) (2002) *The Cambridge companion to Newton*, Cambridge Companions to Philosophy, Cambridge University Press, Cambridge. 3

Coleman, R.A. and Korté, H. (2001) Hermann Weyl: mathematician, physicist, philosopher, in *Hermann Weyl's Raum-Zeit-Materie and a general introduction to his scientific work (Schloss Reisensburg, 1992)*, *DMV Sem.*, vol. 30, Birkhäuser, Basel, pp. 157–386. 2.3.1

Coleman, S. and Glashow, S. (1999) High-energy tests of Lorentz invariance. *Physical Review D*, **59**, 116 008. 2.2.3

Coll, B. and Pozo, J.M. (2006) Relativistic positioning systems: the emission coordinates. *Classical and Quantum Gravity*, **23**, 7395–7416, doi:10.1088/0264-9381/23/24/012. 2.3.3, 26

Contopoulos, G. and Spyrou, N. (1976) The Center of Mass in the Post-Newtonian Approximation of General Relativity. *Astrophysical Journal*, **205**, 592–598, doi:10.1086/154316. 2, 6.1.1

Coope, J.A.R. (1970) Irreducible Cartesian Tensors. III. Clebsch-Gordan Reduction. *Journal of Mathematical Physics*, **11**, 1591–1612, doi:10.1063/1.1665301. 3.10.3

Coope, J.A.R. and Snider, R.F. (1970) Irreducible Cartesian Tensors. II. General Formulation. *Journal of Mathematical Physics*, **11**, 1003–1017, doi:10.1063/1.1665190. 3.10.3

Coope, J.A.R., Snider, R.F., and McCourt, F.R. (1965) Irreducible Cartesian Tensors. *Journal of Chemical Physics*, **43**, 2269–2275, doi:10.1063/1.1697123. 3.10.3, 3.10.3

Courbin, F., Saha, P., and Schechter, P.L. (2002) Quasar Lensing, in *Gravitational Lensing: An Astrophysical Tool, Lecture Notes in Physics, Berlin Springer Verlag*, vol. 608 (eds F. Courbin and D. Minniti), *Lecture Notes in Physics, Berlin Springer Verlag*, vol. 608, pp. 1–+. 7.9.1

Cowsik, R. (2000) Radio Ranging Techniques to Test Relativistic Gravitation. *Gravitation and Cosmology*, **6**, 323–326. 7.5.3

Crosta, M.T. and Mignard, F. (2006) Microarcsecond light bending by Jupiter. *Classical and Quantum Gravity*, **23**, 4853–4871, doi:10.1088/0264-9381/23/15/006. 7.9.1, 7.9.3.2, 7.9.3.2, 7.9.5.2, 7.9.5.2

Crowley, R.J. and Thorne, K.S. (1977) The generation of gravitational waves. II - The postlinear formalism revisited. *Astrophysical Journal*, **215**, 624–635. 7.2

Currie, D.G. (1966) Poincaré-Invariant Equations of Motion for Classical Particles. *Physical Review*, **142**, 817–824, doi:10.1103/PhysRev.142.817. 7.3

Cutri, R.M., Skrutskie, M.F., van Dyk, S., Beichman, C.A., Carpenter, J.M., Chester, T., Cambresy, L., Evans, T., Fowler, J., Gizis, J., Howard, E., Huchra, J., Jarrett, T., Kopan, E.L., Kirkpatrick, J.D., Light, R.M., Marsh, K.A., McCallon, H., Schneider, S., Stiening, R., Sykes, M., Weinberg, M., Wheaton, W.A., Wheelock, S., and Zacarias, N. (2003) *2MASS All Sky Catalog of point sources.*, The IRSA 2MASS All-Sky Point Source Catalog, NASA/IPAC Infrared Science Archive. URL http://irsa.ipac.caltech.edu/applications/Gator/. 9.4.8

Dallas, S.S. (1977) Equations of motion for rotating finite bodies in the extended PPN formalism. *Celestial Mechanics*, **15**, 111–123, doi:10.1007/BF01229052. 6.1.1, 6.3.2, 6.3.4

Damour, T. (1983) Gravitational radiation and the motion of compact bodies, in *Gravitational Radiation* (eds N. Deruelle and T. Piran), pp. 58–144. 1.5.4, 4.3.1, 5.1.1, 5.2.1, 5.2.5.4, 6.1.1, 6.3.1, 6.3.1, 7.3, 7.4, 7.6.1.1

Damour, T. (1984) Strong-field effects and time asymmetry in general relativity and in bimetric gravitation theory. *Foundations of Physics*, **14**, 987–995, doi:10.1007/BF01889249. 7.6.1.1

Damour, T. (1987) The problem of motion in Newtonian and Einsteinian gravity, pp. 128 - 198, in *Three hundred years of gravitation* (eds S. Hawking and W. Israel), Cambridge: Cambridge University Press, pp. 128–198. 2.7.5, 68, 3.10.1, 3.10.1, 4.2.2, 4.2.3, 4.2.3, 5.1.1, 6.3.4, 6.4.1, 7.2

Damour, T. (2000) Experimental Tests of Relativistic Gravity. *Nuclear Physics B Proceedings Supplements*, **80**, 41–50, doi:10.1016/S0920-5632(99)00828-2. 4.1, 4.2.3, 6.4.1

Damour, T. (2009a) The Equivalence Principle and the Constants of Nature. *Space Science Reviews*, **148**, 191–199, doi:10.1007/s11214-009-9533-6. 3.1.1

Damour, T. (2009b) The Equivalence Principle and the Constants of Nature. *Space Science Reviews*, pp. 52–+, doi:10.1007/s11214-009-9533-6. 8.7

Damour, T. and Deruelle, N. (1985) General relativistic celestial mechanics of binary systems. I. The post-Newtonian motion. *Ann. Inst. Henri Poincaré Phys. Théor., Vol. 43, No. 1, p. 107 - 132*, **43**, 107–132. 6.4.1, 6.4.3.2, 6.4.3.2, 7.6.1.4

Damour, T. and Deruelle, N. (1986) General relativistic celestial mechanics of binary systems. II. The post-Newtonian timing formula. *Ann. Inst. Henri Poincaré Phys. Théor., Vol. 44, No. 3, p. 263 - 292*, **44**, 263–292. 6.4.1, 6.4.3.2, 6.4.3.2, 6.4.3.2, 7.6.1.1, 7.6.1.4, 7.6.1.4

Damour, T. and Esposito-Farese, G. (1992) Tensor-multi-scalar theories of gravitation . *Classical and Quantum Gravity*, **9**, 2093–2176. 19, 3.9.3.4, 4.2.1, 4.2.1, 4.3.1, 4.4.1, 4.4.4, 4.4.5, 5.2.1, 5.2.5.4, 9.2.4

Damour, T. and Esposito-Farèse, G. (1992) Testing local Lorentz invariance of gravity with binary-pulsar data. *Physical Review D*, **46**, 4128–4132. 7.6.1.1

Damour, T. and Esposito-Farese, G. (1993) Nonperturbative strong-field effects in tensor-scalar theories of gravitation. *Physical Review Letters*, **70**, 2220–2223. 3.9.3.4

Damour, T. and Esposito-Farèse, G. (1996) Testing gravity to second post-Newtonian order: A field-theory approach. *Physical Review D*, **53**, 5541–5578. 7.1

Damour, T. and Esposito-Farèse, G. (1998a) Gravitational-wave versus binary-pulsar tests of strong-field gravity. *Physical Review D*, **58** (4), 042 001–+. 3.9.3.4, 7.6.1.1

Damour, T. and Esposito-Farèse, G. (1998b) Light deflection by gravitational waves from localized sources. *Physical Review D*,

58 (4), 044 003–+. 7.1, 7.9.5.2

Damour, T., Grishchuk, L.P., Kopejkin, S.M., and Schäfer, G. (1989) Higher-order relativistic dynamics of binary systems, in *Proceedings of the Fifth Marcel Grossmann Meeting on General Relativity, Part A, B (Perth, 1988)*, World Sci. Publ., Teaneck, NJ, pp. 451–459. 68

Damour, T. and Iyer, B.R. (1991a) Multipole analysis for electromagnetism and linearized gravity with irreducible Cartesian tensors. *Physical Review D*, **43**, 3259–3272, doi:10.1103/PhysRevD.43.3259. 3.10.3, 3.10.3, 3.10.3, 3.10.3, 4.1, 6.2.1

Damour, T. and Iyer, B.R. (1991b) Post-newtonian generation of gravitational waves. II. The spin moments. *Ann. Inst. Henri Poincaré Phys. Théor., Vol. 54, No. 2, p. 115 - 164*, **54**, 115 – 164. 4.4.1, 6.2.1

Damour, T. and Nordtvedt, K. (1993a) General relativity as a cosmological attractor of tensor-scalar theories. *Physical Review Letters*, **70**, 2217–2219. 3.9.3.4, 7.9.6.1

Damour, T. and Nordtvedt, K. (1993b) Tensor-scalar cosmological models and their relaxation toward general relativity. *Physical Review D*, **48**, 3436–3450, doi:10.1103/PhysRevD.48.3436. 3.9.3.4, 4.2.3

Damour, T. and Schäfer, G. (1985) Lagrangians for n point masses at the second post-Newtonian approximation of general relativity. *General Relativity and Gravitation*, **17**, 879–905, doi:10.1007/BF00773685. 1.5.4, 5.2.2, 6.4.2, 7.7.1

Damour, T. and Schäfer, G. (1988) Higher-order relativistic periastron advances and binary pulsars. *Nuovo Cimento B Serie*, **101**, 127–176, doi:10.1007/BF02828697. 6.4.3.2

Damour, T. and Schäfer, G. (1991) New tests of the strong equivalence principle using binary-pulsar data. *Physical Review Letters*, **66**, 2549–2552. 7.6.1.1

Damour, T., Soffel, M., and Xu, C. (1991) General-relativistic celestial mechanics. I. Method and definition of reference systems. *Physical Review D*, **43**, 3273–3307. 4.2.2, 4.5.1, 4.5.2.3, 5.1.1, 5.2.1, 5.2.5.1, 5.2.6, 6.1.1, 6.1.3, 6.1.4, 7.7.4, 7.9.3.1

Damour, T., Soffel, M., and Xu, C. (1992) General-relativistic celestial mechanics. II. Translational equations of motion. *Physical Review D*, **45**, 1017–1044. 4.2.2, 4.5.1, 5.1.1, 5.2.1, 5.2.6, 6.1.1, 6.3.4, 7.7.4

Damour, T., Soffel, M., and Xu, C. (1993) General-relativistic celestial mechanics. III. Rotational equations of motion. *Physical Review D*, **47**, 3124–3135. 4.2.2, 6.1.1, 6.2.1, 7.7.4

Damour, T., Soffel, M., and Xu, C. (1994) General-relativistic celestial mechanics. IV. Theory of satellite motion. *Physical Review D*, **49**, 618–635. 5.2.2

Damour, T. and Taylor, J.H. (1992) Strong-field tests of relativistic gravity and binary pulsars. *Physical Review D*, **45**, 1840–1868. 6.4.1, 6.4.3.2, 6.4.3.2, 6.4.3.4, 7.6.1.1, 7.6.1.4

de Felice, A. and Tsujikawa, S. (2010) f(R) Theories. *Living Reviews in Relativity*, **13** (3). 3.9.3.4, 3.9.3.4

de Felice, F. and Clarke, C.J.S. (1990) *Relativity on Curved Manifolds*, Cambridge University Press. 2.1.1, 18, 3.7.5, 42

de Felice, F., Crosta, M.T., Vecchiato, A., Lattanzi, M.G., and Bucciarelli, B. (2004) A General Relativistic Model of Light Propagation in the Gravitational Field of the Solar System: The Static Case. *Astrophysical Journal*, **607**, 580–595, doi:10.1086/383244. 7.9.1

de Felice, F., Vecchiato, A., Crosta, M.T., Bucciarelli, B., and Lattanzi, M.G. (2006) A General Relativistic Model of Light Propagation in the Gravitational Field of the Solar System: The Dynamical Case. *Astrophysical Journal*, **653**, 1552–1565, doi:10.1086/508701. 7.9.1

D'Eath, P.D. (1975a) Dynamics of a small black hole in a background universe. *Physical Review D*, **11**, 1387–1403. 3.10.1, 4.2.2, 4.5.1, 5.1.1, 5.2.1, 6.1.1

D'Eath, P.D. (1975b) Interaction of two black holes in the slow-motion limit. *Physical Review D*, **12**, 2183–2199. 3.10.1, 4.2.2, 4.5.1, 5.1.1, 5.2.1, 6.1.1

Debra, D.B. (2003) Drag-free control for fundamental physics missions. *Advances in Space Research*, **32**, 1221–1226, doi:10.1016/S0273-1177(03)90321-8. 8.4

Dehlinger, P. (1978) *Marine gravity*, Amsterdam ; New York : Elsevier Scientific Pub. Co. : distributors for the U.S. and

Canada, Elsevier North-Holland, 1978. 8.6

Deller, A.T., Bailes, M., and Tingay, S.J. (2009) Implications of a VLBI Distance to the Double Pulsar J0737-3039A/B. *Science*, **323**, 1327–, doi:10.1126/science.1167969. 7.6.1.1

Demiański, M. and Grishchuk, L.P. (1974) Note on the motion of black holes. *General Relativity and Gravitation*, **5**, 673–679, doi:10.1007/BF00761925. 5.2.1, 6.1.1

Deruelle, N., Katz, J., and Ogushi, S. (2004) Conserved charges in Einstein Gauss Bonnet theory. *Classical and Quantum Gravity*, **21**, 1971–1985, doi:10.1088/0264-9381/21/8/004. 3.9.6.3

Deruelle, N. and Mukhanov, V.F. (1995) Matching conditions for cosmological perturbations. *Physical Review D*, **52**, 5549–5555, doi:10.1103/PhysRevD.52.5549. 5.2.2

Descartes, R. (1644) Principia Philosophiae. , Amsterdam: Apud Ludovicum Elzevirium. For a modern edition see: Descartes, René. *Principles of Philosophy*, V. R. Miller and R. P. Miller (trans.), Dordrecht: Kluwer 1983. 1.1.2

Deser, S. (1970) Self-interaction and gauge invariance. *General Relativity and Gravitation*, **1**, 9–18. 3.2.4

Deser, S. and Laurent, B. (1973) Linearity and Parametrisation of Gravitational Effects. *Astronomy & Astrophysics*, **25**, 327–+. 3.8.7

Desloge, E.A. (1989) A theoretical device for space and time measurements. *Foundations of Physics*, **19**, 1191–1213, doi:10.1007/BF00731879. 7

Dewdney, P.E., Hall, P.J., Schilizzi, R.T., and Lazio, T.J.L.W. (2009) The Square Kilometre Array. *Proceedings of the IEEE*, **97**, 1482–1496. URL http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5136190&isnumber=5165124?tag=1. 7.7.1, 7.9.1

Dicke, R.H. (1962a) Long-Range Scalar Interaction. *Physical Review*, **126**, 1875–1877, doi:10.1103/PhysRev.126.1875. 4.1, 6.1.1, 6.1.5, 6.3.4

Dicke, R.H. (1962b) Mach's Principle and Invariance under Transformation of Units. *Physical Review*, **125**, 2163–2167, doi:10.1103/PhysRev.125.2163. 3.1.5, 6.1.5

Dicke, R.H. (1965) The weak and strong principles of equivalence. *Annals of Physics*, **31**, 235–239, doi:10.1016/0003-4916(65)90239-3. 3.1.2, 3.1.4

Dickey, J.O., Bender, P.L., Faller, J.E., Newhall, X.X., Ricklefs, R.L., Ries, J.G., Shelus, P.J., Veillet, C., Whipple, A.L., Wiant, J.R., Williams, J.G., and Yoder, C.F. (1994) Lunar Laser Ranging-a Continuing Legacy of the Apollo Program. *Science*, **265** (5171), 482–490. 3.8.7, 6.1.3, 6.1.5, 7.7.1

Dillon, S.C., Gebel, G., and Pryor, L.L. (1977) Navigation at the prime meridian revisited. *Navigation*, **24**, 264–266. 9.7.2

D'Inverno, R.A. (1992) *Introducing Einstein's relativity*, Oxford University Press. 2.6.3

Dirac, P.A.M. (1937) The cosmological constants. *Nature*, **139**, 323–+, doi:10.1038/139323a0. 4.2.1

Dirac, P.A.M. (1951) Is there an Æther? *Nature*, **168**, 906–907, doi:10.1038/168906a0. 1.1.2

Dixon, W.G. (1979) Extended bodies in general relativity: their description and motion., in *Isolated Gravitating Systems in General Relativity* (ed. J. Ehlers), Amsterdam: North-Holland, pp. 156–219. 2.7.7, 3.8.3, 3.8.3, 3.8.3, 4.2.1, 6.1.1, 6.1.4

Doroshenko, O.V. and Kopeikin, S.M. (1990) High Precision Pulse Timing for Single Pulsars. *Soviet Astronomy*, **34**, 496–+. 7.7.2

Doroshenko, O.V. and Kopeikin, S.M. (1995) Relativistic effect of gravitational deflection of light in binary pulsars. *Monthly Notices of the Royal Astronomical Society*, **274**, 1029–1038. 7.6.1.2, 7.6.1.4, 7.6.1.4, 7.7.2

Dremin, I.M. (2009) Physics at the large hadron collider. *Physics-Uspekhi*, **52** (6), 541–548. URL http://stacks.iop.org/1063-7869/52/i=6/a=R03. 2.7.2

Drever, L.W.P. (1961) A Search for Anisotropy of Inertial Mass Using a Free Precession Technique. *Philosophical Magazine*, **6**, 683–687. 3.1.5

Dubath, F., Alice Gasparini, M., and Durrer, R. (2007) Microlensing modulation by quadrupole variation. *Physical Review D*, **75** (2), 024 015–+, doi:10.1103/PhysRevD.75.024015. 7.1

Dubrovin, B.A., Fomenko, A.T., and Novikov, S.P. (1984) *Modern Geometry*, New York: Springer-Verlag. 1.3.4, 1.4.1, 2.6, 3.9.2.1, 4.2.1, 4.3.1, 4.5.1, 4.5.2, 5.1.2

Dutta Choudhury, S.B. and Bhattacharya, D. (1980) On Birkhoff's theorem in a general class of scalar-tensor theories. *Journal of Physics A Mathematical General*, **13**, 3643–3647, doi:10.1088/0305-4470/13/12/013. 6.3.1

Dymnikova, I.G. (1986) Motion of particles and photons in the gravitational field of a rotating body (In memory of Vladimir Afanas'evich Ruban). *Soviet Physics Uspekhi*, **29**, 215–237, doi:10.1070/PU1986v029n03ABEH003178. 7.6.2.1

Dyson, L., Kleban, M., and Susskind, L. (2002) Disturbing implications of a cosmological constant. *Journal of High Energy Physics*, **2002** (10), 011. URL http://stacks.iop.org/1126-6708/2002/i=10/a=011. 3.9.3.4

Eckhaus, W. (1973) *Matched asymptotic expansions and singular perturbations.*, Amsterdam, North-Holland Pub. Co.; New York, American Elsevier Pub. Co., 1973. 5.1.1, 5.2.1, 5.2.2

Edberg, S.J., Laskin, R.A., Marr, IV, J.C., Unwin, S.C., and Shao, M. (2007) SIM PlanetQuest science and technology: a status report, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 6693, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 6693, doi:10.1117/12.732728. 7.7.4, 7.9.1

Edwards, R.T., Hobbs, G.B., and Manchester, R.N. (2006) TEMPO2, a new pulsar timing package - II. The timing model and precision estimates. *Monthly Notices of the Royal Astronomical Society*, **372**, 1549–1574, doi:10.1111/j.1365-2966.2006.10870.x. 7.6.1.1, 7.6.1.4, 7.7.2

Efroimsky, M. (2002a) Equations for the Keplerian Elements: Hidden Symmetry. *Preprint No 1844 of the Institute for Mathematics and Its Applications, University of Minnesota.* Http://www.ima.umn.edu/preprints/feb02/feb02. 1.4.2, 1.5.1, 1.5.6

Efroimsky, M. (2002b) Implicit gauge symmetry emerging in the N-body problem of celestial mechanics, *Tech. Rep.*, ArXiv Astrophysics e-prints: astro-ph/0212245. Http://arxiv.org/abs/astro-ph/0212245. 1.4.2, 1.5.1, 1.5.6

Efroimsky, M. (2005a) Gauge Freedom in Orbital Mechanics. *New York Academy Sciences Annals*, **1065**, 346–374, doi:10.1196/annals.1370.016. 1.5.7

Efroimsky, M. (2005b) Long-Term Evolution of Orbits About A Precessing Oblate Planet: 1. The Case of Uniform Precession. *Celestial Mechanics and Dynamical Astronomy*, **91**, 75–108, doi:10.1007/s10569-004-2415-z. 1.5.5

Efroimsky, M. and Escapa, A. (2007) The theory of canonical perturbations applied to attitude dynamics and to the Earth rotation. Osculating and nonosculating Andoyer variables. *Celestial Mechanics and Dynamical Astronomy*, **98**, 251–283, doi:10.1007/s10569-007-9080-y. 1.5.8

Efroimsky, M. and Goldreich, P. (2003) Gauge symmetry of the N-body problem in the Hamilton-Jacobi approach. *Journal of Mathematical Physics*, **44**, 5958–5977, doi:10.1063/1.1622447. 1.4.2, 1.5.1, 1.5.4, 1.5.5, 1.5.6

Efroimsky, M. and Goldreich, P. (2004) Gauge freedom in the N-body problem of celestial mechanics. *Astronomy & Astrophysics*, **415**, 1187–1199, doi:10.1051/0004-6361:20034058. 1.4.2, 1.5.1, 1.5.4, 1.5.5

Ehlers, J. (1973) Survey of general relativity theory., in *Relativity, Astrophysics and Cosmology*, *Astrophysics and Space Science Library*, vol. 38 (ed. W. Israel), Reidel: Dordrecht-Holland, *Astrophysics and Space Science Library*, vol. 38, pp. 1–125. 3.8.2, 3.8.4.2

Ehlers, J. and Lämmerzahl, C. (eds) (2006) *Special relativity: Will it Survive the Next 101 Years?*, Lecture Notes in Physics, Springer-Verlag: Berlin. 2.2.3

Ehlers, J., Pirani, F.A.E., and Schild, A. (1972) The geometry of free fall and light propagation, in *General Relativity. Papers in Honor of J. L. Synge* (ed. L. O'Raifeartaigh), Clarendon Press: Oxford, pp. 5–19. 42

Ehlers, J., Rosenblum, A., Goldberg, J.N., and Havas, P. (1976) Comments on gravitational radiation damping and energy loss in binary systems. *Astrophysical Journal Letters*, **208**, L77–L81, doi:10.1086/182236. 7.7.1

Ehlers, J. and Rudolph, E. (1977) Dynamics of extended bodies in general relativity center-of-mass description and quasirigidity. *General Relativity and Gravitation*, **8**, 197–217, doi:10.1007/BF00763547. 3.8.3

Ehresmann, C. (1951) Les connexions infinitésimales dans un espace fibré différentiable, in *Colloque de topologie (espaces fibrés), Bruxelles, 1950*, Georges Thone, Liège, pp. 29–55. 3.4

Einstein, A. (1905) Zur Elektrodynamik bewegter Körper. *Annalen der Physik*, **322**, 891–921, doi:10.1002/andp.19053221004~. English translation by H. M. Schwartz: *The Americal Journal of Physics,* Vol. 45, p. 18 (1977). 2.2.1, 2.2.3, 2.5.6

Einstein, A. (1907) Über die Möglichkeit einer neuen Prüfung des Relativitätsprinzips. *Jahrbuch der Radioaktivität und Elektronik*, **4**, 411–462. 2.6.3

Einstein, A. (1918) Prinzipielles zur allgemeinen Relativitätstheorie. *Annalen der Physik*, **55**, 16–19. 3.1.5, 3.8, 3.8.1

Einstein, A. (1920a) *Ether and the Theory of Relativity: An Address delivered on 5 May 1920, in the University of Leyden*, The Collected Papers of Albert Einstein, Vol. 7. The Berlin Years: Writings, 1918-1921, pp. 305–. 1.1.2

Einstein, A. (1920b) *Relativity: The special and general theory*, New York: Henry Holt and Company . Modern edition: Penguin Books Ltd, London 2006. 2.2.1

Einstein, A., Infeld, L., and Hoffmann, B. (1938) The gravitational equations and the problem of motion. *Ann. of Math. (2)*, **39** (1), 65–100. 2.8.5, 4.1, 5.2.5.4, 6.1.1, 6.3.4

Einstein, A., Podolsky, B., and Rosen, N. (1935) Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Physical Review*, **47**, 777–780, doi:10.1103/PhysRev.47.777. 3.8.7

Einstein, A. and Rosen, N. (1935) The Particle Problem in the General Theory of Relativity. *Physical Review*, **48**, 73–77, doi:10.1103/PhysRev.48.73. 5.2.1

Ellis, G.F.R. and Uzan, J.P. (2005) c is the speed of light, isn't it? *American Journal of Physics*, **73**, 240–247, doi:10.1119/1.1819929. 2.2.1, 63, 83, 7.1, 7.9.4.3

Epstein, R. (1977) The binary pulsar - Post-Newtonian timing effects. *Astrophysical Journal*, **216**, 92–100, doi:10.1086/155449. 6.4.1, 6.4.3.3

Epstein, R. and Wagoner, R.V. (1975) Post-Newtonian generation of gravitational waves. *Astrophysical Journal*, **197**, 717–723, doi:10.1086/153561. 4.4.1

Estabrook, F.B. (1969) Post-Newtonian N-body Equations of the Brans-Dicke Theory. *Astrophysical Journal*, **158**, 81–+, doi:10.1086/150172. 6.3.4

Estabrook, F.B. and Wahlquist, H.D. (1975) Response of Doppler spacecraft tracking to gravitational radiation. *General Relativity and Gravitation*, **6**, 439–447. 7.5.3, 7.8

Euler, L. (1748) *Recherches sur la question des inegalites du mouvement de Saturne et de Jupiter.*, Berlin: Piece qui a remporte le prix de l'academie royale des sciences. For modern edition see: L. Euler *Opera mechanica et astronomica.* (Birkhauser-Verlag, Switzerland, 1999. 1.4.1

Euler, L. (1753) *Theoria motus Lunae exhibens omnes ejus inaequalitates.*, Impensis Academiae Imperialis Scientarum Petropolitanae. For modern edition see: L. Euler *Opera mechanica et astronomica.* (Birkhauser-Verlag, Switzerland, 1999. 1.4.1, 1.4.5

Evans, L. and Bryant, P. (2008) Lhc machine. *Journal of Instrumentation*, **3** (08), S08 001 (VI+158). URL http://iopscience.iop.org/1748-0221/3/08/S08001/. 2.7.2

Everitt, C.W.F., Adams, M., Bencze, W., Buchman, S., Clarke, B., Conklin, J., DeBra, D.B., Dolphin, M., Heifetz, M., Hipkins, D., Holmes, T., Keiser, G.M., Kolodziejczak, J., Li, J., Lockhart, J.M., Muhlfelder, B., Parkinson, B.W., Salomon, M., Silbergleit, A., Solomonik, V., Stahl, K., Turneaure, J.P., and Worden, Jr., P.W. (2008) Gravity Probe B data analysis status and potential for improved accuracy of scientific results. *Classical and Quantum Gravity*, **25** (11), 114 002–+, doi:10.1088/0264-9381/25/11/114002. 8.5

Everitt, C.W.F., Adams, M., Bencze, W., Buchman, S., Clarke, B., Conklin, J.W., Debra, D.B., Dolphin, M., Heifetz, M., Hipkins, D., Holmes, T., Keiser, G.M., Kolodziejczak, J., Li, J., Lipa, J., Lockhart, J.M., Mester, J.C., Muhlfelder, B., Ohshima, Y., Parkinson, B.W., Salomon, M., Silbergleit, A., Solomonik, V., Stahl, K., Taber, M., Turneaure, J.P., Wang, S., and Worden, P.W. (2009a) Gravity Probe B Data Analysis. Status and Potential for Improved Accuracy of Scientific Results. *Space Science Reviews*, **148**, 53–69, doi:10.1007/s11214-009-9524-7. 3.8.3

Everitt, C.W.F., Adams, M., Bencze, W., Buchman, S., Clarke, B., Conklin, J.W., Debra, D.B., Dolphin, M., Heifetz, M., Hipkins, D., Holmes, T., Keiser, G.M., Kolodziejczak, J., Li, J., Lipa, J., Lockhart, J.M., Mester, J.C., Muhlfelder, B., Ohshima, Y., Parkinson, B.W., Salomon, M., Silbergleit, A., Solomonik, V., Stahl, K., Taber, M., Turneaure, J.P., Wang, S., and Worden, P.W. (2009b) Gravity Probe B Data Analysis. *Space Science Reviews*, pp. 65–+, doi:10.1007/s11214-009-9524-7. 8.5

Fairhead, L. and Bretagnon, P. (1990) An analytical formula for the time transformation TB-TT. *Astronomy & Astrophysics*, **229**, 240–247. 9.3.6.1

Fairhurst, S., Guidi, G.M., Hello, P., Whelan, J.T., and Woan, G. (2010) Current status of gravitational wave observations. *General Relativity and Gravitation*, pp. 90–+, doi:10.1007/s10714-010-1009-1. 3.10.5

Fatibene, L., Ferraris, M., Francaviglia, M., and Lusanna, L. (2010) ADM Pseudotensors, Conserved Quantities and Covariant Conservation Laws in General Relativity. *ArXiv e-prints*. 3.9.6.3

Fecko, M. (2006) *Differential Geometry and Lie Groups for Physicists*, Cambridge: Cambridge University Press. 3.7.5

Feissel, M. and Mignard, F. (1998) The adoption of ICRS on 1 January 1998: meaning and consequences. *Astronomy & Astrophysics*, **331**, L33–L36. 9.4.4, 9.4.7, 14

Ferraris, M., Francaviglia, M., Palese, M., and Winterroth, E. (2008) Canonical Connections in Gauge-Natural Field Theories. *International Journal of Geometric Methods in Modern Physics*, **5**, 973–+, doi:10.1142/S0219887808003144. 3.9.7

Ferraro, R. (2007) *Einstein's Space-Time: An introduction to special and general relativity*, New York: Springer Science. 1.1.2

Ferraro, R. and Fiorini, F. (2008) Born-infeld gravity in weitzenböck spacetime. *Physical Review D*, **78** (12), 124 019, doi:10.1103/PhysRevD.78.124019. 3.7.4

Fey, A.L., Boboltz, D., Charlot, P., Fomalont, E., Lanyi, G., Jacobs, C., and K-Q VLBI Survey Collaboration (2009) Extending the ICRF to Higher Radio Frequencies, in *American Astronomical Society Meeting Abstracts*, *American Astronomical Society Meeting Abstracts*, vol. 214, *American Astronomical Society Meeting Abstracts*, vol. 214, pp. 411.03–+. 9.4.8

Fey, A.L., Ma, C., Arias, E.F., Charlot, P., Feissel-Vernier, M., Gontier, A.M., Jacobs, C.S., Li, J., and MacMillan, D.S. (2004) The Second Extension of the International Celestial Reference Frame: ICRF-EXT.1. *Astronomical Journal (USA)*, **127**, 3587–3608, doi:10.1086/420998. 8.6, 9.4.5.1

Feyerabend, P.K. (1994) *Problems of Empiricism*, Cambridge University Press: Cambridge. 7

Fichtengolz, I.G. (1950) Lagrangian for equations of motion in the second approximation of Einstein's theory of gravity. *Journal of Experimental and Theoretical Physics*, **20**, 233–. 3.1.4

Fierz, M. (1956) über die physikalische deutung der erweiterten gravitationstheorie p. jordans. *Helvetica Physica Acta*, **29**, 128–+. 2.8.5, 19, 4.1

Fixler, J.B., Foster, G.T., McGuirk, J.M., and Kasevich, M.A. (2007) Atom Interferometer Measurement of the Newtonian Constant of Gravity. *Science*, **315**, 74–, doi:10.1126/science.1135459. 8.1

Flanagan, É.É. and Hughes, S.A. (2005) The basics of gravitational wave theory. *New Journal of Physics*, **7**, 204–+, doi:10.1088/1367-2630/7/1/204. 3.10.4, 3.10.5

Fock, V. (1957) Three Lectures on Relativity Theory. *Reviews of Modern Physics*, **29**, 325–333, doi:10.1103/RevModPhys.29.325. 6.1.1, 7.9.3.1, 7.9.3.1

Fock, V. (1964) *Theory of Space, Time and Gravitation*, Oxford: Pergamon Press. 2.5.3, 2.8.2, 2.8.3, 3.8.5, 56, 3.9.5.4, 3.10.3, 3.10.5, 4.1, 4.2.1, 4.2.2, 4.2.3, 4.2.3, 4.2.4, 4.3.1, 4.3.1, 4.3.2, 4.4.4, 5.1.1, 5.2.1, 5.2.2, 5.2.5.4, 5.2.5.4, 6.1.1, 6.1.1, 6.3, 6.3.2, 6.3.4, 6.3.4, 7.1, 7.2, 7.2, 7.7.1, 7.9.3.1, 8.2, 8.2

Folkner, W.M. (2007) *JPL "README" file*.

URL ftp://ssd.jpl.nasa.gov/pub/eph/planets/README.txt. 9.5.2

Fomalont, E., Kopeikin, S., Lanyi, G., and Benson, J. (2009a) Progress in Measurements of the Gravitational Bending of Radio Waves Using the VLBA. *Astrophysical Journal*, **699**, 1395–1402, doi:10.1088/0004-637X/699/2/1395. 3.8.6.3, 7.9.6.1, 7.9.6.1

Fomalont, E. and Reid, M. (2004) Microarcsecond astrometry using the SKA. *New Astronomy Review*, **48**, 1473–1482, doi:10.1016/j.newar.2004.09.037. 7.1, 7.7.1, 7.7.3, 7.9.1

Fomalont, E.B. (2005) Phase Referencing Using More Than One Calibrator, in *Future Directions in High Resolution Astronomy*, *Astronomical Society of the Pacific Conference Series*, vol. 340 (eds J. Romney and M. Reid), *Astronomical Society of the Pacific Conference Series*, vol. 340, pp. 460–+. 7.7.3

Fomalont, E.B. and Kopeikin, S. (2002) Phase Referencing Using Several Calibrator Sources, in *Proceedings of the 6th EVN Symposium* (eds E. Ros, R.W. Porcas, A.P. Lobanov, and J.A. Zensus), pp. 53–+. 7.7.3

Fomalont, E.B. and Kopeikin, S. (2008) Radio interferometric tests of general relativity, in *IAU Symposium*, *IAU Symposium*, vol. 248 (eds W.J. Jin, I. Platais, and M.A.C. Perryman), *IAU Symposium*, vol. 248, pp. 383–386, doi:10.1017/S1743921308019613. 1.3.5.1, 7.7.3, 7.9.1, 7.9.4.3

Fomalont, E.B., Kopeikin, S., Titov, O., and Honma, M. (2009b) VLBA/VERA Test of General Relativity. *American Astronomical Society, IAU Symposium #261. Relativity in Fundamental Astronomy: Dynamics, Reference Frames, and Data Analysis 27 April - 1 May 2009 Virginia Beach, VA, USA, #15.03; Bulletin of the American Astronomical Society, Vol. 41, p.890*, **261**, 1503–+. 7.9.4.3

Fomalont, E.B. and Kopeikin, S.M. (2003) The Measurement of the Light Deflection from Jupiter: Experimental Results. *Astrophysical Journal*, **598**, 704–711, doi:10.1086/378785. 4.1, 1.3.5.1, 7.5.2, 7.7.1, 12, 7.7.3, 7.9.1, 7.12, 7.9.4.2, 7.9.4.3, 7.9.4.3, 7.9.6.2, 7.9.6.2

Foster, R.S. and Backer, D.C. (1990) Constructing a pulsar timing array. *Astrophysical Journal*, **361**, 300–308, doi:10.1086/169195. 3.10.5, 9.3

Frank, P. and Rothe, H. (1911) Über die Transformation der Raum-Zeitkoordinaten von ruhenden auf bewegte Systeme . *Annalen der Physik*, **34**, 825–855, doi:10.1002/andp.19113390502. 2.2.2

French, R.G., Nicholson, P.D., Cooke, M.L., Elliot, J.L., Matthews, K., Perkovic, O., Tollestrup, E., Harvey, P., Chanover, N.J., Clark, M.A., Dunham, E.W., Forrest, W., Harrington, J., Pipher, J., Brahic, A., Grenier, I., Roques, F., and Arndt, M. (1993) Geometry of the Saturn system from the 3 July 1989 occultation of 28 SGR and Voyager observations. *Icarus*, **103**, 163–214, doi:10.1006/icar.1993.1066. 7.9.1, 7.9.5.2

Fricke, W., Schwan, H., Corbin, T., Bastian, U., Bien, R., Cole, C., Jackson, E., Jährling, R., Jahreiß, H., Lederle, T., and Röser, S. (1991) Fifth fundamental catalogue. Part 2: The FK5 extension - new fundamental stars. *Veroeffentlichungen des Astronomischen Rechen-Instituts Heidelberg*, **33**, 1–143. 9.4.4

Fricke, W., Schwan, H., Lederle, T., Bastian, U., Bien, R., Burkhardt, G., Du Mont, B., Hering, R., Jährling, R., Jahreiß, H., Röser, S., Schwerdtfeger, H.M., and Walter, H.G. (1988) Fifth fundamental catalogue (FK5). Part 1: The basic fundamental stars. *Veroeffentlichungen des Astronomischen Rechen-Instituts Heidelberg*, **32**, 1–106. 9.4.4

Frittelli, S. (2003a) Aberration by gravitational lenses in motion. *Monthly Notices of the Royal Astronomical Society*, **344**, L85–L87, doi:10.1046/j.1365-8711.2003.07095.x. 7.6.2, 7.9.6.1

Frittelli, S. (2003b) On bending angles by gravitational lenses in motion. *Monthly Notices of the Royal Astronomical Society*, **340**, 457–462, doi:10.1046/j.1365-8711.2003.06302.x. 7.6.2, 7.9.5.2, 7.9.6.1

Frittelli, S., Kling, T.P., and Newman, E.T. (2002) Fermat potentials for nonperturbative gravitational lensing. *Physical Review D*, **65** (12), 123 007–+, doi:10.1103/PhysRevD.65.123007. 7.9.5.2

Frolov, V.P. (1979) The Newman-Penrose method in the theory of general relativity., in *Problems in the general theory of relativity and theory of group representatives* (ed.

N.G. Basov), Moscow: P. N. Lebedev Physics Institute, 1979, pp. 73–185. 2.3.3, 40, 3.9.2.2, 7.2, 7.5.3

Fukushima, T. (1988) The Fermi Coordinate System in the Post-Newtonian Framework. *Celestial Mechanics*, **44**, 61–75, doi:10.1007/BF01230707. 4.1, 5.2.6, 6.1.1, 8.2

Fukushima, T. (2002) Report on astronomical constants. *Highlights of Astronomy*, **12**, 107–112. 4.5.2.3

Fukushima, T. (2004) Efficient Orbit Integration by the Orbital Longitude Method Using Antifocal Anomaly. *Astronomical Journal (USA)*, **128**, 1455–1464, doi:10.1086/423042. 1.3.6, 1.3.6

Fukushima, T., Fujimoto, M.K., Kinoshita, H., and Aoki, S. (1986) Coordinate systems in the general relativistic framework, in *Relativity in Celestial Mechanics and Astrometry. High Precision Dynamical Theories and Observational Verifications*, *IAU Symposium*, vol. 114 (eds J. Kovalevsky and V.A. Brumberg), *IAU Symposium*, vol. 114, pp. 145–168. 8.2

Fushchich, W.I. and Nikitin, A.G. (1987) *Symmetries of Maxwell's equations*, Dordrecht: D. Reidel Publishing Company. 2.3.6

Futamase, T. and Itoh, Y. (2007) The Post-Newtonian Approximation for Relativistic Compact Binaries. *Living Reviews in Relativity*, **10** (2). 4.1, 6.3.1, 6

Futamase, T. and Schutz, B.F. (1983) Newtonian and post-Newtonian approximations are asymptotic to general relativity. *Physical Review D*, **28**, 2363–2372, doi:10.1103/PhysRevD.28.2363. 4.1

Gebel, G. and Matthews, B. (1971) Navigation at the prime meridian. *Journal of the Institute of Navigation*, **18**, 141–146. 9.7.2

Gelfand, I.M. (1964) *Generalized functions.*, New York: Academic Press, 1964. 4.4.1

Gelfand, I.M., Minlos, R.A., and Shapiro, Z.J. (1958) *Representations of the rotation and Lorentz groups and their applications*, Gosudarstv. Izdat. Fiz.-Mat. Lit., Moscow. 5.2.5.3, 5.2.5.4, A, A, A

Gergely, L.Á., Keresztes, Z., and Mikóczi, B. (2006) An Efficient Method for the Evaluation of Secular Effects in the Perturbed Keplerian Motion. *Astrophysical Journal Supplement Series*, **167**, 286–291, doi:10.1086/507707. 6.4.1, 6.4.2

Gibble, K. (2007) Atomic Clocks and Precision Measurements. *International Journal of Modern Physics D*, **16**, 2495–2497, doi:10.1142/S0218271807011383. 3.8.6.1, 8.1

Gill, E., Schastok, J., Soffel, M.H., and Ruder, H. (1989) On the Mashhoon-Theiss "anomaly". *Physical Review D*, **39**, 2441–2443. 8.1

Ginzburg, V.L. (1970) *The propagation of electromagnetic waves in plasmas*, Oxford: Pergamon. 7.3

Göbel, E., Mills, I.M., and Wallard, A.J. (2006) *The International System of Units (SI), 8th ed.*, Sèvres: Bureau International des Poids et Mesures. URL `http://www.bipm.org/utils/common/pdf/si_brochure_8_en.pdf`. 1, 9.3.1

Goldreich, P. (1965) Inclination of satellite orbits about an oblate precessing planet. *Astronomical Journal (USA)*, **70**, 5–9, doi:10.1086/109673. 1.5.7

Goodkind, J.M. (1999) The superconducting gravimeter. *Review of Scientific Instruments*, **70**, 4131–4152, doi:10.1063/1.1150092. 8.1

Gorelik, G. (1993) Vladimir Fock: Philosophy of Gravity and Gravity of Philosophy, in *The Attraction of Gravitation: New Studies in the History of General Relativity* (ed. Earman, J., Janssen, M., & Norton, J. D.), Boston: Birkhaeuser, pp. 308–331. URL `http://people.bu.edu/gorelik/Fock_ES-93_text.htm`. 3.8.5

Gould, A. (1993) Deflection of light by the earth. *Astrophysical Journal Letters*, **414**, L37–L40, doi:10.1086/186990. 7.7.4

Gray, J.J. (1995) Poincaré, Einstein, and others. *The Mathematical Intelligencer*, **17**, 65–67. 2.2.3

Greenberg, O. W. (2002) *CPT* Violation Implies Violation of Lorentz Invariance. *Physical Review Letters*, **89** (23), 231 602, doi:10.1103/PhysRevLett.89.231602. 3.2.1

Grishchuk, L.P. (1977) Graviton Creation in the Early Universe, in *Eighth Texas Symposium on Relativistic Astrophysics*, *Annals of the New York Academy of Sciences*, vol. 302 (ed. M. D. Papagiannis), *Annals of the New York Academy of Sciences*, vol. 302, pp. 439–+, doi:10.1111/j.1749-6632.1977.tb37064.x. 3.10.5

Grishchuk, L.P. (2009) Some Uncomfortable Thoughts on the Nature of Gravity, Cosmology, and the Early Universe. *Space Science Reviews*, **148**, 315–328, doi:10.1007/s11214-009-9509-6. 3.2.4, 3.9.7

Grishchuk, L.P. and Kopeikin, S.M. (1986) Equations of motion for isolated bodies with relativistic corrections including the radiation reaction force, in *Relativity in Celestial Mechanics and Astrometry. High Precision Dynamical Theories and Observational Verifications*, *IAU Symposium*, vol. 114 (eds J. Kovalevsky and V.A. Brumberg), *IAU Symposium*, vol. 114, pp. 19–33. 1.5.4, 68, 5.1.1, 6.1.1, 7.6.1.1

Grishchuk, L.P., Petrov, A.N., and Popova, A.D. (1984) Exact theory of the (Einstein) gravitational field in an arbitrary background spacetime. *Communications in Mathematical Physics*, **94**, 379–396. 3.2.4, 3.9.1, 3.9.7, 3.10

Grishchuk, L.P. and Polnarev, A.G. (1989) Gravitational wave pulses with velocity memory. *Soviet Physics - JETP*, **96**, 1153–1160. 84

Gromov, A., Baryshev, Y., and Teerikorpi, P. (2004) Two-fluid matter-quintessence FLRW models: Energy transfer and the equation of state of the universe. *Astronomy & Astrophysics*, **415**, 813–820. 3.9.3.4

Gruber, T., Abrikosov, O., and Hugentobler, U. (2009) *GOCE Standards. GOCE High Level Processing Facility.*, The European GOCE Gravity Consortium EGG-C, Document No. GO-TN-HPF-GS-0111, Issue 3.1. 8.3, 8.6

Gubanov, V.S. (1997) *Obobshchennyi metod naimenshikh kvadratov*, “Nauka” S.-Peterburg. Otdel., St. Petersburg. Teoriya i primenenie v astrometrii. [Theory and application in astrometry]. 3.8.7

Guinot, B. (1979) Basic problems in the kinematics of the rotation of the earth, in *Time and the Earth’s Rotation*, *IAU Symposium*, vol. 82 (eds D.D. McCarthy and J.D.H. Pilkington), *IAU Symposium*, vol. 82, pp. 7–18. 9.3.1, 9.7, 9.7.2

Guinot, B. (1981) Comments on the terrestrial pole of reference, the origin of the longitudes, and on the definition of UT1, in *IAU Colloq. 56: Reference Coordinate Systems for Earth Dynamics*, *Astrophysics and Space Science Library*, vol. 86 (eds E.M. Gaposchkin and B. Kolaczek), *Astrophysics and Space Science Library*, vol. 86, pp. 125–134. 9.7, 9.7.2

Guinot, B. (1986) Is the International Atomic Time TAI a coordinate time or a proper time? *Celestial Mechanics*, **38**, 155–161. 9.3.1, 9.3.2

Guinot, B. and Felicitas Arias, E. (2005) Atomic time-keeping from 1955 to the present. *Metrologia*, **42**, 20–30, doi:10.1088/0026-1394/42/3/S04. 8.1

Guinot, B. and Petit, G. (1991) Atomic time and the rotation of pulsars. *Astronomy & Astrophysics*, **248**, 292–296. 9.3

Guinot, B. and Seidelmann, P.K. (1988) Time scales - Their history, definition and interpretation. *Astronomy & Astrophysics*, **194**, 304–308. 8.1, 9.3, 9.3.1

Gundlach, J.H., Schlamminger, S., and Wagner, T. (2009) Laboratory Tests of the Equivalence Principle at the University of Washington. *Space Science Reviews*, pp. 105–+, doi:10.1007/s11214-009-9609-3. 3.1.1

Gurfil, P. (2007) Generalized solutions for relative spacecraft orbits under arbitrary perturbations. *Acta Astronautica*, **60**, 61–78. 1.5.4

Gurfil, P. and Belyanin, S. (2008) The gauge-generalized Gyldén Meshcherskii Problem. *Advances in Space Research*, **42**, 1313–1317, doi:10.1016/j.asr.2008.01.019. 1.5.4

Gurfil, P. and Klein, I. (2006) Mitigating the integration error in numerical simulations of Newtonian systems. *International Journal for Numerical Methods in Engineering*, **68**, 267–297. 1.5.2, 1.5.2

Gurvits, L.I. and Mitrofanov, I.G. (1986) Perturbation of the Background Radiation by a Moving Gravitational Lens. *Nature*, **324**, 349–+. 7.6.2.2, 7.9.5.2

Guth, A.H. (1981) Inflationary universe: A possible solution to the horizon and flatness problems. *Physical Review D*, **23**, 347–356, doi:10.1103/PhysRevD.23.347. 2.8.5, 3.9.3.4

Hagihara, Y. (1933) On the theory of secular aberration. *Proceedings of the Physico-Mathemtical Society of Japan*, **15** (4), 155–174. 2.5.8

Hagiwara, Y., Fomalont, E., Tsuboi, M., and Yasuhiro, M. (eds) (2009) *Approaching*

Micro-Arcsecond Resolution with VSOP-2: Astrophysics and Technologies, *Astronomical Society of the Pacific Conference Series*, vol. 402. 7.7.1

Harada, W. and Fukushima, T. (2003) Harmonic Decomposition of Time Ephemeris TE405. *Astronomical Journal (USA)*, **126**, 2557–2561, doi:10.1086/378909. 9.3.2, 9.3.6.1, 9.3.6.1

Hardorp, J. (1985) A Demonstration of the Sun's Motion around the Barycenter of the Solar System, in *Bulletin of the American Astronomical Society*, *Bulletin of the American Astronomical Society*, vol. 17, *Bulletin of the American Astronomical Society*, vol. 17, pp. 592–+. 7.7.2

Hartmann, T., Soffel, M.H., and Kioustelidis, T. (1994) On the use of STF-tensors in celestial mechanics. *Celestial Mechanics and Dynamical Astronomy*, **60**, 139–159, doi:10.1007/BF00693097. 1.2.2

Haugan, M.P. (1985) Post-Newtonian arrival-time analysis for a pulsar in a binary system. *Astrophysical Journal*, **296**, 1–12, doi:10.1086/163414. 6.4.1, 6.4.3.3, 6.4.3.3, 6.4.3.3

Haugen, M.P. and Lämmerzahl, C. (2001) Principles of Equivalence: Their Role in Gravitation Physics and Experiments that Test Them., in *Gyros, Clocks, Interferometers...: Testing Relativistic Gravity in Space.* (ed. Lämmerzahl, C. and Everitt, C.W.F. and Hehl F.W.), Springer-Verlag, pp. 241–244. 3.1.3

Hawking, S.W. and Ellis, G.F.R. (1975) *The Large Scale Structure of Space-Time*, Cambridge University Press: Cambridge. 2.2.4.4, 3.9.5.4, 7.9.4.3

Hazard, C., Sutton, J., Argue, A.N., Kenworthy, C.M., Morrison, L.V., and Murray, C.A. (1971) 3C 273B-Coincidence of radio and optical positions. *Nature Physical Science*, **233**, 89–+. 16

Hehl, F.W. and Kerlick, G.D. (1978) Metric-affine variational principles in general relativity. I - Riemannian spacetime. *General Relativity and Gravitation*, **9**, 691–710, doi:10.1007/BF00760141. 3.5.2, 3.5.3

Hehl, F.W. and Obukhov, Y.N. (2007) Elie Cartan's torsion in geometry and in field theory, an essay. *ArXiv e-prints*. 3.5.2, 31, 3.7.4, 60

Hehl, F.W. and Weinberg, S. (2007) Note on the Torsion Tensor. *Physics Today*, **60** (3), 030 000–+, doi:10.1063/1.2718743. 3.5.2

Heinämäki, P. (2006) The Planck Mission and beyond, in *Stellar Evolution at Low Metallicity: Mass Loss, Explosions, Cosmology*, *Astronomical Society of the Pacific Conference Series*, vol. 353 (eds H.J.G.L.M. Lamers, N. Langer, T. Nugis, and K. Annuk), *Astronomical Society of the Pacific Conference Series*, vol. 353, pp. 403–+. 7.6.2.2

Heinicke, C., Baekler, P., and Hehl, F.W. (2005) Einstein-æther theory, violation of Lorentz invariance, and metric-affine gravity. *Physical Review D*, **72** (2), 025 012–+, doi:10.1103/PhysRevD.72.025012. 3.5.3, 3.5.3, 3.5.4, 3.7.4, 3.8.7, 7.9.4.3

Heiskanen, W.A. and Moritz, H. (1967) *Physical geodesy*, San Francisco: W. H. Freeman. 8.1

Hellings, R.W. (1986a) Erratum: "Relativistic Effects in Astronomical Timing Measurements" [ASTRON. J. 91,650 (1986)]. *Astronomical Journal (USA)*, **92**, 1446, doi:10.1086/114282. 7.1, 7.7.1, 7.7.2, 7.7.3

Hellings, R.W. (1986b) Relativistic effects in astronomical timing measurements. *Astronomical Journal (USA)*, **91**, 650–659, doi:10.1086/114048. 7.1, 7.7.1, 7.7.2, 7.7.3

Hermann, J. (1710) Metodo d'investigare l'orbite de' pianeti, nell'ipotesi che le forze centrali o pure le gravitá delgi stessi pianeti sono in ragione reciproca de' quadrati delle distanze, che i medesimi tengono dal centro, a cui si dirigono le forze stesse. *Giornale de Letterati D'Italia*, **2**, 447–467. 1.3.4

Higgs, P.W. (1964) Broken Symmetries and the Masses of Gauge Bosons. *Physical Review Letters*, **13**, 508–509, doi:10.1103/PhysRevLett.13.508. 2.8.5, 4.1

Hill, H.R. (1967) Instantaneous Action-at-a-Distance in Classical Relativistic Mechanics. *J. Math. Phys.*, **8**, 201–220, doi:10.1063/1.1705187. 7.3

Hilton, J.L., Capitaine, N., Chapront, J., Ferrandiz, J.M., Fienga, A., Fukushima, T., Getino, J., Mathews, P., Simon, J.L., Soffel, M., Vondrak, J., Wallace, P., and Williams, J. (2006) Report of the International Astronomical Union Division I. Working Group on Precession and the Ecliptic.

Celestial Mechanics and Dynamical Astronomy, **94**, 351–367, doi:10.1007/s10569-006-0001-2. 9.6.3, 9.6.5

Hilton, J.L. and Hohenkerk, C.Y. (2004) Rotation matrix from the mean dynamical equator and equinox at J2000.0 to the ICRS. *Astronomy & Astrophysics*, **413**, 765–770, doi:10.1051/0004-6361:20031552. 18

Hinch, E.J. (1991) *Perturbation methods*, Cambridge Texts in Applied Mathematics, Cambridge: Cambridge University Press, 1991. 5.2.1, 5.2.2

Hobbs, G. (2008) Gravitational wave detection using high precision pulsar observations. *Classical and Quantum Gravity*, **25** (11), 114 032–+, doi:10.1088/0264-9381/25/11/114032. 9.3

Hobbs, G. (2010) Pulsar timing array projects, in *IAU Symposium*, *IAU Symposium*, vol. 261 (ed. S. A. Klioner, P. K. Seidelmann, & M. H. Soffel), *IAU Symposium*, vol. 261, pp. 228–233, doi:10.1017/S1743921309990445. 9.3

Hobbs, G.B., Bailes, M., Bhat, N.D.R., Burke-Spolaor, S., Champion, D.J., Coles, W., Hotan, A., Jenet, F., Kedziora-Chudczer, L., Khoo, J., Lee, K.J., Lommen, A., Manchester, R.N., Reynolds, J., Sarkissian, J., van Straten, W., To, S., Verbiest, J.P.W., Yardley, D., and You, X.P. (2009) Gravitational-Wave Detection Using Pulsars: Status of the Parkes Pulsar Timing Array Project. *Publications of the Astronomical Society of Australia*, **26**, 103–109, doi:10.1071/AS08023. 7.6.1.4

Hobbs, G.B., Edwards, R.T., and Manchester, R.N. (2006) TEMPO2, a new pulsar-timing package - I. An overview. *Monthly Notices of the Royal Astronomical Society*, **369**, 655–672, doi:10.1111/j.1365-2966.2006.10302.x. 7.6.1.1, 7.7.2

Hofmann-Wellenhof, B. and Moritz, H. (2005) *Physical Geodesy*, Berlin: Springer, 2005. 8.1, 8.1, 8.2

Høg, E., Fabricius, C., Makarov, V.V., Urban, S., Corbin, T., Wycoff, G., Bastian, U., Schwekendiek, P., and Wicenec, A. (2000) The Tycho-2 catalogue of the 2.5 million brightest stars. *Astronomy & Astrophysics*, **355**, L27–L30. 9.4.8

Holton, G. (1973) *Thematic Origins of Scientific Thought: Kepler to Einstein*, Harvard University Press. 2.2.3

Huang, T.Y., Xu, B.X., Zhu, J., and Zhang, H. (1989) The concepts of International Atomic Time (TAI) and Terrestrial Dynamic Time (TDT). *Astronomy & Astrophysics*, **220**, 329–334. 8.1

Hubbard, W.B., Porco, C.C., Hunten, D.M., Rieke, G.H., Rieke, M.J., McCarthy, D.W., Haemmerle, V., Clark, R., Turtle, E.P., Haller, J., McLeod, B., Lebofsky, L.A., Marcialis, R., Holberg, J.B., Landau, R., Carrasco, L., Elias, J., Buie, M.W., Persson, S.E., Boroson, T., West, S., and Mink, D.J. (1993) The occultation of 28 SGR by Saturn - Saturn pole position and astrometry. *Icarus*, **103**, 215–234, doi:10.1006/icar.1993.1067. 7.9.1, 7.9.5.2

Hughes, J.A., Smith, C.A., and Kaplan, G.H. (eds) (1991) *Proceedings of the 127th Colloquium of the International Astronomical Union on Reference Systems*. 11, 9.4.4

Hughes, V.W., Robinson, H.G., and Beltran-Lopez, V. (1960) Upper limit for the anisotropy of inertial mass from nuclear resonance experiments. *Phys. Rev. Lett.*, **4** (7), 342–344, doi:10.1103/PhysRevLett.4.342. 3.1.5

Hulse, R.A. and Taylor, J.H. (1975) Discovery of a pulsar in a binary system. *Astrophysical Journal Letters*, **195**, L51–L53, doi:10.1086/181708. 6.4.1

Huygens, C. (1669) Rules Concerning the Motion of Bodies after Mutual Impulse. *Philosophical Transactions*, **1**, 173 – 174. For a modern edition see: Huygens, C. 1940. Œuvres Complètes, Vol. 16. Martinus Nijhoff, The Hague. 2.1.3

IAU Secretariat (2010) *IAU Information Bulletin*. URL http://www.iau.org/science/publications/iau/information_bulletin/. 9.1.2

IAU SOFA Center (2009) *Standards of Fundamental Astronomy: SOFA*. URL http://www.iausofa.org/. 9.1.3, 9.4.6, 9.6.6

Ibañez, J., Martin, J., and Ruiz, E. (1984) Gravitational interaction of two spinning particles in general relativity. II. *General Relativity and Gravitation*, **16**, 225–242, doi:10.1007/BF00762538. 7.4

Ichikawa, R., Sekido, M., Ohsaki, H., Koyama, Y., Kondo, T., Ohnishi, T., Yoshikawa, M., Cannon, W., Novikov, A., Bérubé, M., and Nozomi Dvlbi Group (2004) An Evaluation of VLBI Observations for the Deep Space Tracking of the Interplanetary Spacecrafts, in *International VLBI Service for Geodesy and Astrometry 2004 General Meeting Proceedings. Ottawa, Canada, February 9-11, 2004. Edited by Nancy R. Vandenberg and Karen D. Baver, NASA/CP-2004-212255, 2004. Published online at http://ivscc.gsfc.nasa.gov, p. 253* (eds N.R. Vandenberg and K.D. Baver), pp. 253–+. 7.5.2

IERS (2010) *Frequently Asked Questions (FAQs)*. URL `http://www.iers.org/IERS/EN/Service/FAQs/faq.html`. 9.1.3

IERS Rapid Service/Prediction Centre (2010) *IERS Earth rotation data: Bulletins A & B*. URL `http://www.iers.org/nn_10970/IERS/EN/Publications/Bulletins/bulletins.html?__nnn=true`. 9.3.2, 9.3.4, 9.3.6.2

Iess, L. and Asmar, S. (2007) Probing Space-Time in the Solar System:. from Cassini to Bepicolombo. *International Journal of Modern Physics D*, **16**, 2117–2126, doi:10.1142/S0218271807011449. 7.8

Iess, L., Giampieri, G., Anderson, J.D., and Bertotti, B. (1999) Doppler measurement of the solar gravitational deflection. *Classical and Quantum Gravity*, **16**, 1487–1502, doi:10.1088/0264-9381/16/5/303. 7.5.3, 7.6.2.2, 7.7.1, 7.8

Ilyasov, Y.P., Kopeikin, S.M., and Rodin, A.E. (1998) The astronomical timescale based on the orbital motion of a pulsar in a binary system. *Astronomy Letters*, **24**, 228–236. 9.3

Infeld, L. and Plebanski, J. (1960) *Motion and Relativity*, Oxford: Pergamon Press, 1960. 3.8.5, 4.1, 5.2.2, 6.3, 6.3.4

Irwin, A.W. and Fukushima, T. (1999) A numerical time ephemeris of the Earth. *Astronomy & Astrophysics*, **348**, 642–652. 4.5.2.3, 5.2.5.1, 5.2.5.1, 8.1, 9.3.2

The Astronomical Almanac (2010) . 5, 9.1.3, 9.3.2, 9.3.3, 9.3.6.2, 9.4.8, 9.5.4, 9.6.4, 9.6.5, 9.6.7, 9.7.5.1, 9.7.5.2, 9.7.5.4, 9.7.5.5

Itin, Y. and Marolf, D. (2001) Coframe Teleparallel Models of Gravity. *International Journal of Modern Physics D*, **10**, 547–573, doi:10.1142/S0218271801001025. 3.7.4

Jackson, J.D. (1975) *Classical electrodynamics*, New York: Wiley, 1975, 2nd ed. 7.2, 7.2

Jackson, J.D. (1998) *Classical Electrodynamics*, Wiley-VCH, pp. 832. 2.8.4, 3.10.4, 6.3.1

Jacobs, C.S., Lanyi, G.E., Naudet, C.J., Sovers, O.J., Zhang, L.D., Charlot, P., Gordon, D., Ma, C., and KQ VLBI Survey Collaboration (2003) Extending the ICRF to Higher Radio Frequencies: Initial Global Astrometric Results, in *The International Celestial Reference System, Maintenance and Future Realizations*, *IAU Joint Discussion*, vol. 16 (eds R. Gaume, D. McCarthy, and J. Souchay), USNO, *IAU Joint Discussion*, vol. 16, pp. 36–41. 11, 9.4.8

Jacobs, C.S. and Sovers, O.J. (2008) Extending the ICRF to Higher Radio Frequencies: Global Astrometric Results at 32/8 GHz, in *Measuring the Future, Proceedings of the Fifth IVS General Meeting* (eds A. Finkelstein and D. Behrend), Nauka, St. Petersburg, pp. 284–288. 9.4.8

Janssen, T. and Prokopec, T. (2007) Problems and hopes in nonsymmetric gravity. *Journal of Physics A Mathematical General*, **40**, 7067–7074, doi:10.1088/1751-8113/40/25/S63. 3.8.2

Jaranowski, P. and Schäfer, G. (1998) Third post-Newtonian higher order ADM Hamilton dynamics for two-body point-mass systems. *Physical Review D*, **57**, 7274–7291, doi:10.1103/PhysRevD.57.7274. 5.2.2

Jenet, F.A., Hobbs, G.B., Lee, K.J., and Manchester, R.N. (2005) Detecting the Stochastic Gravitational Wave Background Using Pulsar Timing. *Astrophysical Journal Letters*, **625**, L123–L126, doi:10.1086/431220. 7.7.1

Jenet, F.A., Lommen, A., Larson, S.L., and Wen, L. (2004) Constraining the Properties of Supermassive Black Hole Systems Using Pulsar Timing: Application to 3C 66B. *Astrophysical Journal*, **606**, 799–803, doi:10.1086/383020. 7.1

Johnston, K.J., Dorland, B., Gaume, R., Hennessy, G., Olling, R., Zacharias, N., Behr, B., Efroimsky, M., Hajian, A., Harris, H., Hilton, J., Kaplan, G., Monet, D., Munn, J., Pier, J., Vrba, F., Seidelmann, K., Seager, S., Pravdo, S., Coste, K., Danner, R., Grillmair, C., Stauffer, J., Boss, A., Currie, D., Danchi, W., Gould, A., Kopeikin, S., Majewski, S., Makarov, V., McMillan, R., Peterson, D.M., Shaya, E., and Unwin, S. (2006) The Origins Billions Star Survey: Galactic Explorer. *Publications of Astronomical Society of the Pacific*, **118**, 1428–1442, doi:10.1086/508903. 7.7.4, 7.9.1

Johnston, K.J., McCarthy, D.D., Luzum, B.J., and Kaplan, G.H. (eds) (2000) *IAU Colloq. 180: Towards models and constants for sub-microarcsecond astrometry*. 4.3.1, 5.2.5.3, 7.1, 7.7.1, 9.2.1

Jordan, P. (1949) Formation of the stars and development of the Universe. *Nature*, **164**, 637–+, doi:10.1038/164637a0. 2.8.5, 19, 4.1

Jordan, P. (1959) Zum gegenwärtigen Stand der Diracschen kosmologischen Hypothesen. *Zeitschrift fur Physik*, **157**, 112–121, doi:10.1007/BF01375155. 2.8.5, 19, 4.1, 4.2.1

Kadomtsev, B.B., Keldysh, L.V., Kobzarev, I.Y., and Sagdeev, R.Z. (1972) Apropos a. a. tyapkin's paper: "expression of the general properties of physical processes in the space-time metric of the special theory of relativity.". *Soviet Physics Uspekhi*, **15** (2), 230. URL http://stacks.iop.org/0038-5670/15/i=2/a=R07. 3.2.3

Kaiser, N. and Squires, G. (1993) Mapping the dark matter with weak gravitational lensing. *Astrophysical Journal*, **404**, 441–450, doi:10.1086/172297. 7.6.2.1

Kang, Z., Tapley, B., Chen, J., Ries, J., and Bettadpur, S. (2009) Geocenter variations derived from GPS tracking of the GRACE satellites. *Journal of Geodesy*, pp. 7–+, doi:10.1007/s00190-009-0307-4. 8.1

Kaplan, G. (2005a) *Celestial International Origin (CIO) right ascension as a function of time*. URL http://www.usno.navy.mil/USNO/astronomical-applications/software-products/novas/novas-c/CIO_RA.TXT. 9.7.5.3

Kaplan, G., Bangert, J., Bartlett, J., Puatua, W., and Monet, A. (2009) *Naval Observatory Vector Astrometry Software: NOVAS*. URL http://www.usno.navy.mil/USNO/astronomical-applications/software-products/novas. 9.1.3, 9.4.6, 9.6.6

Kaplan, G.H. (1981) The IAU Resolutions on Astronomical Constants, Time Scales, and the Fundamental Reference Frame. *Washington DC: United States Naval Observatory*. 9.4.4

Kaplan, G.H. (1998) High-precision algorithms for astrometry - A comparison of two approaches. *Astronomical Journal (USA)*, **115**, 361–+, doi:10.1086/300189. 7.7.1, 7.7.3, 7.7.3, 7.7.3, 7.7.4, 5

Kaplan, G.H. (2003) Celestial Pole Offsets: Conversion from (dX,dY) to ($d\psi$,$d\epsilon$). *Washington DC: United States Naval Observatory*. 9.6.8

Kaplan, G.H. (2005b) The IAU resolutions on astronomical reference systems, time scales, and earth rotation models : explanation and implementation. *U.S. Naval Observatory Circulars*, **179**. 9.6.6, 9.6.8

Kaplan, G.H. (2005c) The IAU Resolutions on Astronomical Reference Systems, Time Scales, and Earth Rotation Models: Explanation and Implementation. *Washington DC: United States Naval Observatory*. 9.4.4

Kaplan, G.H., Josties, F.J., Angerhofer, P.E., Johnston, K.J., and Spencer, J.H. (1982) Precise radio source positions from interferometric observations. *Astronomical Journal*, **87**, 570–576, doi:10.1086/113131. 16

Karlhede, A. (1982) On the linearity of the parametrised post-Newtonian metric. *Journal of Physics A Mathematical General*, **15**, L103–L104, doi:10.1088/0305-4470/15/3/004. 3.8.7

Kaspi, V.M., Taylor, J.H., and Ryba, M.F. (1994) High-precision timing of millisecond pulsars. 3: Long-term monitoring of PSRs B1855+09 and B1937+21. *Astrophysical Journal*, **428**, 713–728, doi:10.1086/174280. 7.6.1.1

Kates, R.E. (1980a) Gravitational radiation damping of a binary system containing compact objects calculated using matched asymptotic expansions. *Physical Review D*, **22**, 1871–1878, doi:10.1103/PhysRevD.22.1871. 5.2.1

Kates, R.E. (1980b) Motion of a small body through an external field in general relativity calculated by matched asymptotic expansions. *Physical Review D*, **22**, 1853–1870, doi:10.1103/PhysRevD.22.1853. 5.2.1

Kates, R.E. (1981) Underlying structure of singular perturbations on manifolds. *Annals of Physics*, **132**, 1–17. 5.2.1

Kates, R.E. and Kegeles, L.S. (1982) Nonanalytic terms in the slow-motion expansion of a radiating scalar field on a Schwarzschild background. *Physical Review D*, **25**, 2030–2037, doi:10.1103/PhysRevD.25.2030. 4.2.3, 4.2.3

Kates, R.E. and Madonna, R.G. (1982) Gravitational radiating anisotropic spheres in general relativity. *Physical Review D*, **25**, 2499–2508, doi:10.1103/PhysRevD.25.2499. 5.2.1

Kaula, W.M. (1966) *Theory of satellite geodesy. Applications of satellites to geodesy*, Waltham, Mass.: Blaisdell. 4.5.2.3, 8.8.1

Kennefick, D. (2009) Testing relativity from the 1919 eclipse - a question of bias. *Physics Today*, **62** (3), 37–42, doi:10.1063/1.3099578. 7.9.5.1

Khlebnikov, V.I. (1989) The method of geodesic deviation in precision space gravimetry. *Soviet Physics Doklady*, **34**, 523–525. 8.1

Khrapko, R.I. (2000) LETTERS TO THE EDITORS: What is mass? *Physics Uspekhi*, **43**, 1267–1270, doi:10.1070/PU2000v043n12ABEH000778. 2.7.1

Kim, Y.S. and Noz, M.E. (1986) *Theory and Applications of the Poincaré Group*, Dordrecht: D. Reidel Publishing Company. 2.3.6

Kinoshita, H. (1977) Theory of the rotation of the rigid earth. *Celestial Mechanics*, **15**, 277–326, doi:10.1007/BF01228425. 9.6.3

Kinoshita, H. (1993) Motion of the orbital plane of a satellite due to a secular change of the obliquity of its mother planet. *Celestial Mechanics and Dynamical Astronomy*, **57**, 359–368, doi:10.1007/BF00692485. 1.5.7

Kleinert, H. (ed.) (1989) *Gauge Fields In Condensed Matter. Vol. II, Stresses And Defects*, World Scientific: Singapore. 31, 3.5.4

Kleinert, H. (ed.) (2008) *Multivalued Fields in Condensed Matter, Electromagnetism, and Gravitation*, World Scientific: Singapore. 3.5.2, 3.5.3, 3.5.3, 3.5.4, 3.7.5, 42, 57

Kleinert, H. and Pelster, A. (1999) Autoparallels From a New Action Principle. *General Relativity and Gravitation*, **31**, 1439–+. 28, 29

Klioner, S., Capitaine, N., Folkner, W., Guinot, B., Huang, T.Y., Kopeikin, S., Petit, G., Pitjeva, E., Seidelmann, P.K., and Soffel, M. (2009a) Units of Relativistic Time Scales and Associated Quantities. *American Astronomical Society, IAU Symposium 261. Relativity in Fundamental Astronomy: Dynamics, Reference Frames, and Data Analysis 27 April - 1 May 2009 Virginia Beach, VA, USA, #9.03; Bulletin of the American Astronomical Society, Vol. 41, p.883*, **261**, 903–+. 4.5.2.3, 5.2.5.2, 6.1.5, 8.1, 9.3.2

Klioner, S.A. (1991a) General Relativistic Model of VLBI Observables, in *Geodetic VLBI: Monitoring Global Change*, pp. 188–+. 7.6.1.1, 7.6.1.3, 7.7.3

Klioner, S.A. (1991b) Influence of the Quadrupole Field and Rotation of Objects on Light Propagation. *Soviet Astronomy*, **35**, 523–+. 7.9.1, 7.9.5.2

Klioner, S.A. (1996) Angular velocity of rotation of extended bodies in general relativity. (Lecture), in *Dynamics, Ephemerides, and Astrometry of the Solar System*, *IAU Symposium*, vol. 172 (eds S. Ferraz-Mello, B. Morando, and J.E. Arlot), *IAU Symposium*, vol. 172, pp. 309–+. 6.2.2

Klioner, S.A. (2000) EinS: a Mathematica package for computations with indexed objects. *ArXiv General Relativity and Quantum Cosmology e-prints*. 3.9.3.3

Klioner, S.A. (2003a) A Practical Relativistic Model for Microarcsecond Astrometry in Space. *Astronomical Journal (USA)*, **125**, 1580–1597. 7.1, 7.7.1, 7.7.1, 7.7.4, 7.7.4, 7.9.1, 9.2.3, 9.4.6, 9.6.4

Klioner, S.A. (2003b) Light propagation in the gravitational field of moving bodies by means of Lorentz transformation I. Mass monopoles moving with constant velocities. *Astronomy & Astrophysics*, **404**, 783–787, doi:10.1051/0004-6361:20030559. 7.4, 7.6.1.3, 7.7.4, 7.9.6.1

Klioner, S.A. (2004) Physically adequate proper reference system of a test observer and relativistic description of the GAIA attitude. *Physical Review D*, **69** (12), 124 001–+, doi:10.1103/PhysRevD.69.124001. 7.7.4, 7.9.4.2, 8.4, 9.2.4

Klioner, S.A. (2008) Relativistic scaling of astronomical quantities and the system of astronomical units. *Astronomy & Astrophysics*, **478**, 951–958, doi:10.1051/0004-6361:20077786. 9.3.2

Klioner, S.A., Gerlach, E., and Soffel, M.H. (2009b) Relativistic Aspects of Rotational Motion of Celestial Bodies. *American Astronomical Society, IAU Symposium #261. Relativity in Fundamental Astronomy: Dynamics, Reference Frames, and Data Analysis 27 April - 1 May 2009 Virginia Beach, VA, USA, #8.03; Bulletin of the American Astronomical Society, Vol. 41, p.883*, **261**, 112–123. 9.2.4

Klioner, S.A. and Kopeikin, S.M. (1992) Microarcsecond astrometry in space - Relativistic effects and reduction of observations. *Astronomical Journal (USA)*, **104**, 897–914, doi:10.1086/116284. 7.1, 1.3.5.1, 7.6.1.3, 7.7.1, 7.7.1, 7.7.4, 7.7.4, 7.9.1

Klioner, S.A. and Kopeikin, S.M. (1994) The post-Keplerian orbital representations of the relativistic two-body problem. *Astrophysical Journal*, **427**, 951–955, doi:10.1086/174201. 3.8.6.4, 6.4.1, 6.4.3.3, 6.4.3.3, 7.1, 9.2.3

Klioner, S.A. and Soffel, M.H. (2000) Relativistic celestial mechanics with PPN parameters. *Physical Review D*, **62** (2), 024 019, doi:10.1103/PhysRevD.62.024019. 4.1, 4.2.4, 5.1.1, 5.2.1, 5.2.6, 6.1.3, 6.2.2, 6.2.2, 6.3.4, 6.3.4, 6.3.5, 9.2.4

Klioner, S.A. and Voinov, A.V. (1993) Relativistic theory of astronomical reference systems in closed form. *Physical Review D*, **48**, 1451–1461. 5.1.1, 5.1.3, 5.2.1, 7.9.3.1

Kliore, A.J., Anderson, J.D., Armstrong, J.W., Asmar, S.W., Hamilton, C.L., Rappaport, N.J., Wahlquist, H.D., Ambrosini, R., Flasar, F.M., French, R.G., Iess, L., Marouf, E.A., and Nagy, A.F. (2004) Cassini Radio Science. *Space Science Reviews*, **115**, 1–70, doi:10.1007/s11214-004-1436-y. 7.8

Kliore, A.J., Nagy, A.F., Marouf, E.A., French, R.G., Flasar, F.M., Rappaport, N.J., Anabttawi, A., Asmar, S.W., Kahann, D.S., Barbinis, E., Goltz, G.L., Fleischman, D.U., and Rochblatt, D.J. (2008) First results from the Cassini radio occultations of the Titan ionosphere. *Journal of Geophysical Research (Space Physics)*, **113**, 9317–+, doi:10.1029/2007JA012965. 7.8

Kobzarev, I.Y. (1975) FROM THE HISTORY OF PHYSICS: Henri Poincaré's St. Louis lecture, and theoretical physics on the eve of the theory of relativity. *Soviet Physics Uspekhi*, **17**, 584–592, doi:10.1070/PU1975v017n04ABEH004429. 1.1.2, 2.2.3

Kopeikin, S. (2010a) Beyond the standard IAU framework, in *IAU Symposium*, *IAU Symposium*, vol. 261 (ed. S. A. Klioner, P. K. Seidelmann, & M. H. Soffel), *IAU Symposium*, vol. 261, pp. 7–15, doi:10.1017/S1743921309990081. 3.9.3.4, 4.5.2.3, 5.1.1, 6.1.5

Kopeikin, S. (2010b) Beyond the standard IAU framework, in *IAU Symposium*, *IAU Symposium*, vol. 261 (ed. S. A. Klioner, P. K. Seidelmann, & M. H. Soffel), *IAU Symposium*, vol. 261, pp. 7–15, doi:10.1017/S1743921309990081. 9.2.4

Kopeikin, S. and Fomalont, E.B. (2002) Measuring the Speed of Propagation of Gravity, in *Bulletin of the American Astronomical Society*, *Bulletin of the American Astronomical Society*, vol. 34, *Bulletin of the American Astronomical Society*, vol. 34, pp. 1270–+. 1.3.5.1, 7.9.1, 7.12

Kopeikin, S. and Gwinn, C. (2000) Sub-Microarcsecond Astrometry and New Horizons in Relativistic Gravitational Physics, in *IAU Colloq. 180: Towards Models and Constants for Sub-Microarcsecond Astrometry* (eds K.J. Johnston, D.D. McCarthy, B.J. Luzum, and G.H. Kaplan), pp. 303–307. 7.7.1, 7.9.1

Kopeikin, S. and Korobkov, P. (2005) General Relativistic Theory of Light Propagation in the Field of Radiative Gravitational Multipoles. *ArXiv General Relativity and Quantum Cosmology e-prints*. 3.10.5, 7.1, 7.5.2, 7.7.1, 7.7.4, 7.9.3.1, 7.9.3.2, 7.9.4.1, 7.9.4.2, 7.9.4.2, 7.9.5.2

Kopeikin, S., Korobkov, P., and Polnarev, A. (2006) Propagation of light in the field of stationary and radiative gravitational multipoles. *Classical and Quantum Gravity*, **23**, 4299–4322, doi:10.1088/0264-9381/23/13/001. 3.10.4, 3.10.5, 3.10.5, 7.1, 7.5.2, 7.7.1, 7.7.1, 7.7.4, 7.9.3.1, 7.9.3.2, 7.9.4.1, 7.9.4.2, 7.9.4.2, 7.9.5.2

Kopeikin, S. and Mashhoon, B. (2002) Gravitomagnetic effects in the propagation of electromagnetic waves in variable gravitational fields of arbitrary-moving and spinning bodies. *Physical Review D*, **65** (6), 064 025–20. 7.1, 7.6.2.1, 7.7.1, 7.7.1, 7.7.4, 7.9.1, 7.9.3.1, 7.9.3.2, 7.9.4.2

Kopeikin, S. and Vlasov, I. (2004) Parametrized post-Newtonian theory of reference frames, multipolar expansions and equations of motion in the N-body problem. *Physics Reports*, **400**, 209–318, doi:10.1016/j.physrep.2004.08.004. 4.1, 4.1, 4.2.4, 4.4.2, 4.4.4, 4.5.1, 1, 4.5.3, 4.5.3, 5.1.1, 5.2.1, 5.2.2, 6.1.1, 6.1.3, 6.2.2, 6.3.2, 6.3.3, 6.3.4, 6.3.4, 7.9.3.1, 8.2, 9.2.2, 9.2.4

Kopeikin, S. and Vlasov, I. (2008) The Effacing Principle in the Post-Newtonian Celestial Mechanics. pp. 2475–2477, doi:10.1142/9789812834300_0437. 5.2.5.4, 6.1.1, 6.3.1, 6.3.1

Kopeikin, S.M. (1985) General Relativistic Equations of Binary Motion for Extended Bodies with Conservative Corrections and Radiation Damping. *Soviet Astronomy*, **29**, 516–524. 3.1.4, 4.5.1, 5.1.1, 5.2.5.4, 6.1.1, 6.3.1, 6.3.1, 6.3.1, 6.3.4

Kopeikin, S.M. (1989a) Asymptotic Matching of Gravitational Fields in the Solar System. *Soviet Astronomy*, **33**, 665–672. 4.1, 5.1.1, 7.7.4, 8.2, 8.5

Kopeikin, S.M. (1989b) Relativistic Frames of Reference in the Solar System. *Soviet Astronomy*, **33**, 550–555. 4.1, 5.1.1, 7.7.4

Kopeikin, S.M. (1990) Theory of Relativity in Observational Radio Astronomy. *Soviet Astronomy*, **34**, 5–10. 5.2.2, 7.1, 7.7.1, 7.7.3, 7.7.4, 7.8.3, 8.2, 5

Kopeikin, S.M. (1994) Supplementary parameters in the parameterized post-Keplerian formalism. *Astrophysical Journal Letters*, **434**, L67–L70, doi:10.1086/187576. 7.6.1.4, 7.6.1.4

Kopeikin, S.M. (1995) On possible implications of orbital parallaxes of wide orbit binary pulsars and their measurability. *Astrophysical Journal Letters*, **439**, L5–L8, doi:10.1086/187731. 7.6.1.2, 7.6.1.2

Kopeikin, S.M. (1997a) Binary pulsars as detectors of ultralow-frequency gravitational waves. *Physical Review D*, **56**, 4455–4469, doi:10.1103/PhysRevD.56.4455. 9.3

Kopeikin, S.M. (1997b) Millisecond and binary pulsars as nature's frequency standards - I. A generalized statistical model of low-frequency timing noise. *Monthly Notices of the Royal Astronomical Society*, **288**, 129–137. 7.7.2, 9.3

Kopeikin, S.M. (1997c) Propagation of light in the stationary field of multipole gravitational lens. *Journal of Mathematical Physics*, **38**, 2587–2601. 7.1, 7.3, 7.6.2.1, 7.6.2.1, 7.9.1, 7.9.3.2, 7.9.5.2

Kopeikin, S.M. (1999) Millisecond and binary pulsars as nature's frequency standards - II. The effects of low-frequency timing noise on residuals and measured parameters. *Monthly Notices of the Royal Astronomical Society*, **305**, 563–590, doi:10.1046/j.1365-8711.1999.02432.x. 9.3

Kopeikin, S.M. (2001) Testing the Relativistic Effect of the Propagation of Gravity by Very Long Baseline Interferometry. *Astrophysical Journal Letters*, **556**, L1–L5, doi:10.1086/322872. 4.1, 1.3.5.1, 1.3.5.1, 7.7.3, 7.9.1, 7.9.3.2, 7.9.4.2, 7.9.4.3, 7.9.6.2

Kopeikin, S.M. (2003a) Retardation of Gravity in Binary Pulsars, in *Radio Pulsars*, *Astronomical Society of the Pacific Conference Series*, vol. 302 (eds M. Bailes, D.J. Nice, and S.E. Thorsett), *Astronomical Society of the Pacific Conference Series*, vol. 302, pp. 111–114. 7.6.1.4, 7.6.1.4

Kopeikin, S.M. (2003b) The post-Newtonian treatment of the VLBI experiment on September 8, 2002. *Physics Letters A*, **312**, 147–157, doi:10.1016/S0375-9601(03)00613-3. 1, 7.9.6.2

Kopeikin, S.M. (2004) The speed of gravity in general relativity and theoretical interpretation of the Jovian deflection experiment. *Classical and Quantum Gravity*, **21**, 3251–3286, doi:10.1088/0264-9381/21/13/010. 2.2.1, 2.7.5, 3.8.7, 4.1, 4.2.1, 7.1, 1, 12, 7.9.4.2, 7.12, 7.9.4.3, 7.9.6.2, 9.2.1

Kopeikin, S.M. (2005) Comment on 'model-dependence of shapiro time delay and the "speed of gravity/speed of light" controversy'. *Classical and Quantum Gravity*, **22**, 5181–5186. URL `http://stacks.iop.org/0264-9381/22/i=23/a=N01`. 2.2.1, 9.2.1

Kopeikin, S.M. (2006) Gravitomagnetism and the Speed of Gravity. *International Journal of Modern Physics D*, **15**, 305–320, doi:10.1142/S0218271806007663. 7.9.6.2

Kopeikin, S.M. (2007) Relativistic Reference Frames for Astrometry and Navigation in the Solar System, in *American Institute of Physics Conference Series*, pp. 268–283, doi:10.1063/1.2710062. 7.9.3.1

Kopeikin, S.M. (2009) Post-Newtonian limitations on measurement of the PPN parameters caused by motion of gravitating bodies. *Monthly Notices of the Royal Astronomical Society*, **399**, 1539–1552, doi:10.1111/j.1365-2966.2009.15387.x. 7.1, 7.2, 1.3.5.1, 7.9.5.2, 7.9.6, 7.9.6.1

Kopeikin, S.M. and Fomalont, E.B. (2006) Aberration and the Fundamental Speed of Gravity in the Jovian Deflection Experiment. *Foundations of Physics*, **36**, 1244–1285, doi:10.1007/s10701-006-9059-7. 2.7.5, 63, 83, 1, 7.2, 1.3.5.1, 7.6.1.3, 7.9.1, 7.9.3.2, 7.9.4.2, 7.12, 7.9.4.2, 7.9.4.3, 7.9.4.3, 24

Kopeikin, S.M. and Fomalont, E.B. (2007) Gravimagnetism, causality, and aberration of gravity in the gravitational light-ray deflection experiments. *General Relativity and Gravitation*, **39**, 1583–1624, doi:10.1007/s10714-007-0483-6. 3.8.7, 7.1, 7.5.2, 7.6.1.4, 7.9.1, 7.9.4.3

Kopeikin, S.M. and Makarov, V.V. (2006) Astrometric Effects of Secular Aberration. *Astronomical Journal (USA)*, **131**, 1471–1478, doi:10.1086/500170. 2.5.8, 7.9.1, 9.4.3

Kopeikin, S.M. and Makarov, V.V. (2007) Gravitational bending of light by planetary multipoles and its measurement with microarcsecond astronomical interferometers. *Physical Review D*, **75** (6), 062 002–+, doi:10.1103/PhysRevD.75.062002. 7.1, 7.5.2, 7.9.1, 7.9.4.3, 7.9.5.1, 7.9.5.2, 7.9.6.1, 7.9.6.2, 7.9.6.3

Kopeikin, S.M. and Ni, W.T. (2008) Laser Ranging Delay in the Bimetric Theory of Gravity, in *Lasers, Clocks and Drag-Free Control: Exploration of Relativistic Gravity in Space*, *Astrophysics and Space Science Library*, vol. 349 (eds H. Dittus, C. Lämmerzahl, and S.G. Turyshev), *Astrophysics and Space Science Library*, vol. 349, pp. 209–+. 7.9.3.2, 7.9.4.2

Kopeikin, S.M. and Ozernoy, L.M. (1999) Post-Newtonian Theory for Precision Doppler Measurements of Binary Star Orbits. *Astrophysical Journal*, **523**, 771–785, doi:10.1086/307759. 2.5.9, 7.5.3

Kopeikin, S.M., Pavlis, E., Pavlis, D., Brumberg, V.A., Escapa, A., Getino, J., Gusev, A., Müller, J., Ni, W.T., and Petrova, N. (2008) Prospects in the orbital and rotational dynamics of the Moon with the advent of sub-centimeter lunar laser ranging. *Advances in Space Research*, **42**, 1378–1390, doi:10.1016/j.asr.2008.02.014. 3.8.7, 4.5.1, 7.7.1, 8.3, 9.2.4

Kopeikin, S.M., Polnarev, A.G., Schäfer, G., and Vlasov, I.Y. (2007) Gravimagnetic effect of the barycentric motion of the Sun and determination of the post-Newtonian parameter γ in the Cassini experiment. *Physics Letters A*, **367**, 276–280, doi:10.1016/j.physleta.2007.03.036. 3.8.5, 7.1, 7.8, 7.8.3, 7.9.1, 7.9.6.1

Kopeikin, S.M., Ramirez, J., Mashhoon, B., and Sazhin, M.V. (2001) Cosmological perturbations: a new gauge-invariant approach. *Physics Letters A*, **292**, 173–180. 4.1, 4.3.1

Kopeikin, S.M. and Schäfer, G. (1999) Lorentz covariant theory of light propagation in gravitational fields of arbitrary-moving bodies. *Physical Review D*, **60** (12), 124 002–+. 7.1, 7.2, 7.3, 7.3, 7.5.1, 7.6.1.1, 7.6.1.2, 7.6.2, 7.6.2.2, 7.7.1, 7.7.1, 7.7.3, 7.7.4, 7.7.4, 7.8.3, 7.9.1, 7.9.4.2, 7.9.5.2, 7.9.6.1

Kopeikin, S.M., Schäfer, G., Gwinn, C.R., and Eubanks, T.M. (1999) Astrometric and timing effects of gravitational waves from localized sources. *Physical Review D*, **59** (8), 084 023. 3.10.5, 3.10.5, 7.1, 7.3, 7.3, 7.3, 7.5.2, 7.5.2, 7.5.2, 7.6.2.1, 7.6.2.1, 7.7.1, 7.7.3, 7.7.4, 7.9.1, 7.9.3.1, 7.9.4.1, 7.9.4.2, 7.9.4.2, 7.9.4.2, 5

Kopeikin, S.M. (2009) On the two approaches to the data analysis of the cassini interplanetary relativity experiment. *Physics Letters A*, **373** (31), 2605 – 2607. 3.8.5, 7.1, 7.9.1

Kopejkin, S.M. (1988a) Celestial coordinate reference systems in curved spacetime. *Celestial Mechanics*, **44**, 87–115, doi:10.1007/BF01230709. 1.2.3, 3.8.5, 4.1, 4.1, 4.3.1, 4.5.1, 4.5.2.3, 5.1.1, 5.2.1, 5.2.2, 5.2.5.1, 5.2.5.1, 5.2.5.2, 5.2.5.4, 5.2.6, 6.1.1, 6.1.3, 6.1.4, 7.7.4, 7.9.3.1, 8.3

Kopejkin, S.M. (1988b) On the method of solution of the external and internal task in the problem of motion of bodies in general relativity. *Trudy Gosudarstvennogo Astronomicheskogo Instituta*, **59**, 53–71. 5.2.5.4, 6.1.1, 6.3.1, 6.3.2, 6.3.4, 6.3.4

Kopejkin, S.M. (1991a) Relativistic Manifestations of gravitational fields in gravimetry and geodesy. *Manuscripta Geodaetica*, **16**, 301–312. 4.1, 4.1, 5.2.5.1, 7.7.4, 8.1, 8.4, 8.5, 8.7, 8.8.2, 9.2.4

Kopejkin, S.M. (1991b) Relativistic reference systems in the solar system. *Itogi Nauki i Tekhniki Seriia Astronomiia*, **41**, 87–146. 4.1, 4.5.1, 5.1.1, 5.2.1, 5.2.3.1, 5.2.3.1, 5.2.5.1, 6.3.2, 8.2, 8.2, 8.4, 8.4, 8.5, 8.5, 8.6, 8.7

Kostelecky, A. (2008) Theory and Status of Lorentz Violation. *APS Meeting Abstracts*, p. I3.00001. 3.2.1, 4.2.1

Kostelecký, V.A. and Mewes, M. (2007) Lorentz-Violating Electrodynamics and the Cosmic Microwave Background. *Physical Review Letters*, **99** (1), 011 601, doi:10.1103/PhysRevLett.99.011601. 2.2.3, 3.2.1

Koszul, J.L. (1950) Homologie et cohomologie des algèbres de Lie. *Bull. Soc. Math. France*, **78**, 65–127. 3.4

Kovacs, S.J. and Thorne, K.S. (1977) The generation of gravitational waves. III - Derivation of bremsstrahlung formulae. *Astrophysical Journal*, **217**, 252–280. 7.2

Kovacs, Jr., S.J. and Thorne, K.S. (1978) The generation of gravitational waves. IV - Bremsstrahlung. *Astrophysical Journal*, **224**, 62–85, doi:10.1086/156350. 7.2

Kovalevsky, J. (1967) *Introduction to Celestial Mechanics*, Dordrecht: Reidel. 1.4.8

Kovalevsky, J. (1998) HIPPARCOS and Astrometric Results. *Highlights of Astronomy*, **11**, 536–+. 7.7.4

Kovalevsky, J., Mueller, I.I., and Kolaczek, B. (eds) (1989) *Reference frames in astronomy and geophysics*. 4.3.1, 5.1.1

Kovalevsky, J. and Seidelmann, P.K. (2004) *Fundamentals of Astrometry*, Cambridge: Cambridge University Press. 2.5.8, 2.5.8, 21, 7.1

Kox, A.J. and Eisenstaedt, J. (eds) (2005) *Einstein Studies. The Universe of General Relativity*, Boston: Birkhäuser. 3.8.1

Kraichnan, R.H. (1955) Special-relativistic derivation of generally covariant gravitation theory. *Physical Review D*, **98**, 1118–1122. 3.2.4

Kramer, M. (1998) Determination of the Geometry of the PSR B1913+16 System by Geodetic Precession. *Astrophysical Journal*, **509**, 856–860, doi:10.1086/306535. 7.6.1.1

Kramer, M., Stairs, I.H., Manchester, R.N., McLaughlin, M.A., Lyne, A.G., Ferdman, R.D., Burgay, M., Lorimer, D.R., Possenti, A., D'Amico, N., Sarkissian, J.M., Hobbs, G.B., Reynolds, J.E., Freire, P.C.C., and Camilo, F. (2006) Tests of General Relativity from Timing the Double Pulsar. *Science*, **314**, 97–102, doi:10.1126/science.1132305. 6.4.1, 7.6.1.1, 7.6.1.4

Kramer, M. and Wex, N. (2009) TOPICAL REVIEW: The double pulsar system: a unique laboratory for gravity. *Classical and Quantum Gravity*, **26** (7), 073 001–+, doi:10.1088/0264-9381/26/7/073001. 7.6.1.1, 7.6.1.4, 12, 9.3

Kretschmann, E. (1917) Über den physikalischen Sinn der Relativitätspostulat, A Einsteins neue und seine ursprünglische Relativitätstheorie. *Annalen der Physik (Leipzig)*, **53**, 575–614. 3.2.3, 3.8.1

Krisher, T.P., Morabito, D.D., and Anderson, J.D. (1993a) Erratum: "The Galileo solar redshift experiment" [Phys. Rev. Lett. 70, 2213 (1993)]. *Physical Review Letters*, **70**, 3524–+, doi:10.1103/PhysRevLett.70.3524.2. 7.7.1

Krisher, T.P., Morabito, D.D., and Anderson, J.D. (1993b) The Galileo solar redshift experiment. *Physical Review Letters*, **70**, 2213–2216. 7.7.1

Krivov, A.V. (1994) Aberration of light: Relativistic theory and Its applications in cosmic microsecond astrometry. *Astronomy Reports*, **38**, 691–698. 7.7.4

Krori, K.D. and Nandy, D. (1977) Birkhoff's theorem and scalar-tensor theories of gravitation. *Journal of Physics A Mathematical General*, **10**, 993–996, doi:10.1088/0305-4470/10/6/018. 6.3.1

Kundt, W. and Hoffmann, B. (1962) *Determination of gravitational standard time*, Oxford: Pergamon Press, pp. 303–306. 3.8.6.1

Lagrange, J.L. (1778) Sur le Problème de la détermination des orbites des comètes d'aprés trois observations. Premier et Deuxiéme mémoires. *Nouveaux mémoires de l'Académie royale des sciences et* *belles-lettres de Berlin, années 1778*. URL http://math-doc.ujf-grenoble.fr/cgi-bin/oeitem?id=OE_LAGRANGE__4_439_0, Later edition: *Œuvres de Lagrange.* Tome IV, pp. 439-496. Paris: Gauthier-Villars, 1869. 1.4.1, 1.4.6

Lagrange, J.L. (1783) Sur le Probléme de la détermination des orbites des cométes d'aprés trois observations. Troisi.me m.moire dans lequel on donne une solution directe et générale du probléme. *Nouveaux mémoires de l'Académie royale des sciences et belles-lettres de Berlin, années 1783*. URL http://math-doc.ujf-grenoble.fr/cgi-bin/oeitem?id=OE_LAGRANGE__4_439_0, Later edition: *Œuvres de Lagrange.* Tome IV, pp. 496-532. Paris: Gauthier-Villars, 1869. 1.4.1, 1.4.6

Lagrange, J.L. (1788a) *Mécanique analytique. Tome premier*, Paris: Académie des Sciences. URL http://math-doc.ujf-grenoble.fr/cgi-bin/oetoc?id=OE_LAGRANGE__11, later edition: *Œuvres de Lagrange*, Tome XI, pp. 1-497. Paris: Gauthier-Villars, 1888. 1.4.1, 1.4.6

Lagrange, J.L. (1788b) *Mécanique analytique. Tome second.*, Paris: Académie des Sciences. URL http://math-doc.ujf-grenoble.fr/cgi-bin/oetoc?id=OE_LAGRANGE__12, later edition: *Œuvres de Lagrange*, Tome XII, pp. 1-385. Paris: Gauthier-Villars, 1889. 1.4.1, 1.4.6

Lagrange, J.L. (1808a) Mémoire sur la théorie des variations des éléments des planétes, et en particulier des variations des grands axes de leurs orbites. *Mémoires de la premiére classe de l'Institut de France, année 1808*. URL http://math-doc.ujf-grenoble.fr/cgi-bin/oetoc?id=OE_LAGRANGE__6, lu, le 22 août 1808 à l'Institut de France. Later edition: *Œuvres de Lagrange.* Tome VI, pp. 713-768. Paris: Gauthier-Villars, 1873. 1.4.1, 1.4.2, 1.4.6

Lagrange, J.L. (1808b) Mémoire sur la théorie générale de la variation des constantes arbitraires dans tous les problèmes de la mécanique. *Mémoires de la première classe de l'Institut de France, année 1808*. URL http://math-doc.ujf-grenoble.fr/cgi-bin/oetoc?id=OE_LAGRANGE__6, lu, le 13 mars 1809 à l'Institut de France. Later edition: *Œuvres de Lagrange.* Tome VI, pp. 771-805. Paris: Gauthier-Villars, 1873. 1.4.1, 1.4.2, 1.4.6

Lagrange, J.L. (1809) Second mémoire sur la théorie générale de la variation des constantes arbitraires dans tous les problémes de la mécanique, dans lequel on simplifie l'application des formules générales á ces problémes. *Mémoires de la premiére classe de l'Institut de France, année 1809*. URL http://math-doc.ujf-grenoble.fr/cgi-bin/oetoc?id=OE_LAGRANGE__6, lu, le 19 février 1810 à l'Institut de France. Later edition: *Œuvres de Lagrange.* Tome VI, pp. 809 - 816. Paris: Gauthier-Villars, 1873. 1.4.1, 1.4.2, 1.4.6

Lambert, S. and Bizouard, C. (2002) Positioning the Terrestrial Ephemeris Origin in the International Terrestrial Reference Frame. *Astronomy & Astrophysics*, **394**, 317–321, doi:10.1051/0004-6361:20021139. 8.1, 9.7.5.5

Lambert, S.B. and Le Poncin-Lafitte, C. (2009) Determining the relativistic parameter γ using very long baseline interferometry. *Astronomy & Astrophysics*, **499**, 331–335, doi:10.1051/0004-6361/200911714. 3.8.6.3, 7.9.6.1, 7.9.6.1

Lamoreaux, S.K., Jacobs, J.P., Heckel, B.R., Raab, F.J., and Fortson, E.N. (1986) New limits on spatial anisotropy from optically-pumped ^{2}01Hg and ^{199}Hg. *Physical Review Letters*, **57**, 3125–3128, doi:10.1103/PhysRevLett.57.3125. 3.1.5

Lanczos, C. (1972) Einstein's path from special to general relativity., in *General Relativity. Papers in Honor of J. L. Synge* (ed. L. O'Raifeartaigh), Clarendon Press: Oxford, pp. 5–19. 3.8.2

Landau, L.D. and Lifshit's, E.M. (1959) *Theory of elasticity*, London: Addison-Wesley Pub. Co. 1.2.3, 2.8.3, 8.2, 8.2

Landau, L.D. and Lifshitz, E.M. (1959) *Fluid mechanics*, Oxford: Pergamon Press, 1959. 3.9.6.5

Landau, L.D. and Lifshitz, E.M. (1969) *Mechanics*, Oxford: Pergamon Press. 1.5.7, 1.5.7, 1.5.7, 4.5.2.3, 6.2.2

Landau, L.D. and Lifshitz, E.M. (1975) *The classical theory of fields*, Oxford: Pergamon

Press. 1.3.4, 1.4.8, 2.2.4.4, 22, 2.8.1, 2.8.1, 3.9.3.2, 3.9.7, 3.9.7, 3.10.4, 3.10.5, 4.2.1, 4.4.1, 4.4.2, 4.4.5, 4.4.5, 4.4.5, 5.1.1, 5.2.2, 6.3.1, 6.3.4, 6.4.1, 7.1, 7.2, 7.9.4.3, 8.4

Landau, L.D. and Lifshitz, E.M. (1984) *Electrodynamics of continuous media, 2-nd edition*, Elsevier: Butterworth-Heinemann. 4

Lanyi, G., Bagri, D., and Border, J. (2007) Angular position determination of spacecraft by radio interferometry. *Proceedings of the IEEE*, **95** (11), 2193–2201, doi:10.1109/JPROC.2007.905183. 7.5.2

Laplace, P.S. (1805) *Traité de Mécanique Céleste, Vol. 4, pt 2, Book X, chapter VII*, Courcier: Paris, an XIII (reprinted in Œuvres complétes de Laplace, Vol. 4, pp. 314-327. Gauthier-Villars: Paris, 1880). 2.7.5

Laurent, P., Abgrall, M., Clairon, A., Lemonde, P., Santarelli, G., Salomon, C., Picard, F., Sirmain, C., Massonnet, D., and Cacciapuoti, L. (2007) The space program PHARAO/ACES, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 6673, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 6673, doi:10.1117/12.733915. 7.8

Laurent, P., Lemonde, P., Simon, E., Santarelli, G., Clairon, A., Dimarcq, N., Petit, P., Audoin, C., and Salomon, C. (1998) A cold atom clock in absence of gravity. *European Physical Journal D*, **3**, 201–204. 7.8

Lawrence, C.R. (2009) The Planck Mission, in *Bulletin of the American Astronomical Society*, *Bulletin of the American Astronomical Society*, vol. 41, *Bulletin of the American Astronomical Society*, vol. 41, pp. 389–+. 7.6.2.2

Lawrence, J.D. (1972) *A Catalog of Special Plane Curves*, New York: Dover. 7.9.5.3, 7.9.5.3, 7.9.5.3

Le Poncin-Lafitte, C. and Teyssandier, P. (2005) Influence of the multipole moments of a giant planet on the propagation of light: application to GAIA, in *Journées 2004 - systèmes de référence spatio-temporels* (ed. N. Capitaine), pp. 204–209. 7.9.1

Lee, H.R. and Kalotas, T.M. (1975) Lorentz transformations from the first postulate. *The American Journal of Physics*, **43**, 434–437. URL `http://arxiv.org/abs/physics/0304087`. 2.2.2

Lee, T.D. (1981) *Particle physics and introduction to field theory*, *Contemporary Concepts in Physics*, vol. 1, Harwood Academic Publishers, Chur. Translated from the Chinese. 1.1.2

Lemaître, G. (1931) Expansion of the universe, The expanding universe. *Monthly Notices of the Royal Astronomical Society*, **91**, 490–501. 3.9.3.4

Lense, J. and Thirring, H. (1918) Über den Einfluß der Eigenrotation der Zentralkörper auf die Bewegung der Planeten und Monde nach der Einsteinschen Gravitationstheorie. *Physikalische Zeitschrift*, **19**, 156–163. 8.6

Lesovik, G.B., Lebedev, A.V., Mounutcharyan, V., and Martin, T. (2005) Detection of gravity waves by phase modulation of the light from a distant star. *Physical Review D*, **71** (12), 122 001–+, doi:10.1103/PhysRevD.71.122001. 7.1

Levi, B.G. (2003) GRACE Satellites Start to Map Earth's Gravity Field. *Physics Today*, **56** (2), 020 000–+, doi:10.1063/1.1564334. 8.1, 8.6, 8.7

Levi, L. (1967) The "Twin Paradox" Revisited. *American Journal of Physics*, **35**, 968–969, doi:10.1119/1.1973659. 2.6.3

Lewis, A. and Challinor, A. (2006) Weak gravitational lensing of the CMB. *Physics Reports*, **429**, 1–65, doi:10.1016/j.physrep.2006.03.002. 7.6.2.2

Li, W. and Ni, W. (1979) Coupled inertial and gravitational effects in the proper reference frame of an accelerated, rotating observer. *Journal of Mathematical Physics*, **20**, 1473–1480, doi:10.1063/1.524203. 8.4

Lieske, J.H. and Abalakin, V.K. (eds) (1990) *Inertial coordinate system on the sky; Proceedings of the 141st Symposium of the IAU, Leningrad, USSR, Oct. 17-21, 1989*, *IAU Symposium*, vol. 141. 7.7.4, 9.4.4

Lieske, J.H., Lederle, T., Fricke, W., and Morando, B. (1977) Expressions for the precession quantities based upon the IAU /1976/ system of astronomical constants. *Astronomy & Astrophysics*, **58**, 1–2. 9.5, 9.6.3

Lifshitz, E. (1946) On the gravitational stability of the expanding universe. *Journal of Physics*, **10**, 116–129. Russian version of this paper had been published in Zhurnal

Eksperimental'noi i Teoreticheskoi Fiziki, vol. 16, pp. 587-602 (1946). 4.1

Lifshitz, E.M. and Khalatnikov, I.M. (1963) Investigations in relativistic cosmology (Translated by J. L. Beeby). *Advances in Physics*, **12**, 185–249. 4.1

LIGO (2010) *Laser Interferometer Gravitational Wave Observatory*. URL http://www.ligo.caltech.edu/. 3.7.5, 3.10, 3.10.5, 4.1

Linde, A. (2000) Inflationary cosmology. *Physics Reports*, **333**, 575–591. 4.1

Linde, A.D. (1990) *Inflation and quantum cosmology*, New York: Academic Press. 3.9.3.4

Lindegren, L. (2009) Gaia: Astrometric Performance and Current Status of the Project. *American Astronomical Society, IAU Symposium #261. Relativity in Fundamental Astronomy: Dynamics, Reference Frames, and Data Analysis 27 April - 1 May 2009 Virginia Beach, VA, USA, #16.01; Bulletin of the American Astronomical Society, Vol. 41, p.890*, **261**, 1601–+. 4.1, 7.4, 7.7.4, 7.9.1

Lindegren, L. and Dravins, D. (2003) The fundamental definition of "radial velocity". *Astronomy & Astrophysics*, **401**, 1185–1201, doi:10.1051/0004-6361:20030181. 7.5.3

Lindegren, L., Perryman, M.A., and Loiseau, S. (1995) Global astrometric interferometer for astrophysics (GAIA), in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 2477 (ed. R.D. Reasenberg), *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 2477, pp. 91–103. 4.1, 7.4, 7.9.1

Lindegren, L. and Perryman, M.A.C. (1996) GAIA: Global astrometric interferometer for astrophysics. *Astronomy & Astrophysics Supplement*, **116**, 579–595. 4.1, 7.1, 7.4, 7.7.4, 7.9.1

Linder, E.V. (1997) Post-Newtonian Sachs-Wolfe effect. *Astronomy & Astrophysics*, **326**, 23–28. 7.5.3

LISA (2010) *Laser Interferometer Space Antenna*. URL http://lisa.jpl.nasa.gov/. 3.10.5, 4.1

Lobachevsky, N.I. (1999) Nascent non-Euclidean geometry: revisiting *geometric research on the theory of parallels*. *Quantum*, **9** (5), 20–25. Translated from the German. 2.5.3

Logunov, A.A. (1990) *Lectures in Relativity and Gravitation. A Modern Look*, Oxford: Pergamon Press. 2.2.3

Logunov, A.A. (2004) *Henri Poincaré and relativity theory*, Moscow: Nauka. 2.2.3

Logunov, A.A., Mestvirishvili, M.A., and Petrov, A.A. (2004) How were the Hilbert-Einstein equations discovered? *Physics-Uspekhi*, **174**, 607–623. 73

Lorentz, H.A. (1904) Electromagnetic phenomena in a system moving with any velocity less than that of light. *Proceedings of the Royal Netherlands Academy of Arts and Sciences*, **6**, 809–831. URL http://en.wikisource.org/wiki/Electromagnetic_phenomena. 6

Lorentz, H.A. (1909) *The Theory of Electrons.*, Leipzig: Teubner. Later edition: Dover, New York, 1952. 2.2.3

Lorentz, H.A. and Droste, J. (1937) The motion of a system of bodies under the influence of their mutual attraction, according to Einstein's theory. *Collected papers of H. A. Lorentz. Vol. 5* (The Hague: Martinus Nijhoff), **5**, 330–355. 6.3.4

Lorimer, D.R. and Kramer, M. (2004) *Handbook of Pulsar Astronomy*, Cambridge: Cambridge University Press. 1.5.4, 3.8.5, 64, 6.4.1, 7.6.1.1, 7.6.1.2, 7.6.1.2, 7.6.1.2, 7.7.2

Lovelock, D. (1971) The Einstein Tensor and Its Generalizations. *Journal of Mathematical Physics*, **12**, 498–501, doi:10.1063/1.1665613. 3.9.3.4

Lovelock, D. (1972) The 4-Dimensionality of Space and the Einstein Tensor. *Journal of Mathematical Physics*, **13**, 874–876, doi:10.1063/1.1666069. 3.8.4.2, 3.9.3.4

Lovis, C., Mayor, M., Bouchy, F., Pepe, F., Queloz, D., Santos, N.C., Udry, S., Benz, W., Bertaux, J.L., Mordasini, C., and Sivan, J.P. (2005) The HARPS search for southern extra-solar planets. III. Three Saturn-mass planets around HD 93083, HD 101930 and HD 102117. *Astronomy & Astrophysics*, **437**, 1121–1126, doi:10.1051/0004-6361:20052864. 7.5.3

Low, R.J. (1999) Speed limits in general relativity. *Classical and Quantum Gravity*, **16**, 543–549, doi:10.1088/0264-9381/16/2/016. 7.9.1, 7.9.4.3, 9.2.1

Ma, C., Arias, E.F., Eubanks, T.M., Fey, A.L., Gontier, A.M., Jacobs, C.S., Sovers, O.J., Archinal, B.A., and Charlot, P. (1998) The International Celestial Reference Frame as Realized by Very Long Baseline Interferometry. *Astronomical Journal (USA)*, **116**, 516–546, doi:10.1086/300408. 9.4, 9.4.1, 9.4.5.1, 9.5

Ma, C., Arias, F., Bianco, G., Boboltz, D., Bolotin, S., Charlot, P., Engelhardt, G., Fey, A., Gaume, R., Gontier, A.M., Hinkelmann, R., Jacobs, C., Kurdubov, S., Lambert, S., Malkin, Z., Nothnagel, A., Petrov, L., Shurikhina, E., Sokolova, J., Souchay, J., Sovers, O., Tesmer, V., Titov, O., Wang, G., Zharov, V., Bockmann, S., Gipson, J., Gordon, G., Lytvyn, S., and MacMillan, D. (2009) *The Second Realization of the International Celestial Reference Frame by Very Long Baseline Interferometry*, International Earth Rotation and Reference Systems Service, Frankfurt am Main. Technical Note DRAFT (July 2009). 9.4, 11, 9.4.1, 9.4.5.1

Ma, C. and Feissel, M. (eds) (1997) *Definition and Realization of the International Celestial Reference System by VLBI Astrometry of Extragalactic Objects*, *IERS Technical Note*, vol. 23. 9.4.1, 9.4.5.1

MacDonald, A.L. (1992) On the Marzke-Wheeler and Desloge constructions. *Foundations of Physics Letters*, **5**, 493–494, doi:10.1007/BF00690430. 7

Mach, E. (1883) *Die Mechanik in ihrer Entwicklung historisch-kritisch dargestellt*, Open Court Publishing Company, LaSalle IL. English edition: Mach, E. 1960. *The Science of Mechanics; a Critical and Historical Account of its Development*. A reprint of the English translation by Thomas H. McCormack (first published in 1906), with a new introduction by Karl Menger. 3.1.5

Macías, A., Cervantes-Cota, J.L., and Lämmerzahl, C. (2001) *Exact solutions and scalar fields in gravity: recent developments*, Exact Solutions and Scalar Fields in Gravity: Recent Developments. 5.1.1, 9.2.4

MacMillan, D.S. (2005) Quasar Apparent Proper Motion Observed by Geodetic VLBI Networks, in *Future Directions in High Resolution Astronomy*, *Astronomical Society of the Pacific Conference Series*, vol. 340 (ed. J. Romney & M. Reid), *Astronomical Society of the Pacific Conference Series*, vol. 340, pp. 477–+. 12

Makarov, V.V. and Kaplan, G.H. (2005) Statistical Constraints for Astrometric Binaries with Nonlinear Motion. *Astronomical Journal*, **129**, 2420–2427. 9.4.3

Maleki, L., Kohel, J.M., Lundblad, N.E., Prestage, J.D., Thompson, R.J., and Yu, N. (2008) Clocks and Accelerometers for Space Tests of Fundamental Physics, in *Lasers, Clocks and Drag-Free Control: Exploration of Relativistic Gravity in Space*, *Astrophysics and Space Science Library*, vol. 349 (eds H. Dittus, C. Lämmerzahl, and S.G. Turyshev), *Astrophysics and Space Science Library*, vol. 349, pp. 285–296. 8.1

Malkin, Z.M., L'vov, V.N., and Tsekmeyster, C.D. (2009) Future close encounters of planets with radiosources and the opportunities to use them for testing general relativity. *Solar System Research (Astronomicheskii Vestnik)*, **43** (4), 327–332. 7.4, 1.3.5.1, 7.9.3.2, 7.9.6, 7.9.6.2

Manasse, F.K. (1963) Distortion in the Metric of a Small Center of Gravitational Attraction due to its Proximity to a Very Large Mass. *Journal of Mathematical Physics*, **4**, 746–761. 5.2.1

Manchester, R.N. (2008) The Parkes Pulsar Timing Array Project, in *40 Years of Pulsars: Millisecond Pulsars, Magnetars and More*, *American Institute of Physics Conference Series*, vol. 983 (eds C. Bassa, Z. Wang, A. Cumming, and V.M. Kaspi), *American Institute of Physics Conference Series*, vol. 983, pp. 584–592, doi:10.1063/1.2900303. 3.10.5, 7.7.1, 9.3

Manoff, S. (2000) Auto-Parallel Equation as Euler-Lagrange's Equation over Spaces with Affine Connections and Metrics. *General Relativity and Gravitation*, **32**, 1559–1582. 28

Mantovani, F. and Kus, A. (eds) (2004) *The Role of VLBI in Astrophysics, Astrometry and Geodesy*. 7.7.3

Mao, Y., Tegmark, M., Guth, A.H., and Cabi, S. (2007) Constraining torsion with Gravity Probe B. *Physical Review D*, **76** (10), 104 029–+, doi:10.1103/PhysRevD.76.104029. 3.7.4, 47, 3.8.7

Marsh, G.E. and Nissim-Sabat, C. (1999) Comment on "The speed of gravity". *Physics Letters A*, **262**, 257–260, doi:10.1016/S0375-9601(99)00675-1. 2.7.5

Martin, C.F., Torrence, M.H., and Misner, C.W. (1985) Relativistic effects on an earth-orbiting satellite in the barycenter coordinate system. *Journal of Geophysical Research*, **90**, 9403–9410, doi:10.1029/JB090iB11p09403. 8.6

Martinec, Z. and Vaniček, P. (1994) The indirect effect of stokes-helmert's technique for a spherical approximation of the geoid. *Manuscripta Geodaetica*, **19**, 213–219. 8.8.1

Marzke, R.F. and Wheeler, J.A. (1964) Gravitation as Geometry- I: the Geometry of the Space-Time and the Geometrodynamical Standard Meter, in *Gravitation and Relativity* (eds H.Y. Chiu and W.F. Hoffman), New York: Benjamin, pp. 40–64. 3.8.6.1, 8.2, 9.3.1

Marzlin, K. (1994) Fermi coordinates for weak gravitational fields. *Physical Review D*, **50**, 888–891, doi:10.1103/PhysRevD.50.888. 8.4

Mashhoon, B. (1974) Can Einstein's theory of gravitation be tested beyond the geometrical optics limit? *Nature*, **250**, 316–317, doi:10.1038/250316a0. 7.1

Mashhoon, B. (1975) On tidal phenomena in a strong gravitational field. *Astrophysical Journal*, **197**, 705–716, doi:10.1086/153560. 4.5.1

Mashhoon, B. (1985) Gravitational effects of rotating masses. *Foundations of Physics*, **15**, 497–515, doi:10.1007/BF01889284. 5.1.1, 5.2.2, 5.2.5.1, 5.2, 6.1.1, 8.1

Mashhoon, B. (1992) On the strength of a gravitational field. *Physics Letters A*, **163**, 7–14. 7.2

Mashhoon, B. and Grishchuk, L.P. (1980) On the detection of a stochastic background of gravitational radiation by the Doppler tracking of spacecraft. *Astrophysical Journal*, **236**, 990–999, doi:10.1086/157825. 7.5.3

Mashhoon, B., Paik, H.J., and Will, C.M. (1989) Detection of the gravitomagnetic field using an orbiting superconducting gravity gradiometer. Theoretical principles. *Physical Review D*, **39**, 2825–2838. 8.1, 8.7, 8.7, 8.7

Mashhoon, B. and Theiss, D.S. (1982a) Errata to Relativistic Tidal Forces and the Possibility of Measuring Them. *Physical Review Letters*, **49**, 1960, doi:10.1103/PhysRevLett.49.1960.3. 8.1

Mashhoon, B. and Theiss, D.S. (1982b) Relativistic Tidal Forces and the Possibility of Measuring Them. *Physical Review Letters*, **49**, 1542–1545. 8.1

Mashhoon, B. and Theiss, D.S. (1986) Gravitational influence of the rotation of the Sun on the Earth-Moon system. *Physics Letters A*, **115**, 333–337. 8.1

Mashhoon, B. and Theiss, D.S. (1991) Relativistic lunar theory. *Nuovo Cimento B Serie*, **106**, 545–571. 8.1

Mashhoon, B. and Theiss, D.S. (2001) Relativistic Effects in the Motion of the Moon, in *Gyros, Clocks, Interferometers...:Testing Relativistic Gravity in Space* (eds C. Lämmerzahl, C.W.F. Everitt, and F.W. Hehl), Berlin: Springer, pp. 310–316. 5.2.5.1, 5.2

Mathisson, M. (1937) Neue Mechanik materieller Systeme. *Acta Phys. Polon.*, **6**, 163–200. 3.8.3, 6.1.4

Mathúna, D.Ó. (2008) *Integrable Systems in Celestial Mechanics*, Boston: Birkhäuser, doi:10.1007/978-0-8176-4595-3. 1.3.4

Mattingly, D. (2005) Modern Tests of Lorentz Invariance. *Living Reviews in Relativity*, **8**, 5–+. 7.9.4.2

Mattison, E.M. (1989) Ultra stable clocks for use in space. *Advances in Space Research*, **9**, 13–19, doi:10.1016/0273-1177(89)90003-3. 8.1

Maxwell, J.C. (1862) On Physical Lines of Force. III. The Theory of Molecular Vortices Applied to Statical Electricity. *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, Series 4*, **23**, 19–24. On page 19 of this work, the time derivatives in equations (111 - 112) give the displacement current. Maxwell believed that it is a real motion of electrical particles in a sea of æthereal vortices. Though this interpretation has long been abandoned, Maxwell's amendment to Ampère's circuital law remains valid – a varying in time electric field genarates a magnetic field. 2.1.5

Maxwell, J.C. (1865) A Dynamical Theory of the Electromagnetic Field. *Philosophical*

Transactions of the Royal Society of London, **155**, 459–512. 2.1.5

McCarthy, D.D. (2005) Summary of IAU Joint Discussion 16: 'The International Celestial Reference System, Maintenance, and Future Realizations'. *Highlights of Astronomy*, **13**, 595–+. 9.7.5.3

McCarthy, D.D. and Petit, G. (2003a) *IAU Nutation Series*. URL ftp://tai.bipm.org/iers/conv2003/chapter5/tab5.3a.txtandftp://tai.bipm.org/iers/conv2003/chapter5/tab5.3b.txt. 9.6.3, 9.6.6, 9.6.8

McCarthy, D.D. and Petit, G. (2003b) *IERS Conventions (2003)*. URL http://tai.bipm.org/iers/conv2003/conv2003.html. 9.1.3

McCarthy, D.D. and Petit, G. (2003c) *IERS Nutation Subroutine*. URL ftp://tai.bipm.org/iers/conv2003/chapter5/IAU2000A.f. 9.6.8

McCarthy, D.D. and Petit, G. (2004) *IERS Conventions (2003)*, IERS Conventions (2003). Dennis D. McCarthy and Gérard Petit (eds.), International Earth Rotation and Reference Systems Service (IERS). IERS Technical Note, No. 32, Frankfurt am Main, Germany: Verlag des Bundesamtes für Kartographie und Geodäsie, ISBN 3-89888-884-3, 2004, 127 pp. 7.7.1, 7.7.3, 7.7.3, 7.7.3, 7.7.3, 7.7.3, 7.7.3, 8.2, 8.6, 8.6, 9.1.3, 9.2.3, 9.3.2, 9.3.6.2, 9.4.6, 9.4.9, 9.5.1, 9.6.4, 9.6.7, 9.7.5.1, 9.7.5.4, 9.7.5.5, 9.7.5.6, B, B

McGuirk, J.M., Foster, G.T., Fixler, J.B., Snadden, M.J., and Kasevich, M.A. (2002) Sensitive absolute-gravity gradiometry using atom interferometry. *Physical Review A*, **65** (3), 033 608–+, doi:10.1103/PhysRevA.65.033608. 8.1

Melchior, P. (1983) *The tides of the planet earth /2nd edition/*, Oxford, Pergamon Press, 1983. 653 p. 4.5.1, 8.6

Melia, F. (2003) *The edge of infinity. Supermassive black holes in the universe*, Cambridge: Cambridge University Press. ISBN 0-521-81405-7, 2003, IX + 148 pp. 8.6

Melkumova, E.Y. and Khlebnikov, V.I. (1990) Calculation of the Shirokov effect for circular orbits in an axisymmetric static gravitational field. *Soviet Physics Journal*, **33**, 349–352, doi:10.1007/BF00894219. 8.1

Memmesheimer, R.M., Gopakumar, A., and Schäfer, G. (2004) Third post-newtonian accurate generalized quasi-keplerian parametrization for compact binaries in eccentric orbits. *Phys. Rev. D*, **70** (10), 104 011, doi:10.1103/PhysRevD.70.104011. 6.4.3.2

Memmesheimer, R.M. and Schäfer, G. (2005) Third post-Newtonian constrained canonical dynamics for binary point masses in harmonic coordinates. *Physical Review D*, **71** (4), 044 021–+, doi:10.1103/PhysRevD.71.044021. 5.2.2

Mestel, L. (2000) The Pulsar Magnetosphere. *Astrophysics & Space Sciences*, **272**, 283–294, doi:10.1023/A:1002613017285. 5.1.3

Middelberg, E. and Bach, U. (2008) High resolution radio astronomy using very long baseline interferometry. *Reports on Progress in Physics*, **71** (6), 066 901–+, doi:10.1088/0034-4885/71/6/066901. 7.7.3

Mignard, F. (1997) Astrometric Properties of the HIPPARCOS Catalogue, in *Hipparcos - Venice '97*, *ESA Special Publication*, vol. 402 (eds R.M. Bonnet, E. Høg, P.L. Bernacca, L. Emiliani, A. Blaauw, C. Turon, J. Kovalevsky, L. Lindegren, H. Hassan, M. Bouffard, B. Strim, D. Heger, M.A.C. Perryman, and L. Woltjer), *ESA Special Publication*, vol. 402, pp. 5–10. 7.7.4, 9.4.5.2

Mignard, F. (2000) Report of the IAU Working Group on ICRS, in *IAU Colloq. 180: Towards Models and Constants for Sub-Microarcsecond Astrometry* (eds K.J. Johnston, D.D. McCarthy, B.J. Luzum, and G.H. Kaplan), pp. 10–19. 4.5.1

Mignard, F. and Frœschlé, M. (2000) Global and Local Bias in the FK5 from the Hipparcos Data. *Astronomy and Astrophysics*, **354**, 732–739. 9.4.7, 9.4.9

Milani, A., Nobili, A.M., and Farinella, P. (1987) *Non-gravitational perturbations and satellite geodesy.*, Adam Hilger Ltd., Bristol, UK. 5+125 pp. 4.5.2.3, 8.8.1

Miller, A.I. (1974) On Lorentz' methodology. *The British Journal for the Philosophy of Science*, **25**, 29–45. 6

Miller, A.I. (1981) *Albert Einstein's Special Theory of Relativity*, Reading, MA: Addison-Wesley. 2.2.3

Miller, A.I. (1986) *Frontiers of physics: 1900-1911*, Boston: Birkhäuser. 6

Miller, A.I. (1987) A Précis of Edmund Whittaker's "The Relativity Theory of Poincaré and Lorentz". *Archives Internationales d'Histoire des Sciences*, **37**, 93–103. 2.2.3

Minguzzi, E. (2005) The Minkowski metric in non-inertial observer radar coordinates. *American Journal of Physics*, **73**, 1117–1121, doi:10.1119/1.2060716. 2.6.4

Minkowski, H. (1908) *Raum und Zeit*, Leipzig: B. G. Teubner Verlag. 2.2.2

Misner, C.W., Thorne, K.S., and Wheeler, J.A. (1973) *Gravitation*, San Francisco: W.H. Freeman and Co. 2.2.4.4, 2.6.2, 2.6.3, 2.8.1, 2.8.1, 3.4.4.2, 3.6.5, 3.7.1, 3.7.5, 3.7.6.1, 3.7.6.3, 42, 3.8.1, 3.8.2, 3.8.2, 49, 3.9.3.3, 3.9.4.1, 3.9.2.2, 3.9.5.4, 3.9.6.5, 3.10.4, 3.10.4, 3.10.5, 3.10.5, 4.2.1, 4.2.1, 4.2.2, 4.2.3, 4.2.3, 4.2.4, 4.3.1, 4.3.1, 4.5.1, 4.5.2.3, 5.1.1, 5.1.2, 5.1.3, 5.2.5.1, 5.2.5.3, 6.2.1, 6.3.1, 6.3.1, 6.4.1, 7.1, 7.2, 7.3, 7.5.1, 7.5.3, 7.5.3, 7.7.1, 7.9.2, 7.9.3.2, 7.9.4.3, 8.4, 8.4, 8.5, 9

Mitskevich, N.V. (1969) *Physical fields in General Relativity*, Nauka: Moscow. 33, 35, 3.9.2.1, 3.9.2.1, 74, 3.9.4.1

Moffat, J.W. (1979) New theory of gravitation. *Phys. Rev. D*, **19** (12), 3554–3558, doi:10.1103/PhysRevD.19.3554. 3.8.2

Moffat, J.W. (1995) A new nonsymmetric gravitational theory. *Physics Letters B*, **355**, 447–452, doi:10.1016/0370-2693(95)00670-G. 3.8.2, 3.8.7

Moisson, X. and Bretagnon, P. (2001) Analytical Planetary solution VSOP2000. *Celestial Mechanics and Dynamical Astronomy*, **80**, 205–213. 9.3.2

Møller, C. (1943) On homogeneous gravitational fields in the general theory of relativity and the clock paradox. *Danske Videnskab Selskab, Mat.-Fys. Medd.*, **20** (19), 1–25, See also: Møller, C. 1952. *The Theory of Relativity.* Oxford University Press, London 1952. 2.6.3

Møller, C. (1952) *The theory of relativity*, Oxford: Clarendon Press. 2.6, 22, 2.6.3

Molodensky, M.S. (1945) *Basic Questions of Geodetic Gravimetry*, Moscow: Geodezizdat (in Russian). 8.1

Montenbruck, O. and Gill, E. (2000) *Satellite Orbits. Methods, Models and Applications*, Springer: Heidelberg. 8.6, 8.6

Moody, M.V., Chan, H.A., and Paik, H.J. (1986) Superconducting gravity gradiometer for space and terrestrial applications. *Journal of Applied Physics*, **60**, 4308–4315. 8.1, 8.7

Moritz, H. (1980) *Advanced physical geodesy*, Karlsruhe : Wichmann ; Tunbridge, Eng. : Abacus Press, 1980. 6.3.1, 8.1, 8.1

Moritz, H. (1989) Equilibrium Figures in Geodesy and Geophysics., in *Geodesy and Physics of the Earth*, pp. 58–98. 6.3, 6.3.1, 6.3.1, 8.1, 8.2, 8.8.5, 8.8.5

Moritz, H. and Molodensky, M.S. (1971) *Series Solutions of Molodensky's Problem*, München: Bayerische Akademie d. Wissenschaften; Beck in Komm. 8.1, 8.8.1

Moritz, H. and Mueller, I.I. (1987) *Earth rotation: theory and observation*, New York : Ungar, 1987. 6.3.1

Moritz, H. and Yurkina, M.I. (eds) (2000) *M. S. Molodensky. In Memoriam.*, Mitteilungen der Geodätischen Institute der Technischen Universität: Graz, Vol. 88. 8.1

Moyer, T.D. (2003) *Formulation for Observed and Computed Values of Deep Space Network Data Types for Navigation*, John Wiley & Sons, Inc.: Hoboken, New Jersey. 3.8.5, 7.1, 1.3.5.1, 1.3.5.1

Mueller, I.I. and Rapp, R.H. (1989) Horizontal and vertical geodetic datums, in *Reference Frames, Astrophysics and Space Science Library*, vol. 154 (eds J. Kovalevsky, I.I. Mueller, and B. Kolaczek), *Astrophysics and Space Science Library*, vol. 154, pp. 145–162. 8.1

Mukhanov, V. (2005) *Physical Foundations of Cosmology*, Cambridge: Cambridge University Press, 2005. pp. 442, doi:10.2277/0521563984. 2.8.5, 3.9.3.4, 3.9.5.4, 3.9.5.4, 3.9.5.5, 3.10, 4.1

Müller, J., Soffel, M., and Klioner, S.A. (2008a) Geodesy and relativity. *Journal of Geodesy*, **82**, 133–145, doi:10.1007/s00190-007-0168-7. 8.1

Müller, J., Williams, J.G., and Turyshev, S.G. (2008b) Lunar Laser Ranging Contributions to Relativity and Geodesy, in *Lasers, Clocks and Drag-Free Control: Exploration of Relativistic Gravity in Space*, *Astrophysics and Space Science Library*, vol. 349 (ed. H. Dittus, C. Lämmerzahl, &

S. G. Turyshev), *Astrophysics and Space Science Library*, vol. 349, pp. 457–+. 3.1.4, 8.3, 8.6

Munk, W.H. and MacDonald, G.J.F. (1960) *The rotation of the earth; a geophysical discussion*, Cambridge [Eng.] University Press, 1960. 9.6.2

Murray, C.A. (1983) *Vectorial astrometry*, Bristol: Adam Hilger, 1983. 2.1.1, 5.2.5.1, 7.9.3.2, 7.9.5.2

Nash, J. (1956) The imbedding problem for riemannian manifolds. *The Annals of Mathematics*, **63** (1), 20–63. URL http://www.jstor.org/stable/1969989. 3.3.1

Nelson, R.A., McCarthy, D.D., Malys, S., Levine, J., Guinot, B., Fliegel, H.F., Beard, R.L., and Bartholomew, T.R. (2001) The leap second: its history and possible future. *Metrologia*, **38**, 509–529, doi:10.1088/0026-1394/38/6/6. 7, 9.3.5

Newburgh, R. (2007) Did Berkeley foreshadow Mach? *The American Journal of Physics*, **76**, 189–190. 3.1.5

Newhall, X.X., Standish, E.M., and Williams, J.G. (1983) DE 102 - A numerically integrated ephemeris of the moon and planets spanning forty-four centuries. *Astronomy & Astrophysics*, **125**, 150–167. 4.1, 8.3, 8.6

Newman, W.I. and Efroimsky, M. (2003) The method of variation of constants and multiple time scales in orbital mechanics. *Chaos*, **13**, 476–485, doi:10.1063/1.1563355. 1.5.1

Newton, I. (1760) *Philosophiae naturalis principia mathematica, vol. 1 - 4*, Colonia : A. Philibert; 548 p. ; in 8.; DCC.4.221 I through IV. 1.1.2, 44

Ni, W. (2009) Super-ASTROD: probing primordial gravitational waves and mapping the outer solar system. *Classical and Quantum Gravity*, **26** (7), 075 021–+, doi:10.1088/0264-9381/26/7/075021. 3.10.5

Ni, W.T. and Zimmermann, M. (1978a) Inertial and gravitational effects in the proper reference frame of an accelerated, rotating observer. *Physical Review D*, **17**, 1473–1476. 2.6.2

Ni, W.T. and Zimmermann, M. (1978b) Inertial and gravitational effects in the proper reference frame of an accelerated, rotating observer. *Physical Review D*, **17**, 1473–1476, doi:10.1103/PhysRevD.17.1473. 3.8.2, 4.1, 4.2.3, 4.5.1, 7.9.4.2, 8.4

Nobili, A.M., Comandi, G.L., Doravari, S., Bramanti, D., Kumar, R., Maccarrone, F., Polacco, E., Turyshev, S.G., Shao, M., Lipa, J., Dittus, H., Laemmerzhal, C., Peters, A., Mueller, J., Unnikrishnan, C.S., Roxburgh, I.W., Brillet, A., Marchal, C., Luo, J., van der Ha, J., Milyukov, V., Iafolla, V., Lucchesi, D., Tortora, P., de Bernardis, P., Palmonari, F., Focardi, S., Zanello, D., Monaco, S., Mengali, G., Anselmo, L., Iorio, L., and Knezevic, Z. (2009) "Galileo Galilei" (GG) a small satellite to test the equivalence principle of Galileo, Newton and Einstein. *Experimental Astronomy*, **23**, 689–710, doi:10.1007/s10686-008-9128-3. 8.7

Noci, G., Kohl, J.L., and Withbroe, G.L. (1987) Solar wind diagnostics from Doppler-enhanced scattering. *Astrophysical Journal*, **315**, 706–715, doi:10.1086/165172. 7.8

Noether, E. (1918) Invariante Variationsprobleme. *Nachr. d. König. Gesellsch. d. Wiss. zu Göttingen, Math-phys. Klasse*, pp. 235–257. URL http://www.physics.ucla.edu/~cwp/articles/noether.trans/german/emmy235.html. 3.9.4

Noether, E. (1971) Invariant variation problems. *Transport Theory and Statistical Physics*, **1**, 186–207, doi:10.1080/00411457108231446. 79

Noether, E. (2005) Invariant Variation Problems. *ArXiv Physics e-prints*. 79

Nordtvedt, K. (1968a) Equivalence principle for massive bodies. I. Phenomenology. *Physical Review*, **169**, 1014–1016, doi:10.1103/PhysRev.169.1014. 4.5.1, 6.1.1, 6.1.3, 6.3.4

Nordtvedt, K. (1968b) Equivalence Principle for Massive Bodies. II. Theory. *Physical Review*, **169**, 1017–1025, doi:10.1103/PhysRev.169.1017. 4.5.1, 6.1.1, 6.1.3, 6.3.4

Nordtvedt, K. (1971) Equivalence Principle for Massive Bodies. IV. Planetary Bodies and Modified Eötvös-Type Experiments. *Physical Review D*, **3**, 1683–1689, doi:10.1103/PhysRevD.3.1683. 6.3.4

Nordtvedt, K. (1973) Post-Newtonian

gravitational effects in lunar laser ranging. *Physical Review D*, **7**, 2347–2356, doi:10.1103/PhysRevD.7.2347. 6.1.5

Nordtvedt, K. (1985) A post-Newtonian gravitational Lagrangian formalism for celestial body dynamics in metric gravity. *Astrophysical Journal*, **297**, 390–404, doi:10.1086/163538. 7.1

Nordtvedt, K. (1994) Gravitational equation of motion of spherical extended bodies. *Physical Review D*, **49**, 5165–5172, doi:10.1103/PhysRevD.49.5165. 6.3.3, 6.3.4, 6.3.4

Nordtvedt, K. (1999) 30 years of lunar laser ranging and the gravitational interaction. *Classical and Quantum Gravity*, **16**, A101–A112. 7.7.1

Nordtvedt, K.J. (1970) Post-Newtonian Metric for a General Class of Scalar-Tensor Gravitational Theories and Observational Consequences. *Astrophysical Journal*, **161**, 1059–+, doi:10.1086/150607. 4.1, 4.5.1, 6.1.5

Nordtvedt, K.J. and Will, C.M. (1972) Conservation Laws and Preferred Frames in Relativistic Gravity. II. Experimental Evidence to Rule Out Preferred-Frame Theories of Gravity. *Astrophysical Journal*, **177**, 775–+. 4.5.1

Norton, J.D. (1993) General covariance and the foundations of general relativity: eight decades of dispute. *Reports on Progress in Physics*, **56**, 791–858, doi:10.1088/0034-4885/56/7/001. 3.8, 3.8.1

Novikov, I.D. and Frolov, V.P. (1989) *Physics of black holes*, (Fizika chernykh dyr, Moscow, Izdatel'stvo Nauka, 1986, 328 p) Dordrecht, Netherlands, Kluwer Academic Publishers, 1989, 351 p. Translation. Previously cited in issue 19, p. 3128, Accession no. A87-44677. 6.3.1, 6.3.1

Nowojewski, A., Kallas, J., Dragan, A., Shenitzer, A., Mykytiuk, S., and Stillwell, J. (2004) On the appearance of moving bodies. *The American Mathematical Monthly*, **111** (9), 817–821. URL http://www.jstor.org/stable/4145195. 2.5.7

Nutku, Y. (1969a) The Energy-Momentum Complex in the Brans-Dicke Theory. *Astrophysical Journal*, **158**, 991–998. 4.2.4, 4.4.5, 4.4.5

Nutku, Y. (1969b) The Post-Newtonian Equations of Hydrodynamics in the Brans-Dicke Theory. *Astrophysical Journal*, **155**, 999–+. 4.2.4

Ogiyevetsky, V.I. and Barinov, I.V. (1965) Interaction field of spin two and the Einstein equations. *Annals of Physics*, **35**, 167–208. 3.2.4

Ohanian, H.C. (2001) *Special Relativity: a modern introduction*, Lakeville, USA: Physics Curriculum & Instruction, Inc. 2.7.5, 33

Ohta, T., Okamura, H., Kimura, T., and Hiida, K. (1973) Physically Acceptable Solution of Einstein's Equation for Many-Body System. *Progress of Theoretical Physics*, **50**, 492–514. 7.7.1

Okun, L.B. (1989) The concept of mass. *Physics Today*, **42**, 31–46, doi:10.1063/1.881171. 2.7.1

Okun, L.B. (2000) Reply to the letter 'What is mass?' by R I Khrapko. *Physics Uspekhi*, **43**, 1270–1275, doi:10.1070/PU2000v043n12ABEH000889. 2.7.1

O'Raifeartaigh, L. and Straumann, N. (2000) Gauge theory: Historical origins and some modern developments. *Review of Modern Physics*, **72** (1), 1–23, doi:10.1103/RevModPhys.72.1. 3.5.3, 3.9.4.1

Osman, M.K. and Ammar, M. (2006) Variation of the orbital elements for parabolic trajectories due to a small impulse using gauss equations. *Acta Astronautica*, **59** (12), 1111 – 1116, doi:DOI:10.1016/j.actaastro.2006.03.008. URL http://www.sciencedirect.com/science/article/B6V1N-4KY9VTN-1/2/5f303594f97a170f3a9cf26fd95703e4. 15

Ostriker, J.P. and Steinhardt, P. (2001) The Quintessential Universe. *Science*, **284**, 607–623. 3.9.3.4

Overduin, J.M. and Wesson, P.S. (1997) Kaluza-Klein gravity. *Physics Reports*, **283**, 303–378, doi:10.1016/S0370-1573(96)00046-4. 2.8.5

Owen, B.J. and Sathyaprakash, B.S. (1999) Matched filtering of gravitational waves from inspiraling compact binaries: Computational cost and template placement.

Physical Review D, **60** (2), 022 002–+, doi:10.1103/PhysRevD.60.022002. 3.10.5

Paik, H.J. (1981) Superconducting tensor gravity gradiometer. *Bulletin Geodesique*, **55**, 370–381, doi:10.1007/BF02527059. 8.1, 8.7

Paik, H.J. (1989) Tests of general relativity in earth orbit using a superconducting gravity gradiometer. *Advances in Space Research*, **9**, 41–50, doi:10.1016/0273-1177(89)90006-9. 8.1

Paik, H.J. (2008) Detection of the gravitomagnetic field using an orbiting superconducting gravity gradiometer: principle and experimental considerations. *General Relativity and Gravitation*, **40**, 907–919, doi:10.1007/s10714-007-0582-4. 8.1, 8.7

Pais, A. (2005) *Subtle is the Lord*, Oxford University Press. 3.1.5

Papapetrou, A. (1951a) Equations of Motion in General Relativity . *Proceedings of the Physical Society A*, **64**, 57–75. 3.8.3, 4.2.1, 4.2.2, 4.2.4, 5.2.5.4, 6.1.1, 3, 6.1.4

Papapetrou, A. (1951b) Equations of motion in general relativity: Ii the coordinate condition. *Proceedings of the Physical Society A*, **64**, 302–310. 3

Pasquini, L., Cristiani, S., Dekker, H., Haehnelt, M., Molaro, P., Pepe, F., Avila, G., Delabre, B., D'Odorico, S., Liske, J., Shaver, P., Bonifacio, P., Borgani, S., D'Odorico, V., Vanzella, E., Bouchy, F., Dessauges, M., Lovis, C., Mayor, M., Queloz, D., Udry, S., Murphy, M., Viel, M., Grazian, A., Levshakov, S., Moscardini, L., Wiklind, T., and Zucker, S. (2006) CODEX: measuring the acceleration of the universe and beyond, in *The Scientific Requirements for Extremely Large Telescopes*, *IAU Symposium*, vol. 232 (eds P. Whitelock, M. Dennefeld, and B. Leibundgut), *IAU Symposium*, vol. 232, pp. 193–197, doi:10.1017/S1743921306000482. 7.5.3

Pati, M.E. and Will, C.M. (2000) Post-Newtonian gravitational radiation and equations of motion via direct integration of the relaxed Einstein equations: Foundations. *Physical Review D*, **62** (12), 124 015, doi:10.1103/PhysRevD.62.124015. 1.5.4, 5.2.3.1

Pauli, W. (1921) *Die Relativitätstheorie. Encyclopädie der mathematischen Wissenschaften 5.2: 556*, Verlag und Druck von B. G. Teubner, Leipzig. 2.2.2

Pauri, M. and Vallisneri, M. (2000) Märzke-Wheeler Coordinates For Accelerated Observers In Special Relativity. *Foundations of Physics Letters*, **13**, 401–425. URL http://dx.doi.org/10.1023/A:1007861914639, 10.1023/A:1007861914639. 26

Pellinen, L.P. (1987) New possibiliites for the use of gravity data in the realization of geodetic coordinate systems. *Geodeziia i Kartografiia*, pp. 10–13. 8.1

Pellinen, L.P., Neiman, I.M., and Dziaman, D.D. (1980) *Fizicheskaia geodeziia*, Moskva : Vses. in-t nauch. i tekhn. informatsii, 1980. 8.1

Penrose, R. (1968) *Structure of Space-Time*, NY: W. A. Benjamin, Inc. 2.3.3

Penrose, R. (2004) *The road to reality : a complete guide to the laws of the universe*, London: Jonathan Cape. 2.3.2, 2.3.3

Perryman, M.A.C. (2005) Surveying the Galaxy: nano-arcsec astrometry, photometry and radial velocities, in *ESA Special Publication*, *ESA Special Publication*, vol. 588 (eds F. Favata, J. Sanz-Forcada, A. Giménez, and B. Battrick), *ESA Special Publication*, vol. 588, pp. 347–+. 7.1

Perryman, M.A.C. and ESA (eds) (1997) *The HIPPARCOS and TYCHO catalogues. Astrometric and photometric star catalogues derived from the ESA HIPPARCOS Space Astrometry Mission*, *ESA Special Publication*, vol. 1200. 9.4.1, 9.4.5.2, 9.4.6, 9.4.7

Peshkin, M. and Tonomura, A. (1989) *The Aharonov-Bohm Effect*, *Lecture Notes in Physics*, vol. 340, Berlin: Springer Verlag. 57

Peterson, D. and Shao, M. (1997) The Scientific Basis for the Space Interferometry Mission, in *Hipparcos - Venice '97*, *ESA Special Publication*, vol. 402, *ESA Special Publication*, vol. 402, pp. 749–754. 7.1

Petit, G. and McCarthy, D.D. (2005) Updating the IERS Conventions to improve reference frames, in *Journées 2004 - systèmes de référence spatio-temporels* (ed. N. Capitaine), pp. 8–12. 7.7.3, 7.7.3

Petit, G. and Tavella, P. (1996) Pulsars and time scales. *Astronomy & Astrophysics*, **308**, 290–298. 9.3

Petrov, A.N. (2009) Three types of superpotentials for perturbations in the Einstein-Gauss-Bonnet gravity. *Classical and Quantum Gravity*, **26** (13), 135 010–16, doi:10.1088/0264-9381/26/13/135010. 3.9.4.3

Petrov, A.N. (2010) CORRIGENDUM: Three types of superpotentials for perturbations in the Einstein-Gauss-Bonnet gravity Three types of superpotentials for perturbations in the Einstein-Gauss-Bonnet gravity. *Classical and Quantum Gravity*, **27** (6), 069 801–2, doi:10.1088/0264-9381/27/6/069801. 3.9.4.3

Petrov, A.N. and Katz, J. (2002) Conserved currents, superpotentials and cosmological perturbations. *Royal Society of London Proceedings Series A*, **458**, 319–337. 3.9.4.3, 3.9.6.3, 3.9.7

Petrov, A.Z. (1954) The classification of spaces defining gravitational fields. *Uchenye Zapiski Kazanskogo Gosudarstvennogo Universiteta im. V. I. Ulyanovicha-Lenina [Scientific Proceedings of Kazan State University, named after V.I. Ulyanov-Lenin], 114, (8) 55-69 (1954)). Jubilee (1804-1954) Collection.*, **114**, 55–69. 3.7.6.2

Petrov, A.Z. (1969) *Einstein spaces*, Oxford: Pergamon Press, 1969. 3.7.6.2

Petrov, L., Kovalev, Y.Y., Fomalont, E.B., and Gordon, D. (2008) The Sixth VLBA Calibrator Survey: VCS6. *Astronomical Journal (USA)*, **136**, 580–585, doi:10.1088/0004-6256/136/2/580. 9.4.8

Petrova, N.M. (1949) On equations of motion and energy-momentum tensor of matter for a system of extended bodies in general theory of relativity (Ph.D. thesis). *Journal of Experimental and Theoretical Physics (JETP)*, **19**, 989–999. 6.1.1, 6.3.4

Pierre Auger Collaboration (2007) Correlation of the Highest-Energy Cosmic Rays with Nearby Extragalactic Objects. *Science*, **318** (5852), 938–943, doi:10.1126/science.1151124. URL http://www.sciencemag.org/cgi/content/abstract/318/5852/938. 3.2.1

Pijpers, F.P. (1998) Helioseismic determination of the solar gravitational quadrupole moment. *Monthly Notices of the Royal Astronomical Society*, **297**, L76–L80, doi:10.1046/j.1365-8711.1998.01801.x. 7.9.1

Pirani, F.A.E. (1965) *Introduction to Gravitational Radiation Theory*, Lectures on General Relativity, pp. 249–374. 1.2.2, 1.2.2, 3.10.3, 4.4.1

Pireaux, S., Standish, E.M., Pitjeva, E.V., and Rozelot, J.P. (2007) Solar quadrupole moment from planetary ephemerides: present state of the art. *Highlights of Astronomy*, **14**, 473–473, doi:10.1017/S1743921307011489. 7.9.1, 7.9.6.3

Pitjeva, E.V. (1993) Experimental testing of relativistic effects, variability of the gravitational constant and topography of Mercury surface from radar observations 1964-1989. *Celestial Mechanics and Dynamical Astronomy*, **55**, 313–321. 7.7.1

Pitjeva, E.V. (2005a) High-Precision Ephemerides of Planets EPM and Determination of Some Astronomical Constants. *Solar System Research*, **39** (3), 176–186. 7.9.1

Pitjeva, E.V. (2005b) Relativistic Effects and Solar Oblateness from Radar Observations of Planets and Spacecraft. *Astronomy Letters*, **31**, 340–349, doi:10.1134/1.1922533. 7.9.1

Plummer, H.C.K. (1918) *An introductory treatise on dynamical astronomy*, Cambridge: Cambridge University Press. Later editions: Kessinger Publishing LLC, 2007; General Books LLC, 2010. 1.3.5.1, 1.4.8

Poincaré, H. (1896) Sur une forme nouvelle des équations du problème des trois corps. *Comptes Rendus l'Académie des Sciences*, **123**, 1031–1035. 1.5.6

Poincare, H. (1897) Sur une forme Nouvelle des Equations du Probleme des Trois Corps. *Bulletin Astronomique, Serie I*, **14**, 53–67. 1.5.6

Poincaré, H. (1900) La théorie de Lorentz et le principe de réaction. *Archives Néerlandaises des Sciences Exactes et Naturelles*, **5**, 252–278. URL http://en.wikisource.org/wiki/Wikisource:Relativity. 2.2.3

Poincaré, H. (1904) L'état actuel et l'avenir de la Physique mathématique. *Bull. Sci. Math.*, **28**, 302–324. URL http://en.wikisource.org/wiki/The_Principles_of_Mathematical_Physics, lecture delivered on September

24, 1904. English translation in: The Foundations of Science (The Value of Science), 1913, New York: Science Press, Chap. 7-9, pp. 297-320. Another translation: The Principles of Mathematical Physics, Congress of arts and science, universal exposition, St. Louis 1904, (1905), vol. 1, pp. 604-622. 2.2.3

Poincaré, H. (1905) Sur la dynamique de l'electron. *Comptes Rendus*, **140**, 1504–1508, English translaton: http://arxiv.org/abs/gr–qc/0210 005, pp. 241 – 253. URL `http://www.soso.ch/wissen/hist/SRT/P-1905-1.pdf`. 2.2.3

Poincaré, H. (1906) Sur la dynamique de l'electron. *Rendiconti del Circolo Matematico di Palermo*, **21**, 129–176. URL `http://www.univ-nancy2.fr/poincare/bhp/`. 2.7.2

Popova, A.D. and Petrov, A.N. (1988) Exact Dynamic Theories on a Given Background in Gravitation. *International Journal of Modern Physics A*, **3**, 2651–2679, doi:10.1142/S0217751X88001107. 3.2.4, 33, 3.9.1, 3.9.4.1, 3.9.4.3, 3.9.7

Porter, E.K. (2009) An Overview of LISA Data Analysis Algorithms. *GW Notes, no. 1, p. 4*, **1**, 4–+. 3.10.5

Portilla, J.G. and Villareal, F. (2004) An Addendum to the Damour and Deruelle's Solution of the Post-Newtonian Two-Body Problem. *Celestial Mechanics and Dynamical Astronomy*, **89**, 365–381, doi: 10.1023/B:CELE.0000043574.68645.92. 6.4.3.2

Possenti, A., Burgay, M., D'Amico, N., Lyne, A.G., Kramer, M., Manchester, R.N., Camilo, F., McLaughlin, M.A., Lorimer, D., Joshi, B.C., Sarkissian, J.M., and Freire, P.C.C. (2004) The double-pulsar PSR J0737-3039A/B. *Memorie della Societa Astronomica Italiana Supplement*, **5**, 142–+. 6.4.1, 7.6.1.1

Pound, R.V. and Rebka, G.A. (1960) Variation with Temperature of the Energy of Recoil-Free Gamma Rays from Solids. *Physical Review Letters*, **4** (6), 274–275, doi:10.1103/PhysRevLett.4.274. 3.8.6.2

Prestage, J.D., Bollinger, J.J., Itano, W.M., and Wineland, D.J. (1985) Limits for spatial anisotropy by use of nuclear-spin-polarized ^{9}Be$^+$ ions. *Physical Review Letters*, **54**, 2387–2390, doi:10.1103/PhysRevLett.54.2387. 3.1.5

Pyne, T., Gwinn, C.R., Birkinshaw, M., Eubanks, T.M., and Matsakis, D.N. (1996) Gravitational Radiation and Very Long Baseline Interferometry. *Astrophysical Journal*, **465**, 566–577, doi:10.1086/177443. 3.10.5, 7.7.1

Racine, É. and Flanagan, É.É. (2005) Post-1-Newtonian equations of motion for systems of arbitrarily structured bodies. *Physical Review D*, **71**, 044 010. 6.1.5

Ramírez, J. and Kopeikin, S. (2002) A decoupled system of hyperbolic equations for linearized cosmological perturbations. *Physics Letters B*, **532**, 1–2. 3.10, 4.1, 4.3.1

Rappaport, N., Bertotti, B., Giampieri, G., and Anderson, J.D. (1997) Doppler Measurements of the Quadrupole Moments of Titan. *Icarus*, **126**, 313–323, doi:10.1006/icar.1996.5661. 7.8, 7.9.1

Reasenberg, R.D., Shapiro, I.I., MacNeil, P.E., Goldstein, R.B., Breidenthal, J.C., Brenkle, J.P., Cain, D.L., Kaufman, T.M., Komarek, T.A., and Zygielbaum, A.I. (1979) Viking relativity experiment - Verification of signal retardation by solar gravity. *Astrophysical Journal Letters*, **234**, L219–L221, doi:10.1086/183144. 7.7.1

Rendall, A.D. (1992) On the definition of post-Newtonian approximations. *Royal Society of London Proceedings Series A*, **438**, 341–360. 4.2.3

Richtmyer, R.D. (1982) *Principles of Advanced Mathematical Physics*, NY: Springer. 2.3.5.2, 11

Rickman, H. (2001) Transactions of the International Astronomical Union Proceedings of the Twenty-Fourth General Assembly. *Transactions of the International Astronomical Union, Series B*, **24**. 4

Ries, J.C. (2009) Relativity in Satellite Laser Ranging. *American Astronomical Society, IAU Symposium #261. Relativity in Fundamental Astronomy: Dynamics, Reference Frames, and Data Analysis 27 April - 1 May 2009 Virginia Beach, VA, USA, #14.02*, **261**, 1402–+. 4.5.2.3, 8.5, 8.6

Rindler, W. (1960a) Hyperbolic motion in curved space time. *Phys. Rev.*, **119** (6), 2082–2089, doi:10.1103/PhysRev.119.2082. 2.5.2, 2.6.1

Rindler, W. (1960b) *Special relativity*, University Mathematical Texts, Oliver and

Boyd: Edinburgh-London. 2.6.3, 2.6.3, 2.7.6, 2.7.6

Rindler, W. (1994) The Lense-Thirring effect exposed as anti-Machian. *Physical Letters A*, **187**, 236–238. 3.1.5

Rosenfeld, L. (1940) Sur le tenseur d'impulsion-énergie. *Acad. Roy. Belgique. Cl. Sci. Mém. Coll. in* 8°., **18** (fasc. 6), 30. 3.9.4.3, 3.9.6.2

Sachs, R. (1961) Gravitational Waves in General Relativity. VI. The Outgoing Radiation Condition. *Royal Society of London Proceedings Series A*, **264**, 309–338. 4.4.1

Sahni, V. and Starobinsky, A. (2000) The Case for a Positive Cosmological Λ-Term. *International Journal of Modern Physics D*, **9**, 373–443, doi:10.1142/S0218271800000542. 4.1

Sakharov, A. (2000) Vacuum Quantum Fluctuations in Curved Space and the Theory of Gravitation. *General Relativity and Gravitation*, **32**, 365–367. 3.9.3.4

Sakharov, A.D. (1967) Vacuum quantum fluctuations in curved space and the theory of gravitation. *Akademiia Nauk SSSR Doklady*, **177**, 70–+. 3.9.3.4

Sanso', F. (1981) Recent Advances in the Theory of the Geodetic Boundary Value Problem. *Reviews of Geophysics*, **19**, 437–449, doi:10.1029/RG019i003p00437. 8.1

Santiago, D.I. and Silbergleit, A.S. (2000) the Energy-Momentum Tensor of the Scalar Field in Scalar-Tensor Theories of Gravity. *General Relativity and Gravitation*, **32**, 565–582. 2.8.5

Sathyaprakash, B.S. and Schutz, B.F. (2009) Physics, Astrophysics and Cosmology with Gravitational Waves. *Living Reviews in Relativity*, **12** (2). 3.10.5

Satō, I. (1998) A New Anomaly of Keplerian Motion. *Astronomical Journal (USA)*, **116**, 3038–3039, doi:10.1086/300649. 1.3.5.2, 1.3.5.2, 1.3.6

Sauer, T. (2010) Einstein's Unified Field Theory Program, in *The Cambridge Companion to Einstein* (eds M. Janssen and C. Lehner), Cambridge: Cambridge University Press, pp. 1–34. URL http://philsci-archive.pitt.edu/archive/00003293/01/uft.pdf, in press. 3.5.2

Sazhin, M.V. (1978) Opportunities for detecting ultralong gravitational waves. *Soviet Astronomy*, **22**, 36–38. 9.3

Schaefer, G. (1983) On often used gauge transformations in gravitational radiation-reaction calculations. *Nuovo Cimento Lettere*, **36**, 105–108. 3.10.5

Schäfer, B.M. and Bartelmann, M. (2006) Weak lensing in the second post-Newtonian approximation: gravitomagnetic potentials and the integrated Sachs-Wolfe effect. *Monthly Notices of the Royal Astronomical Society*, **369**, 425–440, doi:10.1111/j.1365-2966.2006.10316.x. 7.9.6.2

Schäfer, G. (1982) The Equations of Motion for an Astrophysical Binary with Accuracy $1/c^5$. *Progress of Theoretical Physics*, **68**, 2191–2193, doi:10.1143/PTP.68.2191. 3.8.7, 3.10.5, 3.10.5

Schäfer, G. (1985) The gravitational quadrupole radiation-reaction force and the canonical formalism of ADM. *Annals of Physics*, **161**, 81–100, doi:10.1016/0003-4916(85)90337-9. 3.10.5

Schäfer, G. (1985) The gravitational quadrupole radiation-reaction force and the canonical formalism of ADM. *Annals of Physics*, **161**, 81–100. 5.2.2, 7.6.1.1

Schäfer, G. (1990) Reduced Hamiltonian formalism for general-relativistic adiabatic fluids and applications. *Astronomische Nachrichten*, **311**, 213–217. 7.2

Schäfer, G. (2000) Testing General Relativity, in *Fundamental Physics in Space* (ed. S. Vitale), pp. 1115–1124. 4.1, 4.2.3, 4.2.5

Schäfer, G. and Wex, N. (1993a) Second post-Newtonian motion of compact binaries. *Physics Letters A*, **174**, 196–205, doi:10.1016/0375-9601(93)90758-R. 6.4.3.2

Schäfer, G. and Wex, N. (1993b) Second post-Newtonian motion of compact binaries. *Physics Letters A*, **177**, 461–461, doi:10.1016/0375-9601(93)90980-E. 6.4.3.2

Schaub, H. and Junkins, J.L. (2003) *Analytical Mechanics of Space Systems*, American Institute of Aeronautics and Astronautics, Inc., Reston, Virginia. 1.4.3

Schiff, L.I. (1960) Possible New Experimental Test of General Relativity Theory. *Physical Review Letters*, **4**, 215–217, doi:10.1103/PhysRevLett.4.215. 3.8.3

Schlamminger, S., Choi, K., Wagner, T.A., Gundlach, J.H., and Adelberger, E.G. (2008) Test of the Equivalence Principle Using a Rotating Torsion Balance. *Physical Review Letters*, **100** (4), 041 101–+, doi:10.1103/PhysRevLett.100.041101. 3.1.2, 8.7

Schluessel, E. (2008) Constraining the parameters of binary systems through time-dependent light deflection. *General Relativity and Gravitation*, **40**, 1013–1027, doi:10.1007/s10714-007-0588-y. 7.1

Schlüter, W. and Behrend, D. (2007) The International VLBI Service for Geodesy and Astrometry (IVS): current capabilities and future prospects. *Journal of Geodesy*, **81**, 379–387, doi:10.1007/s00190-006-0131-z. 7.7.3, 8.1

Schlüter, W. and Vandenberg, N.R. (2003) Geodetic VLBI - Wishes and Limitations, in *Astronomical Society of the Pacific Conference Series*, *Astronomical Society of the Pacific Conference Series*, vol. 306 (ed. Y.C. Minh), *Astronomical Society of the Pacific Conference Series*, vol. 306, pp. 21–+. 7.7.3

Schneider, P., Ehlers, J., and Falco, E.E. (1992) *Gravitational Lenses*, Gravitational Lenses, XIV, 560 pp. 112 figs.. Springer-Verlag Berlin Heidelberg New York. Also Astronomy and Astrophysics Library. 7.6.2, 7.6.2.1, 7.6.2.1, 7.9.1

Schouten, J.A. (1954) *Tensor analysis for physicists*, Oxford, Clarendon Press, 1954 (2-d ed.). 2.3.4, 3.8.2, 4.3.1, 4.5.1, 4.5.2, 5.1.2, 8.4

Schröder, U.E. (1994) *Spezielle Relätivittstheorie*, Frankfurt: Verlag Harri Deutsch. 2.2.2

Schuh, H., Fellbaum, M., Campbell, J., Soffel, M., Ruder, H., and Schneider, M. (1988) On the deflection of radio signals in the gravitational field of Jupiter. *Physics Letters A*, **129**, 299–300, doi:10.1016/0375-9601(88)90336-2. 7.9.1

Schutz, B. (1995) *Geometrical Methods of Mathematical Physics*, Cambridge University Press: Cambridge. 2.1.1, 3.3.8.3, 23, 3.5.2, 3.7.3, 3.9.3.3, 3.9.3.3

Sciama, D.W. (1953) On the origin of inertia. *Monthly Notices of the Royal Astronomical Society*, **113**, 34–42. 3.1.5

Sciama, D.W. (1964a) *Erratum*: The Physical Structure of General Relativity. *Reviews of Modern Physics*, **36**, 1103. 3.1.5

Sciama, D.W. (1964b) The Physical Structure of General Relativity. *Reviews of Modern Physics*, **36**, 463–469. 3.1.5

Seelig, C. (1956) *Albert Einstein: A Documentary Biography*, Staples Press: London. 2.2.1, 7.9.4.3

Seidelmann, P.K. (1982) 1980 IAU theory of nutation - The final report of the IAU Working Group on Nutation. *Celestial Mechanics*, **27**, 79–106, doi:10.1007/BF01228952. 9.6.3

Seidelmann, P.K., Archinal, B.A., A'Hearn, M.F., Conrad, A., Consolmagno, G.J., Hestroffer, D., Hilton, J.L., Krasinsky, G.A., Neumann, G., Oberst, J., Stooke, P., Tedesco, E.F., Tholen, D.J., Thomas, P.C., and Williams, I.P. (2007) Report of the IAU/IAG Working Group on cartographic coordinates and rotational elements: 2006. *Celestial Mechanics and Dynamical Astronomy*, **98**, 155–180, doi:10.1007/s10569-007-9072-y. 9.5.4

Seidelmann, P.K., Archinal, B.A., A'Hearn, M.F., Cruikshank, D.P., Hilton, J.L., Keller, H.U., Oberst, J., Simon, J.L., Stooke, P., Tholen, D.J., and Thomas, P.C. (2006) *IAU/IAG Working Group (WG) on Cartographic Coordinates and Rotational Elements*. URL http://astrogeology.usgs.gov/Projects/WGCCRE/. 9.5.4

Seidelmann, P.K. and Fukushima, T. (1992) Why new time scales? *Astronomy & Astrophysics*, **265**, 833–838. 9.3.1, 9.3.2, 9.1

Seidelmann, P.K. and Kovalevsky, J. (2002) Application of the new concepts and definitions (ICRS, CIP and CEO) in fundamental astronomy. *Astronomy & Astrophysics*, **392**, 341–351, doi:10.1051/0004-6361:20020931. 9.4.4

Seidelmann, P.K. and Urban, S.E. (2010) *Explanatory Supplement to the Astronomical Almanac, Third Edition*, vol. 41, Bulletin of the American Astronomical Society. 4.1, 4.3.1, 5.1.1, 6.3.4, 9.3.1, 9.4.6, 9.5.1, 9.7.5.5

Sereno, M. (2002) Gravitational lensing by spinning and radially moving lenses. *Physics Letters A*, **305**, 7–11, doi:10.1016/S0375-9601(02)01361-0. 7.9.5.2

Sereno, M. (2005) On gravitomagnetic

time-delay by extended lenses. *Monthly Notices of the Royal Astronomical Society*, **357**, 1205–1210, doi:10.1111/j.1365-2966.2005.08709.x. 7.9.6.2

Sereno, M. (2008) Aberration in gravitational lensing. *Physical Review D*, **78** (8), 083 003–+, doi:10.1103/PhysRevD.78.083003. 7.6.2, 7.9.6.2

Sereno, M. and Cardone, V.F. (2002) Gravitational lensing by spherically symmetric lenses with angular momentum. *Astronomy & Astrophysics*, **396**, 393–396, doi:10.1051/0004-6361:20021356. 7.9.5.2

Shahid-Saless, B. and Ashby, N. (1988) Relativistic effects in local inertial frames including parameterized-post-Newtonian effects. *Physical Review D*, **38**, 1645–1657, doi:10.1103/PhysRevD.38.1645. 4.1, 5.1.1, 6.1.1

Shang, K., Jian, N., Zhang, S., Shi, X., Dai, C., and Ping, J. (2009) Testing Gravitation Theories with Precise Doppler Tracking in China's Forthcoming YH-1 Mission. *American Astronomical Society, IAU Symposium #261. Relativity in Fundamental Astronomy: Dynamics, Reference Frames, and Data Analysis 27 April - 1 May 2009 Virginia Beach, VA, USA, #9.04; Bulletin of the American Astronomical Society, Vol. 41, p.884*, **261**, 904–+. 7.8

Shapiro, I.I. (1964) Fourth Test of General Relativity. *Physical Review Letters*, **13**, 789–791, doi:10.1103/PhysRevLett.13.789. 3.8.6.3, 7.5.1, 7.7.1

Shapiro, I.I. (1966) Testing General Relativity with Radar. *Physical Review*, **141**, 1219–1222, doi:10.1103/PhysRev.141.1219. 7.5.1

Shapiro, S.S., Davis, J.L., Lebach, D.E., and Gregory, J.S. (2004) Measurement of the Solar Gravitational Deflection of Radio Waves using Geodetic Very-Long-Baseline Interferometry Data, 1979 1999. *Physical Review Letters*, **92** (12), 121 101–+, doi:10.1103/PhysRevLett.92.121101. 3.8.6.3

Shields, A. (1988) Differentiable manifolds: Weyl and Whitney. *Math. Intelligencer*, **10** (2), 5–9, doi:10.1007/BF03028349. URL http://dx.doi.org/10.1007/BF03028349. 20

Shirokov, M.F. (1973) On one new effect of the Einsteinian theory of gravitation. *General Relativity and Gravitation*, **4**, 131–136, doi:10.1007/BF00762799. 8.1

Silagadze, Z. (2008) Relativity without Tears. *Acta Physica Polonica B*, **39**, 811–883. 2.2.2

SIM (2010) *Space Interferometry Mission*. URL http://planetquest.jpl.nasa.gov/SIM/index.cfm. 4.1, 7.9.1

Simon, J.L., Bretagnon, P., Chapront, J., Chapront-Touze, M., Francou, G., and Laskar, J. (1994) Numerical expressions for precession formulae and mean elements for the Moon and the planets. *Astronomy & Astrophysics*, **282**, 663–683. 9.6.6

Slabinski, V. (2005) private communication. 2005. 9.4.9

Smarr, L.L. and Blandford, R. (1976) The binary pulsar - Physical processes, possible companions, and evolutionary histories. *Astrophysical Journal*, **207**, 574–588. 7.6.1.2

Smarr, L.L., Vessot, R.F.C., Lundquist, C.A., Decher, R., and Piran, T. (1983) Gravitational waves and red shifts - A space experiment for testing relativistic gravity using multiple time-correlated radio signals. *General Relativity and Gravitation*, **15**, 129–163. 7.8

Smith, C.A., Kaplan, G.H., Hughes, J.A., Seidelmann, P.K., Yallop, B.D., and Hohenkerk, C.Y. (1989) Mean and apparent place computations in the new IAU system. I - The transformation of astrometric catalog systems to the equinox J2000.0. II - Transformation of mean star places from FK4 B1950.0 to FK5 J2000.0 using matrices in 6-space. *Astronomical Journal (USA)*, **97**, 265–279, doi:10.1086/114979. 9.6.4

Smith, D.E. and Turcotte, D.L. (eds) (1993) *Contributions of Space Geodesy to Geodynamics : Earth Dynamics*. 9.6.1

Smolin, L. (2001) *Three roads to quantum gravity*, New York: Basic Books, viii, 231 p. ISBN 0465078354. 3.9.3.4

Smolin, L. (2006) *The trouble with physics : the rise of string theory, the fall of a science, and what comes next*, Boston: Houghton Mifflin, 2006 xxiii, 392 p. ISBN 9780618551057. 2.2.4.1

Smorodinskiĭ, Y.A. and Ugarov, V.A. (1972) Methodological Notes: Two Paradoxes of the Special Theory of Relativity. *Soviet*

Physics Uspekhi, **15**, 340–346, doi:10.1070/PU1972v015n03ABEH004972. 2.5.7

Snadden, M.J., McGuirk, J.M., Bouyer, P., Haritos, K.G., and Kasevich, M.A. (1998) Measurement of the Earth's Gravity Gradient with an Atom Interferometer-Based Gravity Gradiometer. *Physical Review Letters*, **81**, 971–974, doi:10.1103/PhysRevLett.81.971. 8.1

Sneeuw, N. (2002) Validation of fast pre-mission error analysis of the GOCE gradiometry mission by a full gravity field recovery simulation. *Journal of Geodynamics*, **33**, 43–52, doi:10.1016/S0264-3707(01)00053-9. 8.1

Soffel, M., Gill, E., Ruder, H., and Schneider, M. (1988a) Relativistic theory of gravimetry and gradiometry., in *Relativistic Effects in Geodesy*, pp. 101–105. 8.1

Soffel, M., Herold, H., Ruder, H., and Schneider, M. (1988b) Relativistic theory of gravimetric measurements and definition of the geoid. *Manuscr. Geod., Vol. 13, No. 3, p. 143 - 146*, **13**, 143–146. 8.1, 8.8.1, 8.8.2

Soffel, M., Klioner, S.A., Petit, G., Wolf, P., Kopeikin, S.M., Bretagnon, P., Brumberg, V.A., Capitaine, N., Damour, T., Fukushima, T., Guinot, B., Huang, T.Y., Lindegren, L., Ma, C., Nordtvedt, K., Ries, J.C., Seidelmann, P.K., Vokrouhlický, D., Will, C.M., and Xu, C. (2003) The IAU 2000 Resolutions for Astrometry, Celestial Mechanics, and Metrology in the Relativistic Framework: Explanatory Supplement. *Astronomical Journal (USA)*, **126**, 2687–2706, doi:10.1086/378162. 4.1, 4.1, 4.2.1, 4.2.4, 4.4.5, 4.5.1, 4.5.2.3, 4.5.2.3, 4.5.3, 5.1.1, 5.2.1, 5.2.5.1, 5.2.6, 6.1.4, 6.2.2, 7.9.3.1, 8.3, 9.2.2

Soffel, M., Ruder, H., and Schneider, M. (1986) The dominant relativistic terms in the lunar theory. *Astronomy & Astrophysics*, **157** (2), 357–364. 5.2.2

Soffel, M.H. (1989) *Relativity in Astrometry, Celestial Mechanics and Geodesy*, Relativity in Astrometry, Celestial Mechanics and Geodesy, XIV, 208 pp. 32 figs.. Springer-Verlag Berlin Heidelberg New York. Also Astronomy and Astrophysics Library. 5.2.2, 6.3, 6.3.4, 6.4.1, 6.4.2, 6.4.2, 6.4.3.3, 6.4.3.4, 7.1, 1.3.5.1, 7.7.1, 22, 8.1, 8.6, 8.8.1, 8.8.2

Soffel, M.H. and Brumberg, V.A. (1991) Relativistic reference frames including time scales - Questions and answers. *Celestial Mechanics and Dynamical Astronomy*, **52**, 355–373, doi:10.1007/BF00048451. 5.2.5.2, 9.3.2

Sokolov, V.G. (2009) Holomorphy of Coordinates with Respect to Eccentricities in the Planetary Three-Body Problem. *Solar System Research*, **43**, 53–60. 1.3.6

Sotiriou, T.P. (2009) f(R) gravity, torsion and nonmetricity. *Classical and Quantum Gravity*, **26** (15), 152 001. URL http://stacks.iop.org/0264-9381/26/i=15/a=152001. 3.5.3

Sotiriou, T.P. and Liberati, S. (2007) Metric-affine f(R) theories of gravity. *Annals of Physics*, **322**, 935–966, doi:10.1016/j.aop.2006.06.002. 3.5.3, 3.7.4

Souchay, J., Loysel, B., Kinoshita, H., and Folgueira, M. (1999) Corrections and new developments in rigid earth nutation theory. III. Final tables "REN-2000" including crossed-nutation and spin-orbit coupling effects. *Astronomy & Astrophysics Supplement*, **135**, 111–131, doi:10.1051/aas:1999446. 9.6.3

Souchay, J. and R., G. (2010) *The International Celestial Reference Frame: ICRF*. URL http://hpiers.obspm.fr/webiers/. 8.6, 9.4.5.1

Sovers, O.J., Fanselow, J.L., and Jacobs, C.S. (1998) Astrometry and geodesy with radio interferometry: experiments, models, results. *Reviews of Modern Physics*, **70**, 1393–1454. 7.1, 7.7.3, 7.9.6, 8.1

Spyrou, N. (1975) The N-body problem in general relativity. *Astrophysical Journal*, **197**, 725–743, doi:10.1086/153562. 5.2.2, 2, 6.1.1, 6.3.2, 6.3.4

Spyrou, N. (1978) Relativistic equations of motion of extended bodies. *General Relativity and Gravitation*, **9**, 519–530, doi:10.1007/BF00759546. 2, 6.1.1

Stairs, I. (2005) Overview of Pulsar Tests of General Relativity, in *Binary Radio Pulsars*, *Astronomical Society of the Pacific Conference Series*, vol. 328 (eds F.A. Rasio and I.H. Stairs), *Astronomical Society of the Pacific Conference Series*, vol. 328, pp. 3–+. 7.6.1.1

Stairs, I.H., Arzoumanian, Z., Camilo, F., Lyne, A.G., Nice, D.J., Taylor, J.H., Thorsett, S.E., and Wolszczan, A. (1998) Measurement of Relativistic Orbital Decay in the PSR B1534+12 Binary System. *Astrophysical Journal*, **505**, 352–357, doi:10.1086/306151. 7.6.1.1

Stairs, I.H., Thorsett, S.E., and Arzoumanian, Z. (2004) Measurement of Gravitational Spin-Orbit Coupling in a Binary-Pulsar System. *Physical Review Letters*, **93** (14), 141 101–+, doi:10.1103/PhysRevLett.93.141101. 7.6.1.1

Standish, E.M. (1998a) JPL Planetary and Lunar Ephemerides, DE405/LE405. Interoffice Memorandum IOM 312.F-98-048, *Tech. Rep.*, Jet Propulsion Laboratory. 9.4.8, 9.5.2

Standish, E.M. (1998b) Time scales in the JPL and CfA ephemerides. *Astronomy & Astrophysics*, **336**, 381–384. 9.3.2, 9.5.2

Standish, E.M. (2004) An approximation to the errors in the planetary ephemerides of the Astronomical Almanac. *Astronomy & Astrophysics*, **417**, 1165–1171, doi:10.1051/0004-6361:20035663. 7.9.6.3, 9.5.1

Standish, E.M. and Fienga, A. (2002) Accuracy limit of modern ephemerides imposed by the uncertainties in asteroid masses. *Astronomy & Astrophysics*, **384**, 322–328, doi:10.1051/0004-6361:20011821. 9.5.3

Standish, E.M., Newhall, X.X., Williams, J.G., and Folkner, W.M. (2010) *JPL Planetary and Lunar Ephemerides on CD-ROM*. URL http://www.willbell.com/software/jpl.htm. 9.5.2

Stebbins, A. (1988) Cosmic strings and the microwave sky. I - Anisotropy from moving strings. *Astrophysical Journal*, **327**, 584–614, doi:10.1086/166218. 7.8.4

Stephani, H., Kramer, D., MacCallum, M., Hoenselaers, C., and Herlt, E. (2003) *Exact solutions of Einstein's field equations*, Cambridge monographs on mathematical physics. Cambridge, UK: Cambridge University Press. 45

Stiefel, E. (1976) Remarks on the numerical integration of near-parabolic orbits. *Celestial Mechanics*, **14**, 85–90, doi:10.1007/BF01247134. 15

Subbotin, M.F. (1936a) A New Anomaly that Includes the Eccentric, True, and Tangential Anomalies as Particular Cases. (in Russian). *Proceedings of the Academy of Sciences of the USSR*, **6**, 167–169. 1.3.6, 1.3.6, 1.3.6

Subbotin, M.F. (1936b) A New Anomaly that Includes the Eccentric, True, and Tangential Anomalies as Particular Cases. (in Russian). *Proceedings of the Academy of Sciences of the USSR*, **7**, 9–20. 1.3.6, 1.3.6, 1.3.6

Suchting, Z. (1967) Berkeley's critisims of Newton on space and motion. *Isis*, **58**, 186–197. 3.1.5

Suen, W.M. (1986) Multipole moments for stationary, non-asymptotically-flat systems in general relativity. *Physical Review D*, **34**, 3617–3632, doi:10.1103/PhysRevD.34.3617. 4.5.1, 4.5.2.3, 8.4

Synge, J.L. (1964) *Relativity: The general theory*, Amsterdam: North-Holland. 2.6, 23, 42, 3.8.6.4, 7.4, 7.5.3, 7.5.3, 8.7, 8.7

Szabados, L.B. (2009) Quasi-Local Energy-Momentum and Angular Momentum in General Relativity. *Living Reviews in Relativity*, **12** (4). URL http://www.livingreviews.org/lrr-2009-4. 3.9.1, 3.9.2.1, 3.9.4.1, 3.9.4.3, 3.9.6.1

Taylor, J.H. and Weisberg, J.M. (1989) Further experimental tests of relativistic gravity using the binary pulsar PSR 1913 + 16. *Astrophysical Journal*, **345**, 434–450, doi:10.1086/167917. 6.4.1, 6.4.3.3

Taylor, Jr., J.H. (1991) Millisecond pulsars - Nature's most stable clocks. *IEEE Proceedings*, **79**, 1054–1062. 9.3

Taylor, Jr., J.H. (1994) Binary pulsars and relativistic gravity. *Reviews of Modern Physics*, **66**, 711–719. 3.1.4, 7.6.1.1

Teerikorpi, P., Gromov, A., and Baryshev, Y. (2003) Limits on dark energy-matter interaction from the Hubble relation for two-fluid FLRW models. *Astronomy & Astrophysics*, **407**, L9–L12. 3.9.3.4

Tegmark, M., Strauss, M.A., Blanton, M.R., Abazajian, K., Dodelson, S., Sandvik, H., Wang, X., Weinberg, D.H., Zehavi, I., Bahcall, N.A., Hoyle, F., Schlegel, D., Scoccimarro, R., Vogeley, M.S., Berlind, A., Budavari, T., Connolly, A., Eisenstein, D.J., Finkbeiner, D., Frieman, J.A., Gunn, J.E., Hui, L., Jain, B., Johnston, D., Kent, S., Lin, H., Nakajima, R., Nichol, R.C., Ostriker, J.P., Pope, A., Scranton, R., Seljak, U., Sheth, R.K., Stebbins, A., Szalay, A.S., Szapudi, I., Xu, Y., Annis, J., Brinkmann, J., Burles, S., Castander, F.J., Csabai, I., Loveday, J., Doi, M., Fukugita, M., Gillespie, B., Hennessy, G., Hogg, D.W., Ivezić, Ž., Knapp, G.R., Lamb, D.Q., Lee, B.C., Lupton, R.H., McKay, T.A., Kunszt, P., Munn, J.A., O'Connell, L., Peoples, J., Pier, J.R., Richmond, M., Rockosi, C., Schneider, D.P., Stoughton, C., Tucker, D.L., vanden Berk, D.E., Yanny, B., and York, D.G. (2004) Cosmological parameters from SDSS and WMAP. *Physical Review D*, **69** (10), 103 501–+, doi:10.1103/PhysRevD.69.103501. 3.9.3.4

Teyssandier, P., Le Poncin-Lafitte, C., and Linet, B. (2008) A Universal Tool for Determining the Time Delay and the Frequency Shift of Light: Synge's World Function, in *Lasers, Clocks and Drag-Free Control: Exploration of Relativistic Gravity in Space*, *Astrophysics and Space Science Library*, vol. 349 (eds H. Dittus, C. Lämmerzahl, and S.G. Turyshev), *Astrophysics and Space Science Library*, vol. 349, pp. 153–180. 7.4

Thomas, L.H. (1926) The Motion of the Spinning Electron. *Nature*, **117**, 514, doi:10.1038/117514a0. 2.6.2

Thorne, K.S. (1969) Nonradial Pulsation of General-Relativistic Stellar Models.IV. The Weakfield Limit. *Astrophysical Journal*, **158**, 997–+, doi:10.1086/150259. 5.2.1

Thorne, K.S. (1980) Multipole expansions of gravitational radiation. *Reviews of Modern Physics*, **52**, 299–340, doi:10.1103/RevModPhys.52.299. 1.2.2, 3.10.1, 3.10.1, 3.10.1, 3.10.3, 3.10.3, 3.10.3, 3.10.4, 3.10.5, 4.1, 4.2.5, 4.4.1, 4.4.1, 4.4.2, 4.4.3, 4.4.4, 4.5.2.1, 5.2.5.3, 6.1.1, 6.3.1, 6.3.1, 6.3.1, 6.3.1, 7.1, 7.9.3.1, 7.9.3.2, A, A, A

Thorne, K.S. and Hartle, J.B. (1985) Laws of motion and precession for black holes and other bodies. *Physical Review D*, **31**, 1815–1837. 4.5.1, 4.5.2.3, 4.5.2.3, 4.5.3, 5.1.1, 5.2.1, 5.2.5.1, 5.2.5.1, 5.2.5.2, 5.2.5.4, 6.1.1, 6.1.3, 6.2.1, 7.9.1, 7.9.3.1

Thorne, K.S. and Kovacs, S.J. (1975) The generation of gravitational waves. I - Weak-field sources. *Astrophysical Journal*, **200**, 245–262. 7.2

Thorne, K.S., Price, R.H., and MacDonald, D.A. (1986) *Black holes: The membrane paradigm*, Hew Haven: Yale University Press. 8.6

Thorne, K.S. and Winstein, C.J. (1999) Human gravity-gradient noise in interferometric gravitational-wave detectors. *Physical Review D*, **60** (8), 082 001–+, doi:10.1103/PhysRevD.60.082001. 3

Titov, O. and Malkin, Z. (2009) Effect of asymmetry of the radio source distribution on the apparent proper motion kinematic analysis. *Astronomy & Astrophysics*, **506**, 1477–1485, doi:10.1051/0004-6361/200912369. 9.4.3

Titov, O.A. (2007) Apparent proper motions of radio sources from geodetic VLBI data. *Astronomy Letters*, **33**, 481–487, doi:10.1134/S1063773707070055. 9.4.3

Torge, W. (1990) Absolute gravimetry as an operational tool for geodynamics research. *Lecture Notes in Earth Sciences, Berlin Springer Verlag*, **29**, 15–28, doi:10.1007/BFb0009877. 8.1

Torge, W. (1991) *Geodesy*, Berlin: W. de Gruyter, 2nd ed. 5

Torretti, R. (1983) *Relativity and geometry.*, Oxford: Pergamon Press. 2.2.2

Touboul, P. and Bernard, A. (2003) Electrostatic accelerometers for the Equivalence Principle test in space. *Advances in Space Research*, **32**, 1357–1360, doi:10.1016/S0273-1177(03)90345-0. 8.1

Touboul, P. and Rodrigues, M. (2001) The MICROSCOPE space mission. *Classical and Quantum Gravity*, **18**, 2487–2498, doi:10.1088/0264-9381/18/13/311. 8.7

Touboul, P., Rodrigues, M., Willemenot, E., and Bernard, A. (1996) Electrostatic accelerometers for the equivalence principle test in space. *Classical and Quantum Gravity*, **13**, 67–78. 8.1

Trautman, A. (2006) Einstein-Cartan Theory, in *Encyclopedia of Mathematical Physics* (eds J.P. Francoise, G. Naber, and S. Tsou), Oxford: Elsevier, pp. 189–195. URL http://arxiv.org/PS_cache/gr-qc/pdf/0606/0606062v1.pdf. 3.5.2, 3.7.4

Treuhaft, R.N. and Lowe, S.T. (1991) A measurement of planetary relativistic deflection. *Astronomical Journal (USA)*, **102**, 1879–1888, doi:10.1086/116010. 7.7.3, 7.9.1

Trujillo, C.A. and Brown, M.E. (2003) The Caltech Wide Area Sky Survey. *Earth Moon and Planets*, **92**, 99–112, doi:

10.1023/B:MOON.0000031929.19729.a1. 7.1

Tsamparlis, M. (2010) *Special Relativity: An Introduction with 200 Problems and Solutions*, Springer: Hiedelberg, pp. 595. 2.7.6

Tulczyjew, W. (1959) Motion of multipole particles in general relativity theory. *Acta Phys. Polon.*, **18**, 393–409. 2.7.7, 3.8.3

Turneaure, J.P., Everitt, C.W.F., Parkinson, B.W., Bardas, D., Buchman, S., Debra, D.B., Dougherty, H., Gill, D., Grammer, J., Green, G.B., Gutt, G.M., Gwo, D.H., Heifetz, M., Kasdin, N.J., Keiser, G.M., Lipa, J.A., Lockhart, J.M., Mester, J.C., Muhlfelder, B., Parmley, R., Silbergleit, A.S., Sullivan, M.T., Taber, M.A., van Patten, R.A., Vassar, R., Wang, S., Xiao, Y.M., and Zhou, P. (2003) Development of the Gravity Probe B flight mission. *Advances in Space Research*, **32**, 1387–1396, doi:10.1016/S0273-1177(03)90351-6. 8.5

Turyshev, S.G. (2009a) REVIEWS OF TOPICAL PROBLEMS: Experimental tests of general relativity: recent progress and future directions. *Physics Uspekhi*, **52**, 1–27, doi:10.3367/UFNe.0179.200901a.0003. 4.1, 7.9.6

Turyshev, V.G. (2009b) Relativistic gravitational deflection of light and its impact on the modeling accuracy for the Space Interferometry Mission. *Astronomy Letters*, **35**, 215–234, doi:10.1134/S106377370904001X. 7.7.4, 7.9.1, 7.9.3.1, 7.9.3.2, 7.9.3.2

Tyapkin, A.A. (1972) Expression of the general properties of physical processes in the space-time metric of the special theory of relativity. *Soviet Physics Uspekhi*, **15** (2), 205. URL http://stacks.iop.org/0038-5670/15/i=2/a=R06. 3.2.3

U.S. Navigation Center (2010) *Global Positioning System: Serving the World*. URL http://www.gps.gov/. 7.8

Uzan, J., Deruelle, N., and Turok, N. (1998) Conservation laws and cosmological perturbations in curved universes. *Physical Review D*, **57**, 7192–7199, doi:10.1103/PhysRevD.57.7192. 3.9.6.3

Valenti, J.A., Butler, R.P., and Marcy, G.W. (1995) Determining Spectrometer Instrumental Profiles Using FTS Reference Spectra. *Publications of Astronomical Society of the Pacific*, **107**, 966–+. 7.5.3

Van Flandern, T. (1998) The speed of gravity - What the experiments say. *Physics Letters A*, **250**, 1–11, doi:10.1016/S0375-9601(98)00650-1. 2.7.5

Van Flandern, T. (1999) Reply to: "Comment on The speed of gravity". *Physics Letters A*, **262**, 261–263, doi:10.1016/S0375-9601(99)00676-3. 2.7.5

van Leeuwen, F. (ed.) (2007a) *Hipparcos, the New Reduction of the Raw Data*, *Astrophysics and Space Science Library*, vol. 350. 9.4.5.2

van Leeuwen, F. (2007b) Validation of the new Hipparcos reduction. *Astronomy and Astrophysics*, **474**, 653–664, doi:10.1051/0004-6361:20078357. 9.4.5.2

Van Patten, R.A. and Everitt, C.W.F. (1976) Possible experiment with two counter-orbiting drag-free satellites to obtain a new test of einstein's general theory of relativity and improved measurements in geodesy. *Physical Review Letters*, **36** (12), 629–632, doi:10.1103/PhysRevLett.36.629. 8.6

van Straten, W., Bailes, M., Britton, M., Kulkarni, S.R., Anderson, S.B., Manchester, R.N., and Sarkissian, J. (2001) A test of general relativity from the three-dimensional orbital geometry of a binary pulsar. *Nature*, **412**, 158–160. 7.6.1.1, 7.6.1.4

Vaniček, P. and Krakiwsky, E.J. (1986) *Geodesy, the concepts*, New York, Elsevier Science Pub. Co. 2nd ed. 8.8.1

Vaniček, P. and Martinec, Z. (1994) The stokes-helmert scheme for the evaluation of a precise geoid. *Manuscripta Geodaetica*, **19**, 119–128. 8.8.1

Vecchiato, A., Lattanzi, M.G., Bucciarelli, B., Crosta, M., de Felice, F., and Gai, M. (2003) Testing general relativity by micro-arcsecond global astrometry. *Astronomy & Astrophysics*, **399**, 337–342, doi:10.1051/0004-6361:20021785. 7.9.1

Verbiest, J.P.W., Bailes, M., Coles, W.A., Hobbs, G.B., van Straten, W., Champion, D.J., Jenet, F.A., Manchester, R.N., Bhat, N.D.R., Sarkissian, J.M., Yardley, D., Burke-Spolaor, S., Hotan, A.W., and You, X.P. (2009) Timing stability of millisecond pulsars and prospects for gravitational-wave detection. *Monthly Notices of the Royal Astronomical Society*, **400**, 951–968, doi:10.1111/j.1365-2966.2009.15508.x. 3.10, 9.3

Vincent, M.A. (1986) The relativistic equations of motion for a satellite in orbit about a finite-size, rotating earth. *Celestial Mechanics*, **39**, 15–21, doi:10.1007/BF01232285. 6.3.4

Vinti, J.P. (1998) *Orbital and Celestial Mechanics*, Progress in Astronautics and Aeronautics Series, Vol.177, American Institute of Aeronautics and Astronautics (AIAA). 1.4.8

von Ignatowski, W.A. (1910) Einige allgemeine Bemerkungen zum Relativitatsprinzip. *Physikalische Zeitschrift*, **11**, 972–976. 2.2.2

Vozmischeva, T.G. (2003) *Integrable Problems of Celestial Mechanics in Spaces of Constant Curvature*, *Astrophysics and Space Science Library*, vol. 295, Dordrecht: Kluwer Academic Publishers. 1.3.4

Wagoner, R.V. (1979) Low-frequency gravitational radiation from collapsing systems. *Physical Review D*, **19**, 2897–2901, doi:10.1103/PhysRevD.19.2897. 4.4.1

Wagoner, R.V. and Will, C.M. (1976) Post-Newtonian gravitational radiation from orbiting point masses. *Astrophysical Journal*, **210**, 764–775, doi:10.1086/154886. 6.4.3.3, 6.4.3.3

Wahr, J.M. (1981) The forced nutations of an elliptical, rotating, elastic and oceanless earth. *Geophysical Journal International*, **64**, 705–727, doi:10.1111/j.1365-246X.1981.tb02691.x. 9.6.3

Wald, R.M. (1984) *General relativity*, University of Chicago Press: Chicago. 2.2.4.4, 3.2.1, 3.6.5, 3.7.5, 4.2.1, 4.2.3, 6.3.1, 6.4.1, 7.9.4.3

Walker, M.J.H., Owens, J., and Ireland, B. (1985) A set of modified equinoctial orbit elements. *Celestial Mechanics*, **36**, 409–419, doi:10.1007/BF01227493. 1.4.6

Walter, H.G. and Sovers, O.J. (2000) *Astrometry of fundamental catalogues : the evolution from optical to radio reference frames*, Berlin: Springer-Verlag, 2000. Astronomy and Astrophysics Library. 7.9.6

Ward, W.R. (1981) Orbital inclination of Iapetus and the rotation of the Laplacian plane. *Icarus*, **46**, 97–107, doi:10.1016/0019-1035(81)90079-8. 1.5.6

Weber, J. (1961) *General relativity and gravitational waves*, New York: Interscience. 3.10.5, 3.10.5

Weinberg, S. (1965) Photons and gravitons in perturbation theory: derivations of Maxwell and Einstein equations. *Physical Review B*, **138**, 988–1002. 3.2.4

Weinberg, S. (1972) *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*, John Wiley & Sons. 2.8.3, 3.1.5, 3.7.6.2, 42, 3.8.1, 3.8.3, 3.8.6.3, 3.10.5, 4.1, 4.2.1, 4.2.3, 7.1, 7.2, 7.2, 7.5.1, 7.5.2, 8.5, 9

Weinstock, R. (1982) Dismantling a centuries-old myth: Newton's Principia and inverse-square orbits. *American Journal of Physics*, **50**, 610–617, doi:10.1119/1.12788. 5, 1.3.2

Weisberg, J.M. and Taylor, J.H. (2002) General Relativistic Geodetic Spin Precession in Binary Pulsar B1913+16: Mapping the Emission Beam in Two Dimensions. *Astrophysical Journal*, **576**, 942–949, doi:10.1086/341803. 7.6.1.1

Weitzenböck, R. (1923) *Invarianten Theorie*, Germany, Groningen: Nordhoff. 3.7.4

Wesson, P.S. (2000) *Space, time, matter. Modern Kaluza-Klein theory.*, World Scientific: Singapore. 2.8.5

Westpfahl, K. (1985) High-Speed Scattering of Charged and Uncharged Particles in General Relativity. *Fortschritte der Physik*, **33**, 417–493, doi:10.1002/prop.2190330802. 7.2

Wex, N. (1995) *Relativistische Effekte in der Beobachtung und Bewegung von Pulsaren, Ph.D. Thesis*, Ph.D. thesis, Friedrich Schiller University of Jena, Germany. 1.3.5.1, 7.6.1.1, 7.6.1.3, 7.6.1.4

Wex, N. and Kopeikin, S.M. (1999) Frame Dragging and Other Precessional Effects in Black Hole Pulsar Binaries. *Astrophysical Journal*, **514**, 388–401, doi:10.1086/306933. 1.3.5.1, 7.6.1.3

Weyl, H. (1950) *Space-Time-Matter*, Translated from the German by Henry L. Brose. First American Printing of the Fourth Edition (1922), Dover Publications, Inc., New York. 2.3.1

Weyl, H. (1964) *The concept of a Riemann surface*, Translated from the third German edition by Gerald R. MacLane. ADIWES International Series in Mathematics, Addison-Wesley Publishing Co., Inc., Reading, Mass.-London. 3.3.1

Whipple, A.L., Jefferys, W.H., Benedict, G.F., Hemenway, P.D., McArthur, B., McCartney, J., Nelan, E., Shelus, P.J., Story, D., Duncombe, R., van Altena, W.F., Franz, O.G., Wasserman, L.H., and Fredrick, L.W. (1996) Observation of the Gravitational Deflection of Light by Jupiter, in *Bulletin of the American Astronomical Society*, *Bulletin of the American Astronomical Society*, vol. 28, *Bulletin of the American Astronomical Society*, vol. 28, pp. 1187–+. 7.9.1

White, M. and Hu, W. (1997) The Sachs-Wolfe effect. *Astronomy & Astrophysics*, **321**, 8–9. 7.5.3

Whitney, H. (1936) Differentiable manifolds. *Ann. of Math. (2)*, **37** (3), 645–680, doi:10.2307/1968482. URL http://dx.doi.org/10.2307/1968482. 3.3.1

Whittaker, E.T. (1953) *A History of the Theories of Æther and Electricity, Vol.2*, London: Nelson [Later reprint: Tomash NY 1987]. 2.2.3

Wigner, E.P. (1939) On unitary representations of the inhomogeneous Lorentz group. *Annals of Mathematics*, **40**, 149–204. 2.3.6

Will, C.M. (1971) Relativistic Gravity in the Solar System. II. Anisotropy in the Newtonian Gravitational Constant. *Astrophysical Journal*, **169**, 141–155. 8.1

Will, C.M. (1993) *Theory and Experiment in Gravitational Physics*, Cambridge: Cambridge University Press. 2.2.3, 2.7.5, 2.8.2, 3.1.2, 3.2.2.2, 19, 3.5.3, 3.5.3, 46, 3.8.4.1, 3.8.5, 3.8.7, 86, 4.1, 4.1, 4.2.1, 4.2.1, 4.2.2, 4.2.3, 4.2.3, 4.2.4, 4.2.5, 4.2.5, 4.3.1, 4.4.1, 4.4.3, 4.4.4, 4.5.1, 4.5.2.3, 5.2.1, 5.2.2, 5.2.5.4, 6.1.1, 6.1.3, 6.1.3, 6.3, 6.3.2, 6.3.3, 6.3.4, 6.3.4, 6.4.1, 6.4.2, 6.4.2, 6.4.2, 6.4.3.3, 7.1, 1.3.5.1, 10, 7.7.1, 7.9.4.2, 7.9.4.3, 7.9.5.2, 7.9.6.1, 8.1, 9, 9.2.4

Will, C.M. (2003) Propagation Speed of Gravity and the Relativistic Time Delay. *Astrophysical Journal*, **590**, 683–690, doi:10.1086/375164. 1, 7.7.1, 7.7.1

Will, C.M. (2006) The Confrontation between General Relativity and Experiment. *Living Reviews in Relativity*, **9** (3). 3.5.3, 3.8.7, 4.1, 4.2.1, 6.4.2, 7.1, 7.9.1, 7.9.4.2, 7.9.4.3, 24

Williams, G.V., Helin, E.F., Pravdo, S., Lawrence, K., Rabinowitz, D., Brown, M.E., Trujillo, C.A., Margot, J.L., Hicks, M., Kowal, C.T., Jordan, A.B., Conselice, C.J., and Marsden, B.G. (2002) 2002 LM60. *Minor Planet Electronic Circulars 2002-T34*, pp. 34–+. 7.1

Williams, J.G., Newhall, X.X., and Dickey, J.O. (1996) Relativity parameters determined from lunar laser ranging. *Physical Review D*, **53**, 6730–6739. 7.7.1

Williams, J.G., Turyshev, S.G., and Boggs, D.H. (2009) Lunar Laser Ranging Tests of the Equivalence Principle with the Earth and Moon. *International Journal of Modern Physics D*, **18**, 1129–1175, doi:10.1142/S021827180901500X. 8.7

Worden, P., Torii, R., Mester, J.C., and Everitt, C.W.F. (2000) The STEP Payload and Experiment. *Advances in Space Research*, **25**, 1205–1208, doi:10.1016/S0273-1177(99)00988-6. 4.1, 8.7

Worden, P. and Everitt, C. (1974) Tests of the equivalence of gravitational and inertial mass based on cryogenic techniques, in *Proceedings of the International School of Physics "Enrico Fermi", 15+574 p.*, *Astrophysics and Space Science Library*, vol. 56 (ed. B. Bertotti), London: Academic Press, *Astrophysics and Space Science Library*, vol. 56, pp. 382–402. 8.7

Wu, N. and Ruan, T. (2003) Problems on Foundations of General Relativity. *ArXiv High Energy Physics - Theory e-prints*. 37

Wucknitz, O. and Sperhake, U. (2004) Deflection of light and particles by moving gravitational lenses. *Physical Review D*, **69** (6), 063 001–+, doi:10.1103/PhysRevD.69.063001. 7.9.5.2, 7.9.6.1

Xie, Y. and Kopeikin, S. (2010) Reference frames, gauge transformations and gravitomagnetism in the post-Newtonian theory of the lunar motion, in *IAU Symposium*, *IAU Symposium*, vol. 261 (ed. S. A. Klioner, P. K. Seidelmann, & M. H. Soffel), *IAU Symposium*, vol. 261, pp. 40–44, doi:10.1017/S1743921309990123. 3.8.5, 4.4.5, 4.5.1, 5.2.2, 9.2.4

Xu, C., Wu, X., and Schäfer, G. (1997) Binary systems with monopole, spin, and quadrupole moments. *Physical Review D*, **55**, 528–539, doi:10.1103/PhysRevD.55.528. 6.1.5, 6.3.4

Xu, C., Wu, X., and Soffel, M. (2005) General-relativistic perturbation equations for the dynamics of elastic deformable astronomical bodies expanded in terms of generalized spherical harmonics. *Physical Review D*, **71** (2), 024 030–+, doi:10.1103/PhysRevD.71.024030. 4.1, 4.2.2

Xu, C., Wu, X., Soffel, M., and Klioner, S. (2003) Relativistic theory of elastic deformable astronomical bodies: Perturbation equations in rotating spherical coordinates and junction conditions. *Physical Review D*, **68** (6), 064 009–+, doi:10.1103/PhysRevD.68.064009. 4.1, 4.2.2, 4.5.1, 6.3.1

Yakovlev, O.I. (2002) *Space Radio Science*, London: Taylor & Francis. 64, 7.3

Yallop, B.D., Hohenkerk, C.Y., Smith, C.A., Kaplan, G.H., Hughes, J.A., and Seidelmann, P.K. (1989) Mean and Apparent Place Computations in the New IAU System. II. Transformation of Mean Star Places from FK4 B1950.0 to FK5 J2000.0 Using Matrices in 6-Space. *Astronomical Journal (USA)*, **97**, 274–+, doi:10.1086/114980. 9.6.4

Yoshino, T., Takahashi, Y., and Schuh, H. (1989) Relativistic radio wave deflection VLBI experiment due to the Jupiter gravitational field., in *Earth Rotation, Astrometry, and Geodesy*, pp. 119–124. 7.9.1

Zacharias, N., McCallon, H.L., Kopan, E., and Cutri, R.M. (2005) Extending the ICRF into the infrared: 2MASS-UCAC astrometry. *Highlights of Astronomy*, **13**, 603–604. 9.4.8

Zacharias, N., Urban, S.E., Zacharias, M.I., Wycoff, G.L., Hall, D.M., Monet, D.G., and Rafferty, T.J. (2004) The Second US Naval Observatory CCD Astrograph Catalog (UCAC2). *Astronomical Journal (USA)*, **127**, 3043–3059, doi:10.1086/386353. 9.4.8

Zaglauer, H.W. (1992) Neutron stars and gravitational scalars. *Astrophysical Journal*, **393**, 685–696, doi:10.1086/171537. 7.6.1.1

Zahar, E. (1989) *Einstein's Revolution: A Study in Heuristic*, Chicago: Open Court Publishing Company. 2.2.3

Zakharov, A.F. and Sazhin, M.V. (1998) Gravitational microlensing. *Uspekhi Fizicheskikh Nauk*, **41**, 945–982. 7.6.2, 7.6.2.1, 7.6.2.1

Zeh, H. (2001) *The physical basis of the direction of time*, Berlin: Springer, 4th edition, X + 231 pp. 2.1.4

Zel′manov, A.L. (1944) *Chronometric Invariants: On Deformations and the Curvature of Accompanying Space*, American Research Press: Rehoboth, NM USA. Translated in 2006 from the Russian by Dmitri Rabounski and Stephen J. Crothers. Formulae corrected by Larissa Borissova. ISBN 1-59973-011-1.. 3.8.5, 3.10.5

Zel′manov, A.L. and Agakov, V.G. (1989) *Elements of General Theory of Relativity*, Moscow: Nauka. "Elementy obshchei teorii otnositelnosti" (in Russian with an English summary), 238 pp. 2.5.6, 22, 3.8.2, 3.8.5, 59, 3.10.5

Zhang, X. (1986) Multipole expansions of the general-relativistic gravitational field of the external universe. *Physical Review D*, **34**, 991–1004, doi:10.1103/PhysRevD.34.991. 8.4

Zharkov, V.N. and Trubitsyn, V.P. (1978) *Physics of planetary interiors*, Astronomy and Astrophysics Series, Tucson: Pachart. 4.2.2, 6.3, 6.3.1, 6.3.1

Zucker, S. and Alexander, T. (2007) Spectroscopic Binary Mass Determination Using Relativity. *Astrophysical Journal Letters*, **654**, L83–L86, doi:10.1086/510799. 7.5.3

Zwillinger, D. (1998) *Handbook of differential equations (3-d ed.)*, San Diego: Academic Press, 1998. 5.2.1

Index