

THE FRONTIERS COLLECTION

Mendel Sachs

QUANTUM MECHANICS AND GRAVITY



Springer

THE FRONTIERS COLLECTION

Springer

Berlin

Heidelberg

New York

Hong Kong

London

Milan

Paris

Tokyo

Physics and Astronomy



ONLINE LIBRARY

springeronline.com

THE FRONTIERS COLLECTION

Series Editors:

D. Dragoman M. Dragoman A.C. Elitzur M.P. Silverman J. Tuszynski H.D. Zeh

The books in this collection are devoted to challenging and open problems at the forefront of modern physics and related disciplines, including philosophical debates. In contrast to typical research monographs, however, they strive to present their topics in a manner accessible also to scientifically literate non-specialists wishing to gain insight into the deeper implications and fascinating questions involved. Taken as a whole, the series reflects the need for a fundamental and interdisciplinary approach to modern science. It is intended to encourage scientists in all areas to ponder over important and perhaps controversial issues beyond their own speciality. Extending from quantum physics and relativity to entropy, time and consciousness – the Frontiers Collection will inspire readers to push back the frontiers of their own knowledge.

Quantum Mechanics and Gravity

By M. Sachs

Mind, Matter and Quantum Mechanics

By H. Stapp

Quantum–Classical Correspondence

By A.O. Bolivar

Quantum–Classical Analogies

By D. Dragoman and M. Dragoman

Quo Vadis Quantum Mechanics?

Edited by A. C. Elitzur, S. Dolev, N. Kolenda

Mendel Sachs

QUANTUM MECHANICS AND GRAVITY



Springer

Prof. Mendel Sachs (Prof. em.)

University at Buffalo, State University of New York, Department of Physics,
Buffalo, New York 14260, USA email: msachs@buffalo.edu

Series Editors:

Prof. Daniela Dragoman

University of Bucharest, Physics Faculty, Solid State Chair, PO Box MG-11,
76900 Bucharest, Romania email: danieladragoman@yahoo.com

Prof. Mircea Dragoman

National Research and Development Institute in Microtechnology, PO Box 38-160,
023573 Bucharest, Romania email: mircead@imt.ro

Prof. Avshalom C. Elitzur

Bar-Ilan University, Unit of Interdisciplinary Studies,
52900 Ramat-Gan, Israel email: avshalom.elitzur@weizmann.ac.il

Prof. Mark P. Silverman

Department of Physics, Trinity College,
Hartford, CT 06106, USA email: mark.silverman@trincoll.edu

Prof. Jack Tuszynski

University of Alberta, Department of Physics, Edmonton, AB,
T6G 2J1, Canada email: jtus@phys.ualberta.ca

Prof. H. Dieter Zeh

University of Heidelberg, Institute of Theoretical Physics, Philosophenweg 19,
69120 Heidelberg, Germany email: zeh@urz.uni-heidelberg.de

Cover figure: Detail from a real image of the Barth sextic, a degree six algebraic surface in threespace admitting the maximum number of nodes (65). Courtesy of Stephan Endrass, Micronas GmbH

ISSN 1612-3018

ISBN 3-540-00800-4 Springer-Verlag Berlin Heidelberg New York

Library of Congress Cataloging-in-Publication Data.

Sachs, Mendel. Quantum mechanics and gravity/Mendel Sachs. p.cm. – (The frontiers collection, ISSN 1612-3018) Includes bibliographical references and index. ISBN 3-540-00800-4 (acid-free paper) 1. Quantum theory. 2. General relativity (Physics) I. Title. II. Series. QC174.12.S2246 2003 530.12-dc22 2003059074

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer-Verlag Berlin Heidelberg New York

a member of BertelsmannSpringer Science+Business Media GmbH
springeronline.com

© Springer-Verlag Berlin Heidelberg 2004 Printed in Germany

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting by Stephen Lyle using a Springer $\text{\TeX}$ macro package

Final processing by Frank Herweg, Leutershausen

Cover design by KünkelLopka, Werbeagentur GmbH, Heidelberg

Printed on acid-free paper SPIN: 10885648 57/3141/tr - 5 4 3 2 1 0

To My Grandchildren
Ariana, Yoni, Tali, Noah, Sam, Jonah and Alexandra

and to our proposition
There is More to the Real World than Meets the Eye

Preface

Theoretical physics is presently at a very exciting time in the history of scientific discovery. For we are at a precipice facing two conflicting 20th century revolutionary movements in physics, each purporting to be basic truths of nature – the quantum theory and the theory of relativity.

In the 20th century the mathematical expression of the quantum theory yielded correct predictions of a great deal of the data on the behavior of the molecular, atomic, nuclear and elementary particle domains of matter. In the same period, the theory of relativity successfully described new features of material systems. In special relativity, the relativistic Doppler effects (transverse and longitudinal) of electromagnetic radiation, and the mechanics of matter that moves at speeds close to the speed of light, revealing, for example, the energy mass relation, $E = mc^2$, revolutionized our thinking. In its form of general relativity, it has yielded a formalism that successfully predicted features of the phenomenon of gravity, also predicted by the classical Newtonian theory, but in addition, features not predicted by the classical theory, thereby superceding Newton's theory of universal gravitation.

The problem we are now faced with, in these early decades of the 21st century, is that in their precise mathematical forms and their conceptual bases, the theory of relativity and the quantum theory are both logically and mathematically incompatible. They each entail opposing paradigms on the true nature of matter and radiation, as well as opposing epistemologies. For the purpose of description of particular phenomena, and so long as the physical conditions that require the use of the quantum theory and the theory of relativity do not overlap, these theories may be expressed separately. But in general, for the purpose of explanation, we must consider the conditions where both theories would be required simultaneously to correctly represent the laws of nature.

Examples of such attempts are high energy particle physics, called relativistic quantum field theory, and the theory of general relativity, to represent the gravitational manifestations of matter.

Further, even in those cases where the empirical conditions for each of these theories do not overlap, as an explanation of the behavior of matter we cannot accept both simultaneously, if they are in principle both logically and mathematically incompatible. This would yield a logically inconsistent understanding of the nature of matter. Thus we are at a precarious precipice where we must choose one of these theories or the other, but not both, lest we fall into the mire of conceptual inconsistencies, below the edge of the cliff!

My research program, as it has developed over the past 40 years in the literature, started with the attempt to answer the following question: is it possible that the formal structure of quantum mechanics is not more than a linear approximation for a theory of matter that is based on the foundations of Einstein's theory of general relativity, as a general theory of matter? For if this were the case it would signify a genuine paradigm change in physics, from linearity, indeterminism and the fundamental role of probability and measurement, to views of holism, continuity, nonlinearity and determinism, where measurement and probability play no fundamental role. It would be a change from an atomistic model based on the epistemology of logical positivism, to a holistic view of the universe, based on the epistemology of realism. In the latter holistic ontology, the 'things' that we identify, for example, with electron, proton, DNA molecule, people, planets, galaxies, etc., are not separable, singular entities; rather, each is of the infinite distribution of real, distinguishable manifestations of the single continuous universe – of its correlated, though inseparable modes.

The basic attempt, then, in this book, is to initiate a study of a single, coherent theory of matter applicable to all domains – from elementary particle physics to cosmology. The approach taken is that of a fully exploited theory of general relativity, from its mathematical and conceptual bases.

The reader should not assume that the claim is made of full accomplishment of this goal, nor that such a total theory could ever be accomplished. It is, rather, that it is the belief of the author that the attitude toward such a coherent theory should be conducive toward real progress in our understanding of physics in all domains.

I am most grateful to my loving wife, Yetty, for her encouragement and her indispensable role as a 'sounding board' for ideas, during the

entire development of the research that led to the final fruition of this presentation.

Buffalo, New York
August 2003

Mendel Sachs

Contents

1. Introduction	1
1.1 History and Philosophy of the Quantum Theory	2
1.1.1 Blackbody Radiation	2
1.1.2 Photoelectric Effect	3
1.1.3 Compton Effect	4
1.1.4 Atomic Spectra and the Bohr Atom	4
1.1.5 The Seminal Experiment: Electron Diffraction	6
1.1.6 Interpretations of Quantum Mechanics	9
1.2 History and Philosophy of Theories of Gravity	
Leading to the Theory of General Relativity	11
1.2.1 Newton's Third Law of Motion	15
1.2.2 Predictions	
of Einstein's Theory of General Relativity	18
1.3 A Quantum Theory of Gravity	21
2. The Theory of General Relativity:	
Einstein's Formulation	25
2.1 The Spacetime Language	26
2.2 On Invariance of the Speed of Light	27
2.3 The Riemannian Metric	28
2.4 Expression of Einstein's Field Equations	30
2.5 Einstein's Equations from the Variational Principle	31
2.6 The Vacuum Equation	34
2.7 The Geodesic Equation	36
2.8 The Schwarzschild Solution and the Newtonian Limit	39
3. A Unified Field Theory	43
3.1 Einstein's Field Theory in Quaternion Form	44
3.2 Spin Affine Connection	47
3.3 The Quaternion Variables in a Riemannian Spacetime	48

3.4	Derivation of the Quaternion Metrical Field Equations from the Principle of Least Action	50
3.5	A Symmetric Tensor–Antisymmetric Tensor Representation of General Relativity	53
3.5.1	Einstein’s Field Equation from the Symmetric Tensor Part	54
3.5.2	The Maxwell Field Equations from the Antisymmetric Part	55
3.6	The Geodesic Equation in Quaternion Form	58
3.7	Summary	61
4.	Quantum Mechanics from a Theory of Inertial Mass in Relativity	63
4.1	Introduction	64
4.2	Discovery of Quantum Mechanics	67
4.3	Inertial Mass from General Relativity	71
4.4	The Matter Field Equations in General Relativity	73
4.5	Gauge Covariance	73
4.6	The Elementary Interaction	75
4.7	Proof of the Attractive Gravitational Force in the Newtonian Limit and the Oscillating Universe Cosmology	76
4.8	From the Mach Principle to the Generalized Mach Principle	77
5.	Electromagnetism	79
5.1	Introduction	80
5.2	Interpretation of Maxwell’s Equations in the Holistic Field Theory	83
5.3	The Elementary Interaction Formulation	84
5.4	A Spinor Formulation of Electromagnetic Theory in Special Relativity	85
5.5	Invariants and Conservation Equations in the Spinor Formalism	86
5.6	Lagrangian for the Spinor Formulation of Electromagnetism	88
5.7	Faraday’s Approach and the Mach Principle	90
5.8	Spinor Formulation of Electromagnetism in General Relativity	90
5.9	Extension of the Spinor Conservation Laws of Electromagnetism in General Relativity	91

5.10 The Electromagnetic Interaction Functional in the Matter Field Equations	94
5.11 Delayed Action at a Distance	95
6. The Pauli Principle and Pair Creation/Annihilation	97
6.1 Introduction	98
6.2 The Individual Particle Model	100
6.3 The Free Field Limit	100
6.4 Conservation of Interaction	101
6.5 The Pauli Exclusion Principle	103
6.6 Sufficiency of the Three Conditions for Proof of the Pauli Principle	108
6.7 Fermi–Dirac Statistics from the Nonrelativistic Approximation for Ψ	110
6.8 Bound Particle–Antiparticle Pairs. Ground State	111
6.9 Energy and Momentum of the Bound Pair in its Ground State	113
6.10 Dynamical Properties of the Pair in its Ground State ..	114
6.11 Pair Creation, the Physical Vacuum, and Blackbody Radiation	117
7. Atomic and Elementary Particle Physics	121
7.1 Introduction	122
7.2 Hydrogen	124
7.2.1 Linearization of the Hydrogen Field Equation ...	124
7.2.2 The Lamb Splitting	127
7.3 The Neutron	130
7.3.1 Binding Energy of the Neutron: A Phenomenological Determination	131
7.3.2 The Neutron Lifetime	132
7.3.3 The Neutron Magnetic Moment	134
7.4 Mass Doublets: The Electron–Muon	135
7.5 Infinite Lepton Spectrum	136
7.6 The Pion	138
7.6.1 Ratio of Neutral and Charged Pion Masses	139
7.6.2 Ratio of Neutral and Charged Pion Lifetimes	140
7.7 CP Violation in Neutral Kaon Decay	141
7.8 Charge Quantization in General Relativity	143

8. Astrophysics and Cosmology in General Relativity ..	147
8.1 Introduction	148
8.2 Principle of Equivalence	150
8.3 The Quaternion Geodesic for a Stationary Orbit	152
8.4 Planetary Motion	153
8.5 The Schwarzschild Problem	155
8.6 The Radial Solution and Perihelion Precession	158
8.7 The Hubble Law and Cosmology	159
8.7.1 The Oscillating Universe Cosmology	160
8.7.2 Dynamics of the Expansion and Contraction of the Universe	161
8.7.3 Spiral Structure of the Universe	163
8.8 Black Holes and Pulsars	166
8.8.1 Possible Model of a Pulsar	168
8.8.2 Damped Oscillatory Motion and Pulsars	170
8.9 Separation of Matter and Antimatter in the Early Universe	171
Bibliography	175
Index	189

1. Introduction

In this chapter, I will spell out as clearly as possible what are the foundational ideas that underlie the quantum theory, and the theory of gravity according to the theory of general relativity, and how these theories of matter differ from the classical views.

On the quantum theory, we will first discuss how it appeared on the scene in physics in the 1920s, including the ensuing philosophical basis of this theory, and then move on to some of the differing interpretations – of Schrödinger, Bohr/Heisenberg, Born, de Broglie, Bohm and Einstein.

The next section will focus on the history and philosophy of the theories of gravity. This section starts with ideas of Aristotle, of the ancient times, to ideas of Galileo and Newton, of the Renaissance period, and then to Einstein, of the contemporary period. Conceptual differences will be discussed between the classical views of Newton (action at a distance and atomism) and the theory of general relativity (holism, continuity, and forces propagating between interacting bodies at a finite speed).

Comments will then be directed to the possibility of a quantum theory of gravity. There has been a great deal of discussion on this topic in the current literature. Nevertheless, I will argue that because of the incompatible underlying concepts and mathematical expressions of the quantum and general relativity theories, it is not possible, in principle, to unify them into a quantum theory of gravity.

The approach of my research program will then be spelled out, which starts at the outset with the most general expression of the theory of general relativity, as foundational, while discarding the formal expression of the quantum theory. It is found that the formal (Hilbert space) expression of the quantum theory appears as a linear approximation for a generally covariant field theory of the inertia of matter in general relativity, which originates in a unified field theory. The origin of the field unification of the inertial manifestation of matter and its gravitational and electromagnetic force manifestations are shown to

follow from fully exploiting the algebraic as well as the geometrical implications of the principle of covariance – the underlying axiom of the theory of general (and special) relativity. The Mach principle, in a generalized form, and gauge covariance also will be shown to play essential roles in the structuring of this unified field theory.

1.1 History and Philosophy of the Quantum Theory

1.1.1 Blackbody Radiation

Near the turn of the 20th century it was discovered by Max Planck that the spectral distribution of blackbody radiation could be fitted to a model whereby the energy in each of the modes of the frequency components of radiation enclosed in a cavity is linearly proportional to its frequency.

The experiment is as follows: a cavity with a small window in its side is maintained in a heat bath at a constant temperature. The walls of the cavity are in thermodynamic equilibrium with the radiation that is emitted and reabsorbed at the same rate by these walls. This is called blackbody radiation. A filter is placed in the window of the cavity that will only transmit a single frequency radiation component. The intensity of this radiation is then measured. The observation is then continued for the visible spectrum of frequencies, by placing the appropriate filters in the window of the cavity. One then plots the intensity of the radiation field in the cavity as a function of frequency. This is the spectral curve for blackbody radiation. Plotting the intensity as a function of frequency ν , it is found to have a characteristic shape with a maximum, and then go to zero at zero frequency and at infinite frequency. Repeating the experiment at different temperatures yields the same general shape for the spectral curve of blackbody radiation.

The blackbody radiation curves were not the expected ones, according to the analysis using classical statistics (Rayleigh–Jeans). Their result depended on the assumption that the energy in each radiation mode in the cavity depends on the square of its frequency. Their prediction was that, as the wavelength $\lambda = c/\nu \rightarrow 0$, where c is the speed of light, the intensity of the radiation diverges to infinity. This is the so-called ultraviolet catastrophe. In contrast with this prediction, the empirical result is that as the wavelength decreases, the intensity goes through a maximum and then decreases to approach zero (in the ultraviolet region) as the wavelength approaches zero.

Planck analyzed this problem with the assumption that the energy of a mode of the enclosed radiation field in the cavity is linearly proportional to its frequency, $E_\nu = h\nu = hc/\lambda$. The constant of proportionality h is Planck's constant. It was found to be a universal constant of nature. With this assumption, an analysis by Planck, based on Maxwell–Boltzmann classical statistics for the ‘gas’ of radiation modes in the cavity (where each mode is distinguishable, i.e., ‘tagged’), he correctly predicted the curve that was observed for blackbody radiation.

Planck's discovery of the linear relation $E_\nu = h\nu$ between the energy and the frequency of a radiation mode in the cavity led to the idea of the ‘quantization’ of electromagnetic radiation, and the idea of a particle of light, called a photon. This theory is referred to as the old quantum theory.

It is important to note at this stage of the discussion that the ‘photon’ is not the same sort of particle as the ‘electron’. The electron has inertial mass. It can thus be slowed down or speeded up by an external electric force. But, as we will see later on in our discussion of relativity theory, the ‘photon’ can only propagate at a constant speed c , the speed of light. Since it has no inertial mass, it cannot be slowed down or speeded up from the constant speed c . All that can happen to a photon (after it has been created with energy E_γ and speed c) is that it can be annihilated when it is absorbed by matter, which in turn elevates the energy of the matter by $E_i - E_j = E_\gamma = h\nu$ units of energy.

1.1.2 Photoelectric Effect

A beam of monochromatic light (single frequency) impinges on a piece of metal that in turn is in an electrical circuit, with a battery, a resistor and an ammeter. The ammeter reads the current in the circuit, $I = V/R$, where V is the voltage across the resistance R . At first, the current is constant. When the frequency of the light is increased to some threshold value ν_0 , the current reading increases. Continuously increasing the light frequency from that threshold value increases the current reading linearly. This is the photoelectric effect.

The voltage change ΔV across the resistor corresponding to the frequency change $\Delta\nu$, multiplied by the electron charge e is

$$\Delta E = e\Delta V = eR\Delta I = h\Delta\nu .$$

Thus, Planck's constant h appears again in the photoelectric effect, the same constant that appeared in the fit to the blackbody radiation spectral curve.

There is much more to say about the photoelectric effect, but this is the essence of it. It was a further substantiation of the quantization of monochromatic light in terms of its energy $h\nu$. In spite of these results from blackbody radiation, analyzed by Planck, and the photoelectric effect, analyzed by Einstein, neither Einstein nor Planck accepted the concept of the ‘photon’ as an elementary particle of light. In a letter to Einstein (6 July, 1907), Planck said: “For I do not seek the meaning of the quantum of action (light quantum) in the vacuum, but at its site of absorption and emission.” In a letter to Laub (4 November, 1910), Einstein said: “I am very hopeful that I will solve the radiation problem and I will do it without light quanta.”¹

In the theory developed by Einstein, the photon plays the role of a virtual field that propagates as a signal between interacting electrically charged matter, to affect their mutual interaction. But the ‘photon’ is not a thing on its own, in this view. Similarly, the view of Planck was that what is ‘quantized’ is a gas of radiation in the cavity – a set of modes of radiation that couple charged matter of the walls of the cavity and the matter that constitutes the measuring apparatus that ‘looks into’ the cavity to measure the frequencies of these modes.

1.1.3 Compton Effect

Another important experiment that ushered in the ‘old quantum theory’ was the Compton effect. Here, a photon scatters from an electron, thereby changing its energy (and therefore its frequency) to a lower value, while the electron increases its energy by the same amount. It was found that the scattering angle of the electron was dependent on the same constant h determined in the blackbody radiation spectrum and the photoelectric effect, thereby verifying the assumption of the quantization of light.

1.1.4 Atomic Spectra and the Bohr Atom

A seminal observation in the early days of the ‘old quantum theory’ related to the measurement of the emission spectrum of excited atoms. The Ritz combination rule said that the frequencies of any of the various lines of the emission spectrum of hydrogen added as follows:

$$\nu_1 = \nu_2 + \nu_3 = \nu_4 + \nu_5 + \nu_6 = \dots ,$$

¹ Both quotations are from Anna Beck, *The Collected Papers of Albert Einstein*, Vol. 5: The Swiss Years Writings, 1902–1914, English translation, Princeton, 1994.

where ν_j is the frequency of any line in the observed spectrum. Niels Bohr explained this rule with his model of the atom, whereby the orbital electrons are in ‘quantized orbits’, with respective energies $E_1, E_2, \dots, E_6$, relative to the positively charged nucleus of the atom. The idea is the following. When the orbital electron drops in energy from the j th state to the i th state, it loses energy $E_j - E_i$, which then transforms into the energy of a created photon $E_\gamma = h\nu_{ij}$. It then follows that for a series of de-excitations between the different states of the atom, the energy losses (its emissions) are:

$$\begin{aligned} E_f - E_i &= (E_f - E_1) + (E_1 - E_i) \\ &= (E_f - E_3) + (E_3 - E_5) + (E_5 - E_i), \end{aligned}$$

and so on. Dividing this equation by Planck’s constant h , and using the relation

$$\frac{E_j - E_k}{h} = \nu_{ki},$$

we have the empirically verified Ritz combination rule for the frequencies in the emission spectrum:

$$\nu_{fi} = \nu_{f1} + \nu_{1i} = \nu_{f3} + \nu_{35} + \nu_{5i},$$

and so on.

Questions that arise in regard to the Bohr model are as follows:

- What is the physical cause for the de-excitation of the atomic electron from the higher to the lower energy levels?
- What is the physical cause for the creation of a photon when the atom de-excites? Of course, the latter process conserves energy in that the energy lost by the atomic electron is given to the created photon. But it does not explain how the photon was created from a vacuum (i.e., from no photons).
- When the electron loses energy by dropping from one energy level to a lower one, but before it reaches the lower energy level, it would have lost energy while the photon is not created until the electron does reach the lower energy level. Thus it seems that energy is conserved only when one sees the electron in one energy state or another, but not when it is in transition between energy levels. How is this explained in the context of an energy-conserving system? That is, does the law of conservation of energy apply only when one is looking at the atom in one state or another, but not when the electron is in transition between the states of the atom? The idea of the

Copenhagen school is that the conservation laws apply only when a macro-observer is observing the atom in one state or another, but not when the atomic system is in transition between its states. The physical properties of a microsystem of matter that is measured by a macro-observer then defines the micromatter. This is a crucial aspect in the theory of matter according to the new quantum view that was to emerge with the Copenhagen school.

Summing up, the experiments of the ‘old quantum theory’ at the turn of the 20th century, on blackbody radiation, the Compton effect, the photoelectric effect, atomic spectra, were all explained in terms of the ‘photon’, and the new universal constant, Planck’s constant h , that entailed the quantum of radiation $h\nu$. This evolved into the new quantum theory, called quantum mechanics, in the 1920s. This development will now be outlined in an historical context.

1.1.5 The Seminal Experiment: Electron Diffraction

The seminal experiment that led to the ideas of the Copenhagen School was the set of observations in 1927, by C.J. Davisson and L.N. Germer in the US and by G.P. Thomson, in the UK, that electrons can scatter from a crystal lattice as though they are continuous waves. These experiments were preceded three years earlier by the theoretical speculation of Louis de Broglie on the possibility that an electron (or any other elementary particle with mass) has a wave nature. According to de Broglie’s hypothesis, its particle-like nature – its momentum p – relates to its wave-like nature – its wavelength λ – according to the reciprocal relation $p = h/\lambda$.

The experiments on electron diffraction verified de Broglie’s speculation about the ‘matter wave’. A question then arises. If the electron is indeed a massive particle, then one should expect that its scattering from a crystal lattice would reveal, on an absorbing screen, a geometrical mapping of the atoms of the crystal lattice. Instead, what was seen was that, when the lattice spacing is the same order of magnitude as the de Broglie wavelength of the electron $\lambda = h/p$, the distribution of scattered electrons on the absorbing screen was not unlike the pattern of X-rays that have been scattered by a crystal lattice. That is, there are regions of constructive interference, where the electrons bunch together, and there are regions of destructive interference where no electrons are seen to land. Of course, one expects scattered X-rays to land this way on the absorbing screen (as had been seen earlier

by Bragg) because we know at the outset that X-rays are electromagnetic waves, whose wavelength is the order of magnitude of the lattice spacing of the crystal. It is the expected diffraction pattern for the scattering of waves from a crystal lattice.

But why should electrons diffract in this way if indeed they are discrete particles of matter? One answer is that it may be an illusion that electrons are truly point particles; rather, they may be, most fundamentally, (matter) waves at the outset. This is a difficult explanation because there are experimental circumstances wherein electrons appear to be point particles of matter, such as the cathode ray experiment of J.J. Thomson.

With these two empirical facts in mind, Niels Bohr and the Copenhagen school proposed a resolution to extend Einstein's idea of wave-particle dualism from the (massless) photons of electromagnetic radiation to material particles with finite mass, such as electrons. The idea was then that, when one does an experiment to view the electron as a discrete particle, it is such a particle at that time. But when an experiment is carried out to view the electron as a wave, it is a wave at that time. Both statements – that the electron can be a particle and that it can be a wave – are taken to be true, so long as the observations of these states of the electron are not seen by the macro-observer simultaneously.

This assertion fits in with the epistemological stand of logical positivism. It is a philosophical approach, first proposed by Ernst Mach and the Vienna Circle, around 1900. It is the idea that in principle the only meaningful statements about nature must be verifiable with the human senses or their measuring instruments. This is called 'the principle of verifiability'. It was claimed by Bohr and Heisenberg to be a natural epistemological basis of knowledge about the natural world. In his initial paper on quantum mechanics, W. Heisenberg said: "The present paper seeks to establish a basis of theoretical quantum mechanics founded exclusively upon relationships between quantities which in principle are observable."² This is the view of logical positivism, based on the principle of verifiability. Bertrand Russell gave a well-known refutation of the principle of verifiability: "This principle is not, in itself, verifiable by the human senses or instruments. Thus, if it is true, it must be false. Therefore it is false".

According to Bohr's and Heisenberg's positivistic view, then, the observer must be involved in the definition of what an electron is, rather than an outside experimenter who probes the nature of the ob-

² W. Heisenberg, *Zeits f. Physik* **33**, 879 (1925).

jective thing, called ‘electron’, independent of himself (or herself). The Copenhagen view then defines the elements of matter in an irreducibly subjective manner.

An epistemological view different than this one is the idea of realism, wherein the elements of matter are what they are, independent of who or what may be probing their properties. In this approach, what we see in experimentation must then be rationally interpreted to arrive at assertions about what the nature of this matter is, independent of ourselves as observers and our mode of measurement.

Niels Bohr and Werner Heisenberg led the philosophical view of positivism in physics. Most of the physics community has followed this approach from the 1920s until this time. But there were some very notable physicists in the 20th century who took the stand of realism in physics, such as Einstein, Planck, Schrödinger, de Broglie and Bohm. The former positivistic philosophy is that of the quantum theory. The latter realist philosophy is that of the theory of general relativity. It is the realist philosophy that I believe is the one where the truth lies in science, and will flourish in 21st century physics.

Not too long after the experimental discoveries of the wave nature of matter, in the 1920s, Erwin Schrödinger discovered the equation whose solutions are the matter waves. This is the so-called Schrödinger wave equation. His formalism correctly predicted the energy levels of atoms and the transitions between its states. Around the same time Werner Heisenberg discovered an equation in terms of matrices of numbers, representing the discrete observables such as energy values, in particular states of an atom, as well as predicting correctly the transitions between the states of the system. This formalism is called Heisenberg’s matrix mechanics. Thus both the continuous wave theory of Schrödinger and the discrete matrix theory of Heisenberg made identical predictions for the atomic states of matter.

Not long afterward, it was shown by C. Lanczos, and independently by Schrödinger, that the Schrödinger representation of the quantum theory and the Heisenberg representation could be mathematically transformed into one another. Lanczos showed this equivalence by transforming the differential equation of Schrödinger into an integral equation. From there, he was able to demonstrate the Heisenberg form of matrix mechanics. Thus the Schrödinger equation and the Heisenberg equation are equivalent formalisms; this is the reason that they gave identical predictions, though this was not obvious to the physicists at first glance! They called these mathematical expressions, which

correctly gave predictions of the physical properties of atomic matter, quantum mechanics.

1.1.6 Interpretations of Quantum Mechanics

Soon after the discovery of quantum mechanics, Max Born found that one could express the formalism in terms of a probability calculus. Thus, Born, Bohr and Heisenberg interpreted Schrödinger's matter waves as waves of probability. This was to be the (complex number) amplitude whose absolute square is the probability that the particle of matter whose properties are being measured is at a particular point of space. The probability was then tied to quantum mechanics as a theory of measurement – made by a macro-observer on micro-matter. This view was then in line with the positivistic philosophy, whereby the elements of matter are defined subjectively in terms of the measurements of their properties, expressed with a probability calculus. These ideas will be discussed in more detail in Chap. 4. Proponents of these ideas are said to belong to the 'Copenhagen school'.

While the majority of physicists have accepted the truth of the ideas of the Copenhagen school, there have been other interpretations of the empirically successful equations of quantum mechanics. I will now briefly describe some of these.

Schrödinger's View. E. Schrödinger did not accept the probability interpretation of his wave equation at the outset. He thought of his wave solutions – the matter waves – as real waves, just like ocean waves or waves of electromagnetic radiation. His idea, after seeing the results of the electron diffraction experiments, was to complete the Maxwell formulation of electromagnetic field theory by properly expressing the real-number-valued source terms (on their right-hand side) in terms of complex functions that are the de Broglie matter waves ψ – the solutions of Schrödinger's wave equation. He found that these source terms might be factorized into a product of ψ – a complex function – and its complex conjugate $\bar{\psi}$. Thus the charge density becomes $\rho = e\bar{\psi}\psi$ and the current density becomes $\mathbf{j} \propto \text{Re}(\bar{\psi}\nabla\psi)$, where the overline denotes (henceforth) the complex conjugate, e is the electron charge and ∇ is the gradient operator.

Thus Schrödinger believed that the wave nature of electrons (and any other electrically charged elementary matter) is implicit in the real number-valued source terms of Maxwell's field equations for electromagnetism. It is not 'unfolded', that is, available to be observed, until experiments such as electron diffraction are carried out.

De Broglie, Bohm and the Hidden Variable View. Not long after the successes of quantum mechanics, Louis de Broglie suggested his ‘double solution’ interpretation of the quantum formalism. This was an approach, like Schrödinger’s, that attempts to restore determinism to physics. His idea was that, in addition to the probability calculus of quantum mechanics, there must be variables that relate to the objective electron, independent of anyone’s observation of it. He saw the Schrödinger wave equation as a subjective part of the theory of the electron, relating only to the measurements of its properties. But he believed that, buried inside of the probability wave there must be a singular function ζ representing the real electron, independent of any measurements on its properties. Thus, to complete the description of the electron, there must be another mathematical equation in ζ depending on the space and time variables. It is this equation that entails the actual dynamics of the real electron.

De Broglie argued that this added function must be a point singularity, coupled to the Schrödinger wave ψ . The latter complex function, in turn, must influence the function ζ , because, he felt, the probability wave must in some way guide the point singularity wave of the electron. He then concluded that both of these functions must influence each other, and that the equation in ζ must be nonlinear, while the equation in ψ must remain linear, since it is to represent a solution for a probability calculus. The added function ζ is then not directly observed – it is a hidden variable that relates to the deterministic electron. This view is called de Broglie’s double-solution interpretation of quantum mechanics. The analysis has yet to be further analyzed and taken to completion.

In the 1950s, in trying also to restore determinism to physics, David Bohm took a different view of hidden variables to de Broglie. What he did was to add to the independent variables, additional hidden parameters that the matter field depends on:

$$\psi(\mathbf{r}, t) \longrightarrow \psi(\mathbf{r}, t, \Lambda) ,$$

where $(\mathbf{r}, t)$ are the ordinary space and time independent variables and $\{\Lambda\}$ are the additional hidden parameters. Then, the change of Λ with time is to denote the actual velocity of the particle’s trajectory. The matter wave is viewed as the dependent variable $\psi(\mathbf{r}, \Lambda(t))$. The equation in the matter field must then also entail a dependence on the changes of the hidden parameter Λ , to restore determinism in the description of the particle. What appears in this regard is an extra ‘potential’ in the equation for the matter field.

Proceeding from the single-particle description in this way to a many-particle system, the matter field becomes the dependent variable

$$\psi(\mathbf{r}, \Lambda_1(t), \Lambda_2(t), \dots) .$$

The current density for the velocity of any of these trajectories $\Lambda_i(t)$ then depends on the rate of change of the matter field ψ with respect to the hidden variable $\Lambda_i(t)$. This implies that the fundamental description of a single particle, say an electron, depends on the trajectories of all of the other particles of the system, in other places in space. Thus the given electron has a nonlocal description.

The question that then arises is: can there be a hidden variable formalism that is local? There is reason to believe that one cannot indeed maintain the Hilbert space formalism, required of quantum mechanics, within the hidden variable theory. This is based on the so-called von Neumann theorem.³

Einstein's View. The application of Einstein's theory of general relativity is that this problem of the quantum theory may be resolved by abandoning the model of matter in terms of discrete particles altogether. Instead, one has a single, holistic continuum, wherein what were formerly called discrete, separable particles of matter are instead the infinite number of distinguishable, though correlated manifestations of this continuum, that in principle is the universe. Hence, wave-particle dualism, which is foundational for the quantum theory, is replaced by wave (continuous field) monism.

In my research program, based on general relativity theory, it will be shown (Chap. 4) that the formal (Hilbert space) expression of quantum mechanics is a linear approximation for a generally covariant, nonlinear field theory of the inertia of matter. Thus, as Einstein originally anticipated, quantum mechanics is a derivative feature of matter in the local domain, where the curved spacetime approaches the flat spacetime representation, as an approximation. Instead of hidden variables, one now has a theory in terms of 'exposed variables' – the field variables that solve the equations for matter in general relativity.

1.2 History and Philosophy of Theories of Gravity Leading to the Theory of General Relativity

In the period of ancient Greece, about 2300 years ago, Aristotle claimed that the reason why a material body above the ground (say an apple

³ J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton, 1955).

on a tree) falls to the ground, is that Earth is down and the apple is up, down being the natural place for matter to be. Aristotle learned from his teacher Plato as well as earlier Greek philosophers, that there are four material constituents in the world: air, fire, earth and water. The natural place for air and fire is up. He considered the stars of the night sky to consist of fire, the natural place for fire is then in the heavens. Several centuries before Aristotle, the Greek philosopher Thales considered water to be the ultimate stuff of all matter, asserting that all matter comes out of water and returns to it.

Aristotle's law of gravity is a statement that any material body falls from above to below because it seeks its correct place in space. This place is at the bottom. That is, to Aristotle, the bottom, which is the center of space, is where the Earth sits at rest in the universe, and it is the natural place for matter to be.

Claiming to see that the quantity of matter in the universe is finite, Aristotle concluded that the space of the universe must also be finite in extent, since the logical basis of space is that it exists to be occupied by matter.

One of the features of gravity perceived since the ancient times is that it is a force that is only attractive, i.e., matter is drawn to other matter. For example, the freely falling apple is drawn to the Earth. An important question in physics is then: why does matter never repel other matter? We will see in our discussion of the theory of general relativity in this book that there are indeed instances in nature whereby the gravitational force can be repulsive, side by side with the attractive aspect of the gravitational force. Which one of these aspects of gravity predominates depends on the physical conditions. This problem arises in discussions of cosmology – the laws of the universe as a whole. At the present stage in the evolution of the universe, it is observed that, rather than moving toward each other, the galaxies are moving away from each other. This is called the expanding universe. It is indeed an example of repulsive gravitational forces.

Aristotle in ancient Greece not only theorized. He also depended on empirical verification to verify his speculations. A great scientist, in the Renaissance period, seventeen centuries after Aristotle, who studied the laws of gravity from the experimental as well as the theoretical side, was Galileo Galilei. He insisted that the experimental verification of a theoretical speculation was essential in order to claim a scientific truth.

Galileo's law of gravity for a body in 'free fall' was $g = \text{constant}$, where g is the acceleration of a body in its fall. Galileo discovered that

g is independent of the weight or size of the body and that, numerically, $g = 9.8 \text{ m/s}^2$.

At first sight, this seems to be a false statement, since we do see heavier bodies falling more rapidly to the ground than lighter bodies. For example, a lead weight reaches the ground before a feather, being dropped from the same height. Galileo argued that the reason for this is that the air resistance to the lead weight in its fall is less than it would be for the feather. To prove his claim, Galileo constructed a device that excluded the air – an evacuated jar. Triggering both bodies, of different weights, to fall at the same time in the jar, Galileo saw that they landed at the bottom at the same time, both falling with the acceleration g , thus proving his law of gravity.

Galileo used the experimental method to make an accurate measurement of g . He deduced that if a body at the top of an (almost) frictionless inclined plane is released at time $t = 0$, the distance it would move down the plane in the time t would be

$$s = \frac{1}{2}g \sin \theta t^2 \text{ m} ,$$

where θ is the angle of inclination of the plane and t is the time traveled. Experimentally verifying that the ratio of two distances traveled is equal to the ratio of the square of the times taken, $s_1/s_2 = (t_1/t_2)^2$, he could then determine the numerical value of g for a given run as

$$g = \frac{2s}{t^2 \sin \theta} = 9.8 \text{ m/s}^2 .$$

Newton was born in 1642, the same year that Galileo died. He was the next great physicist to explore the phenomenon of gravity. He was greatly influenced by Galileo, and also by J. Kepler. Newton said that, had he not been standing on the shoulders of two giants, Kepler and Galileo, he would not have been able to see as far as he did.

Kepler was a contemporary of Galileo. He insisted on the existence of general laws of nature. Thus when he determined that the orbit of Mars is elliptical with reference to the Sun's position at one of the foci of its elliptical orbit, he said that the orbits of all of the planets should also be elliptical, with the Sun at one of the foci of their elliptical orbits.

This was indeed a revolutionary statement to be made in the 17th century. All scientists since Plato and Aristotle, as well as Galileo, in Kepler's time, believed that all orbits in the heavens were circular. Plato argued that the circle is the most perfect of all shapes, since

viewed from its center the circle looks the same from all angles. He argued that since the Heavens must be perfect, so must the orbits of the planets be perfect, and therefore circular. However, Kepler saw that the orbit of Mars was not circular, but instead elliptical. He then concluded that all of the orbits of the planets are elliptical. So convinced was he of Plato's argument, Galileo would not accept Kepler's finding in his lifetime!⁴

Newton set about finding the explicit law of gravity that would explain Kepler's discoveries about planetary motion, as well as Galileo's discovery of the constancy of g near the Earth's surface.

First, Newton described Galileo's principle of inertia by Newton's first law of motion. This law states that if a body were at rest, or in motion at a constant speed in a straight line, it would remain this way forever, unless compelled to change this state by some external influence. (This was contrary to Aristotle's law that said that it is only the state of rest that is natural for bodies, while any type of motion requires some external influence.) Galileo's principle was then incorporated into Newton's second law of motion, $\mathbf{F} = m\mathbf{a}$. Here, $\mathbf{F}$ is a 'force' – the outside influence – that causes the body to change its velocity by accelerating at the rate $\mathbf{a}$. Thus if $\mathbf{F} = 0$, then $\mathbf{a} = 0$, leading us back to Galileo's law of inertia and Newton's first law of motion – if there is no outside influence on the body, then $\mathbf{v} = \text{constant}$, i.e., it will be at rest ($\mathbf{v} = 0$) or it will be moving at a constant, non-zero velocity $\mathbf{v}$. It should be noted that bold face symbols are used for the force, acceleration and velocity to signify that they are vectors – numbers that have magnitude as well as direction in space. The law, $\mathbf{F} = m\mathbf{a}$, is a compact way of signifying three laws of motion in three-dimensional space, one for each of the three coordinate directions, x , y and z . It might be that $F_x = 0$ while $F_y \neq 0$. This is the case of gravity, where $F_x = 0$ and $F_y = -mg$, if y denotes the downward vertical direction (toward the center of the earth). This combination of forces in the x and y directions led Galileo to the prediction of projectile motion. Indeed it was Galileo who recognized the vectorial nature of force, acceleration and velocity.

If the force on a body is not zero, then Newton's second law of motion, $\mathbf{F} = m\mathbf{a}$, asserts that in all cases the relation between the force that acts on a body and the acceleration it causes is a linear one. This is not only with regard to gravity, but also for all other possible types of force. It was found in the 20th century that the force laws of general relativity theory refute this linear law, although it is a good

⁴ See A. Koestler, *The Sleepwalkers* (Macmillan, New York 1968).

approximation in those cases that Newton was considering in classical physics.

Thus Newton saw that the magnitude of the force of the Earth on a body near its surface, to give it ‘weight’, is $F = W = mg$. It should be noted that if a body is at rest on a platform on the Earth’s surface, its weight is then the force $-m_g g$, where m_g is called the gravitational mass of the body. What this means is that if the platform should suddenly disappear, the body would fall freely at the acceleration $-g \text{ m/s}^2$. When it is in motion, the mass of the body is its inertial mass m_i . The empirical evidence is that $m_i = m_g$. This equality has been puzzling to physicists since the early days of attempts to understand gravity. But with Einstein’s theory of general relativity, which superseded Newton’s law of gravity, the equality can be explained, as we will discuss in Chap. 4. It asserts that there is only one kind of mass, and that it is the inertial mass, which we will signify by m . This mass is, by definition, a measure of the resistance of a body that is in a state of rest or in a state of constant motion to a change of this state. It will be discussed in detail in Chap. 4.

1.2.1 Newton’s Third Law of Motion

Newton’s third law of motion did not come from Galileo, as his first two laws did. It was a new concept in Newton’s investigation of mechanics that indeed had a great deal of influence on Einstein’s theory of general relativity, that was to come three centuries later.

Newton’s third law of motion is the assertion that for every force exerted by one body A on another B, there is an equal magnitude, but oppositely directed force of the body B exerted on body A. This implies that, rather than an open system, we must always consider the material system to be closed. Further, the existence of the closed system implies that the laws of motion are necessarily nonlinear, rather than linear. This may be seen as follows. If a force acting on B, spatially separated from A but due to A, is $\mathbf{F}_B(A)$, then B would accelerate at the rate $\mathbf{a}_B$, according to Newton’s second law of motion, $\mathbf{F}_B(A) = m_B \mathbf{a}_B$. This force, according to Newton, depends on the distance r_{AB} from A to B. The reacting force exerted on A by B, $\mathbf{F}_A(B) = -\mathbf{F}_B(A)$, according to Newton’s third law, then causes A to accelerate from the position it was at, initially, thereby changing the magnitude of the force exerted by A on B from $\mathbf{F}_B(A)$ to $\mathbf{F}_B(A')$, because the distance from A to B has then changed. Thus, the acceleration $\mathbf{a}_A$ of A before it acted on B changes to $\mathbf{a}_{A'}$ by virtue of its acting on B. We see, then,

that the force exerted by B on A depends on the parameters of A itself; that is to say, the force $\mathbf{F}_B(A)$ on the left-hand side of the force law is $\mathbf{F}_B(A(\mathbf{a}_A)) = m\mathbf{a}_B$ is itself a function of $\mathbf{a}_A$. This equation is not linear in the parameters of the body A. Nonlinearity is a general feature of the laws of a material system, such as this, which is closed. This is the case for the laws of motion in general relativity theory, as a theory of gravity, for this automatically applies to a closed system, as we will discuss in the next chapter.

Finally, from his observations of the motion of the Moon relative to the Earth, Newton discovered that the force of gravity generally depends inversely on the square of the distance between the interacting bodies, and directly on the masses of the interacting bodies, i.e., $F_G = -Gm_A m_B / r_{AB}^2$, directed only along the line joining the centers of the interacting bodies. The constant G , called Newton's gravitational constant, is the constant of proportionality between the force between A and B, and their masses and the separation between them. The minus sign signifies that the force of gravity is always attractive. Further, with Newton's discovery of calculus, he was able to write the acceleration as a second derivative, viz., $\mathbf{a} = a = d^2r/dt^2$. Here we drop the vectorial denotation since the force is directed only along the line of centers of the interacting bodies, of length r . Thus, for the case of the gravitational force between bodies A and B, Newton's second law of motion, for B acted upon by A, becomes

$$F_B(A) = m_B \frac{d^2r}{dt^2} = -\frac{Gm_B m_A}{r^2}.$$

This is a differential equation in the variable r that is the separation between the gravitationally interacting bodies A and B. The function $r(t)$ is the dependent variable that is to be determined. The time measure t is the independent variable. The equation of motion is then a second order differential equation, implying that its solution $r(t)$ is unique when two boundary conditions are specified, such as the location (in space) and the speed of one of the bodies relative to the other, at some initial time t_0 .

Newton showed that, assuming the Earth to be spherically symmetric in shape, with a homogeneous matter distribution, then in its interaction with another body with mass m , one may consider all of the mass M of the Earth to be at its center of mass, which is the center of the spherical Earth. If R is the radius of the Earth, then the force of weight acting on a given body with mass m on the surface of the Earth is $W = GMm/R^2$. Since this weight is equal to mg , Newton

discovered that Galileo's $g = GM/R^2$. With the determination of the gravitational constant G from a test case, and the mass M and radius R of the Earth, g was then found to match Galileo's value of 9.8 m/s^2 .

This is the value of g at the surface of the Earth. If someone were to stand on the Moon, his/her weight would be $W(\text{Moon}) = GmM(\text{Moon})/R(\text{Moon})^2$. With the mass of the Moon and its radius in this formula, it turns out that the weight of the person on the Moon would be $1/6$ of his/her weight on the earth. The acceleration of a falling body near the surface of the Moon would be one sixth of 9.8 m/s^2 , which is 1.6 m/s^2 . The weight of a body and its acceleration toward the surface of other massive bodies such as the other planets of the solar system would then vary according to Newton's formula for W and g .

Newton's theory of universal gravitation predicted many gravitational phenomena in addition to the weight of a body. For example, it correctly predicted the planetary orbits, the rotational motion of the Moon relative to the Earth, as well as the ocean tides due to the gravitational interaction between the Moon and our oceans. It was therefore natural for the scientist of Newton's day to extend his theory of gravitation to regions beyond the solar system – to the universe as a whole. With this view, the theory predicts that all stellar systems of the universe must rotate about each other in stationary orbits, just as the Moon was considered, from the data, to rotate about the Earth in a stationary orbit and the planets were believed, from the data, to rotate about the Sun in stationary orbits.

The stationary orbit comes from a description in which the spatial coordinates of a body are independent of its time coordinate. We will see later on that the theory of relativity predicts that the space and time coordinates of a body must unify into a single spacetime coordinate system. In the latter description, the space and time measures are not totally independent, in the sense that they transform into each other, in changing reference frames that describe a material body. That is, the statement of 'where' and 'when' in relativity theory depends on the reference frame in which the body is described. One may not generally refer to the location of a body without also saying when it is there. This theory then implies that there cannot be stationary orbits of bodies, in general. That is to say, what appears to be a stationary orbit in one frame of reference (i.e., independent of the time measure) would not be a stationary orbit in other frames of reference. Since the frame of reference is not an absolute feature of the physical system, i.e., it is a subjective aspect in the description of the body, it must be

concluded that, independently of reference frames, there are generally no stationary orbits (according to the theory of relativity) and that the observation of a stationary orbit by Newton was only an illusion, though it may have been a good approximation for a non-stationary orbit.

1.2.2 Predictions of Einstein's Theory of General Relativity

The conclusion of relativity theory that there are no stationary orbits for the objects in the sky was empirically verified by the observation, in the 1920s, of an expanding universe. E. Hubble observed that each of the constituent galaxies of the universe is continually moving away from the others. The latter observation was predicted successfully by Einstein's theory of general relativity, 300 years after Newton's discovery of the theory of universal gravitation. Thus, it took 300 years to supersede Newton's theory and to explain the phenomenon of gravity in a more scientifically satisfactory way.

In the 19th century, an important experimental observation that was not in agreement with Newton's theory of gravitation was that of the astronomer U. Leverrier. He saw that Mercury's orbit about the Sun is not stationary. He observed that it took extra time in each of Mercury's years in its orbit to reach the same place relative to the Sun's position. Thus Leverrier saw that Mercury – the smallest planet of our solar system – does not trace out a stationary orbit. This advance of the perihelion is called the perihelion precession of a planetary orbit. The perihelion is the point of closest approach to the elliptical focal point – which is the position of the center of the Sun – on the planetary orbit. The effect was predicted correctly by Einstein's theory of general relativity, both qualitatively and quantitatively.

At first, it was thought that Newton's theory was not refuted, but rather that account was not yet taken of the other planets' gravitational effects on Mercury. That is, the prediction of the stationary orbit was based on the idea that only the Sun influences Mercury's motion. The Sun's gravity is, of course, the largest effect on Mercury's motion, but the other planets do have a smaller gravitational perturbing effect on Mercury that would throw it off of its cyclic, stationary path. The astronomers knew the exact positions of the other planets of our solar system, relative to Mercury's position, at all times. Thus they were able to compute the perturbing effect that they thought would explain Mercury's observed anomalous behavior. Nevertheless, these calculations of the perturbations on Mercury's orbit by its sister planets in the solar system were not enough to explain its deviation from a

cyclic path. The precise orbit of Mercury was not explained until the appearance of the theory of general relativity, in the early part of the 20th century, which gave the empirical result, both qualitatively and quantitatively.

Another anomalous gravitational effect, not predicted by Newton's theory, was a bending of the path of starlight as it passes the rim of the Sun. The observation, made by Arthur Eddington and his group at Cambridge University around 1918, agreed precisely with the (qualitative and quantitative) prediction of general relativity theory. Here, it was a consequence of the curvature of spacetime in the vicinity of the Sun, as predicted by Einstein's theory. That is, according to general relativity theory, the starlight must move along a geodesic path which, due to the Sun's presence, is a curved path.

A third anomalous prediction of Einstein's theory of general relativity, not predicted by Newton's theory of universal gravitation, was that the spectral lines of a radiating atomic or nuclear system would be shifted toward the red end of the spectrum, increasingly as the gravitational potential at the radiating system increases. This is called the gravitational redshift. It was not actually observed conclusively until the 1950s, when R. Pound and his collaborators at Harvard University saw the effect. What they measured was the gamma spectrum of radioactive cobalt, in one gravitational potential – in the basement of their building – compared with the measurement in a different gravitational potential – on the roof of the building! They were able to use a diagnostic technique with the newly discovered Mössbauer effect. It had sufficient resolution to see the difference in the gamma ray frequencies on the roof and in the basement of the building, to one part in 10^{15} . The observation agreed quantitatively with the numerical prediction of the gravitational redshift of Einstein's theory.

It is only in a mathematical approximation that Einstein's equations for the theory of general relativity yield the formal expression of Newton's theory of gravity. Thus, Einstein's theory predicts everything that Newton's theory had predicted successfully, but in addition, other gravitational effects not predicted at all by the classical theory. Thus, in the early years of the 20th century, Einstein's theory of general relativity and its prediction of gravity met all of the conditions to supersede Newton's theory of universal gravitation.

It is important to note at this stage of the discussion, several important conceptual differences between Einstein's theory of gravity and that of Newton. First, Newton's theory is based on a model of matter in terms of atomism. It is also based on 'action-at-a-distance'. That is,

bodies are fundamentally discrete entities, representing an open system, and they act on each other with a force that depends only on their mutual separation, but not on the time. Thus, if one body is 100 million miles from another, their interaction would be spontaneous, depending only on their mutual separation in space. If one of these bodies should move to a different location relative to the other, or if it should suddenly extinguish, the second body would respond to this change instantaneously (irrespective of how far apart they may be!). Newton was not satisfied with this situation – which seemed like mysticism to him! But he said that he did not form hypotheses, and since the formulas worked, the action-at-a-distance concept must be true scientifically.

Fully exploiting Einstein's theory of general relativity, there is no action-at-a-distance. Instead, all forces, including those of gravity, propagate at a finite speed between the interacting bodies. The maximum value of this speed is the same as the speed of light in a vacuum $c = 3 \times 10^8$ m/s. We will discuss the reasons for this conclusion in more depth in the following chapter. Thus, we see that Newton's view of action-at-a-distance as fundamental in the description of interacting bodies was an illusion. The speed with which these forces propagate is so great that, in his day, Newton would not have been able to detect that there is a finite speed of propagation of the gravitational interaction between massive bodies, such as the Earth and the Sun. Indeed, propagating at the speed c , it takes the gravitational force exerted by the Sun on the Earth about 8 minutes to arrive here. Thus, if the Sun should extinguish, it would take 8 minutes for the Earth to become aware of this, and fly off on a tangent into outer space! That is, the inhabitants of Earth would have 8 minutes to live from the time that the Sun disappears! This force of gravity, as well as all other types of force, according to relativity theory, propagates as a continuous field, analogous to the ocean breakers moving against the seashore.

An important difference between the classical theory and Einstein's theory of gravity is the contrast between atomism and continuous holism. In general relativity theory, the basic variables of matter are continuous fields, rather than discrete particles. Instead of the latter, general relativity implies the continuous field concept as foundational. Rather than separable, singular bits of matter, the universe in relativity theory is one continuous whole, with an infinite number of distinguishable manifestations or modes. The latter are the 'things' that one interprets as atoms, people, stars, etc., in the Newtonian or quantum mechanical atomistic theories. But according to the theory

of general relativity, they are not in reality separable things. They are instead correlated modes of the single continuum that is the universe. The continuum ontology is then one of holism.

The debate between these two ontologies – one of atomism versus one of holism – has been ongoing throughout our history since ancient times. The atomistic view has been the one that has normally been accepted since it meets the intuition of humans more readily than the holistic view, because of the reactions of our human senses. But we must recognize that our senses do not directly reveal the truths that underlie the laws of nature. Epistemologically, the attitude that our senses tell us all that underlies nature is one of naive realism and positivism, while the ontology of holism takes the view of abstract realism. My use of the adjective ‘abstract’ refers to the fact that aspects of the real world are not directly revealed to our senses or our invented instruments. Rather, we must come to these truths by means of the process of rational deductions, starting from the premises that we intuitively propose to test the truths of the natural world and then logically deducing its conclusions to be compared with the empirical facts of nature. This is an approach akin to Plato’s cave analogy in his *Republic*. Plato envisions slaves chained to the floor, facing the shadows on the cave wall. Most of them deduce that the shadows are all that there is to the real world. But one slave manages to break his chains and he looks at the opening of the cave, where he sees that the Sun is reflecting from objects that are the cause of the shadows on the wall.

We see, then, that Einstein’s theory of general relativity explained all gravitational effects known to the scientific world until the middle of the 20th century, including effects that were not predicted at all by Newton’s theory of universal gravitation. In superseding Newton’s theory, Einstein’s theory of gravitation changed the concepts of Newton’s view from atomism and action-at-a-distance, to holism, continuity and no action-at-a-distance, where all forces propagate at a finite speed.

1.3 A Quantum Theory of Gravity

In the contemporary period of theoretical physics, at the beginning of the 21st century, a great deal of effort is being expended on attempts to formulate a quantum theory of gravity. I will argue that, however brilliant and innovative these attempts have been, such a unification is impossible to achieve. This is because the quantum theory and the

theory of general relativity (our current explanation of the force of gravity) are incompatible, both mathematically and conceptually.

For example, from the mathematical side, and independently of any approximations, the quantum theory is necessarily a linear formalism in terms of a Hilbert space with an eigenvalue/eigenfunction form. The reason for this is that quantum mechanics is, by definition, a theory of measurement in terms of a probability calculus, and probability theory must be expressed with a linear formalism. That is, any arbitrary sum of solutions of the equations of the quantum theory must also be a solution of the same equations.

On the other hand, the theory of general relativity is necessarily a nonlinear mathematical formalism. There are two main reasons for this:

- the mathematics of the field equations is to represent a closed system,
- this is a field theory – a continuum with correlated modes – that must be represented in a curved spacetime.

These features of general relativity theory will be discussed in more detail in the next chapter.

Conceptually, there are several points where the quantum theory and general relativity differ:

- indeterminism versus determinism,
- the ontology of separable, discrete atoms of matter versus the ontology of holism, implying no separability, where the ‘things’ of matter are correlated, distinguishable modes of a single continuum,
- wave–particle dualism and positivism versus wave (continuum) monism and realism,
- a subjective view of the world wherein the nature of matter depends on the observation versus a fully objective model, wherein an ‘observer’ plays no fundamental role in the definition of elementary matter.

With the quantum view, the quantitative values of the physical properties of elementary matter are not predetermined; they depend on the nature of the observation (by a macro-observer). Thus this is a ‘nondeterministic’ theory of matter. In this theory, there are two distinct types of variable: the classical variables – to represent the macro-observer – and the quantum variables – to represent the ‘observed matter’. Once it is decided, subjectively, which is the ‘observed’ and which is the ‘observer’ of the interacting, measured system, these two

types of variable are not interchangeable, thus resulting in a basic subjectivity of the theory.

In the theory of relativity (whether in the special or the general form) all variables are interchangeable, and predetermined by objective laws of matter. The theory is fully deterministic. This refers not only to the ordering parameter that is the time measure, but to all other parameters – the independent variables – that indicate the logical connections between elements of the closed system.

Summing up, from the view of both logical and mathematical consistency, it is not possible to unify the quantum theory and the theory of general relativity. In the language of quantum mechanics, the latter attempt requires the ‘quantization’ of the field equations of Einstein’s theory of general relativity, with the emergence of a new particle – the quantum of the Einstein field called the graviton. But such unification has never been accomplished, for the reasons of inconsistency spelled out above, and there is no reason to believe in the existence of the graviton.

My research program, which will be developed in this book, resolves the problem of the incompatibility of the quantum and general relativity theories by basing an explanation of the behavior of matter, in any domain from the microscopic to the cosmological, solely in terms of the theory of general relativity. It abandons the concepts and the mathematical formalism of the quantum theory, as fundamental, replacing it with a covariant field theory of matter in general relativity. But this field theory does recover the Hilbert space form of the quantum theory in a linear approximation, as we will see in Chap. 4, for a generally covariant field theory of the inertia of matter. The natural incorporation of the inertial manifestation of matter is found to be a key ingredient in the formulation of a truly unified field theory.

It is shown that the origin of the formulation of a unified field theory that incorporates the inertia of matter and its electromagnetic and gravitational force manifestations is the principle of covariance – the underlying axiom of the theory of general relativity. Also necessarily incorporated in this formulation is the generalized Mach principle – asserting that not only the inertia of matter, but all of its other claimed intrinsic properties in the particle theories are instead measures of coupling within a closed system of matter – in principle the universe. In this way, atomism is totally exorcised from this explanation of matter. A further necessary ingredient in the unified field theory is the imposition of gauge invariance. It is the necessary and sufficient condition for the inclusion in the field theory of relations in the curved spacetime

whose local limits, in a flat spacetime, are the laws of conservation of energy and of linear and angular momentum.

2. The Theory of General Relativity: Einstein's Formulation

We introduce the idea in this chapter that Einstein's theory of general relativity is a general theory of matter in all domains, and not only a theory of gravity in the large domain. This theory is based on a single idea – the principle of covariance. It is the assertion that all of the laws of nature must be fully objective. This means that their expressions are independent of any reference frame in which they may be described, that is, from the view of any particular reference frame.

One of the first implications of this restraint on the laws of nature is that the space and time measures in the language of these laws must merge in a single, fused spacetime language. This implies that a purely spatial (or purely temporal) measure in one reference frame would have to be expressed as a particular combination of spatial and temporal measures in other (relatively moving) reference frames, if the expressions of the laws of nature are to remain objective, i.e., independent of the reference frame.

A logical implication of this fusion of the space and time measures is that there must be a universal constant with the dimension of speed, independent of any reference frame. This turns out to be the speed of light in a vacuum. In the theory of relativity, this is the maximum speed of propagation of any type of force between interacting matter.

Einstein found that the geometrical logic of the four-dimensional spacetime expressed with the rules of Euclidean geometry is not adequate for the language of the laws of matter. Thus he generalized the geometry of the spacetime language to the non-Euclidean Riemannian geometry. With this new language for the laws of matter, the natural (unobstructed) paths of bodies are curves rather than straight lines.

An implication of general relativity is that the (continuous) geometrical properties of spacetime correspond with the existence of continuously distributed matter. The Einstein field equations then yield the geometrical variables (on the left-hand side of these equations) given the matter fields (on the right), or vice versa. Thus these equations are interpreted here as identities – the geometrical fields on the

left-hand sides of these equations are another way of talking about the matter source fields on the right, and vice versa. We therefore reject the source-free solutions of these equations (representing the 'vacuum', everywhere) as unphysical, in the context of this interpretation.

The equation of motion of a test body in this theory is the so-called geodesic equation in the curved spacetime. This is the path of minimal separation between any of its points. In a Euclidean space, such a path is a straight line. In the Riemannian space, it is a curve. Such equations of motion in general relativity have yielded several predictions, to be discussed in this chapter, of physical effects that are not predicted by the classical Newtonian theory of gravity, as well as predicting all of the successful results of Newton's theory. Thus it is seen that Einstein's theory of general relativity superseded Newton's theory as a correct explanation for the phenomenon of gravity.

2.1 The Spacetime Language

Initially in the analysis of the theory of special relativity, wherein the reference frames are inertial, that is, moving in a straight line at a constant speed relative to each other, it was found that the spatial and temporal measures that express the 'words' of the language of the laws, must be unified in a single spacetime measure, rather than expressing them separately.

That is to say, a purely spatial measure in the expression of a law of nature in one reference frame must be represented as a particular combination of spatial and temporal measures in other reference frames that are in motion relative to this one, if this law is to be covariant – if the form of this law is to be independent of the reference frame. Similarly, a purely temporal measure in one reference frame must be expressed as a combination of temporal and spatial measures in the language of the same law of nature in other reference frames. It is only such a unified spacetime that has objective connotation.

This is in the same sense that the words of a verbal language we use to express meanings with sentences, must transform in other languages to changed words and syntax such that the meanings of the sentences of the original language are preserved in all other languages. If, in English, one says: 'the sky is blue', and in French one says: 'le ciel est bleu', the words have changed, but the meaning in the sentence is preserved. This is the idea of covariance. In the changed 'mathematical language' of laws of nature, the 'words' that denote spatial measure

and those that denote temporal measure must fuse into words that fuse the two measures, according to the covariance requirement of relativity theory.

The unification of space and time measures in special relativity carries over to general relativity, where the different reference frames in which one compares the forms of the laws of nature are in arbitrary kinds of relative motion.

2.2 On Invariance of the Speed of Light

The unification of the space and time measures in the languages of the laws of nature implies that they must be expressed in the same units. Thus instead of referring to t , t' , t'' , and so on, seconds as temporal measures in the different reference frames, we must refer to ct , ct' , ct'' , and so on, centimeters, as the temporal measures. At this stage, c is a conversion factor – a constant with the dimension of centimeters/second – a speed. It must be a universal constant, independent of reference frame, if we are to distinguish physically between spatial and temporal measures. For example, the law in physics of conservation of energy refers to the constancy of energy in time, distinguishing thereby the temporal measure from a spatial measure. Thus, the universal speed c must be a part of the spacetime language that enters into the expressions of all of the laws of nature.

The first law examined by Einstein, wherein he discovered special relativity theory, was the Maxwell formulation of electromagnetism. The constant c appears in this law as the speed of light in a vacuum. The further examination of the theory of relativity reveals that c is the maximum speed of propagation of (any type of) force between interacting bodies. Thus, contrary to Newton's theory, wherein we have action-at-a-distance – that is, bodies interact with each other spontaneously, independently of their mutual separation – all forces, according to the theory of relativity, propagate at a finite speed, whose maximum value is the speed of light in a vacuum, c , so that it takes a finite amount of time for one body to respond to another body that it interacts with. We see here, then, that it is not necessary to axiomatize the invariance of the speed of light in a vacuum. The invariance of c is indeed a logical consequence of the fusion of space and time into spacetime, following in turn from the principle of covariance, in the languages of the laws of nature. The only underlying axiom of the theory of relativity (of the special and general forms) is then the principle of covariance – the assertion of the independence of the forms

of the laws of nature with respect to transformations of the language elements between all possible reference frames. This is the idea that the laws of nature must be fully objective.

The usual language of the laws of nature is in terms of continuously varying parameters called space and time measures, as the words of this language. The logic of this language, analogous to the syntax of ordinary verbal language, is in two parts – geometry and algebra. The geometrical logic refers to the relations of points to points in the sense of relative congruence, mapping, angle preserving transformations, etc. The algebraic logic is in the sense of rules of combination of the space-time points, commutativity/anticommutativity, associativity, etc. The algebra of spacetime is expressed most compactly in terms of a symmetry group. The group of special relativity is the Poincaré group, that of general relativity is the Einstein group.

2.3 The Riemannian Metric

Because the metric relations between the points of space and time must relate to the matter distribution whose laws are to be expressed with the spacetime parameters, and because the matter distribution is generally variable and the metric relations of a Euclidean space and time are constant everywhere, Euclidean geometry is inadequate as the correct geometry for laws of nature in general relativity. Thus it is necessary to change the geometrical logic of relativity theory from Euclidean geometry to a non-Euclidean geometry. Einstein chose the differential geometry of Riemann for the spacetime language of the laws of nature. The differential metric of Riemann has the feature that it approaches the flat (Euclidean) spacetime metric asymptotically in the limit as the mutual interaction between matter becomes vanishingly small. That is, as the mutual interaction between matter components of a closed system becomes vanishingly small, the limits on the squared differential metric become:

$$ds^2(\text{Riemann}) = g^{\mu\nu}(x)dx_\mu dx_\nu \longrightarrow ds^2(\text{Euclid}) = dx_0^2 - d\mathbf{r}^2. \quad (2.1)$$

The usual convention is used where the symbols $\mu, \nu = 0$ (temporal measure), 1, 2, 3 (spatial measures) are summed over, $d\mathbf{r}$ is the 3-dimensional spatial differential, and x stands for the functional dependence on all four space and time coordinates. $g^{\mu\nu}(x)$ is a second-rank tensor, called the metric tensor. Thus we see that the limits on the metric tensor components, as the mutual interactions of matter become vanishingly small, are those of the flat spacetime metric:

$$\begin{aligned}
g^{00}(x) &\longrightarrow 1, & g^{kk}(x) &\longrightarrow -1 \quad (k = 1, 2, 3), \\
g^{\mu\nu}(x) &\longrightarrow 0 \quad (\mu \neq \nu).
\end{aligned}
\tag{2.2}$$

Clearly the metric tensor is symmetric, i.e., $g^{\mu\nu} = g^{\nu\mu}$. This follows because of the following equality of sums: $g^{\mu\nu}dx_\mu dx_\nu \equiv g^{\nu\mu}dx_\nu dx_\mu = g^{\nu\mu}dx_\mu dx_\nu$, since dx_μ and dx_ν commute under multiplication. It then follows that the symmetric second-rank metric tensor has ten independent components: $g^{00}, g^{11}, g^{22}, g^{33}, g^{01}, g^{02}, g^{03}, g^{12}, g^{13}, g^{23}$. These are ten continuous functions that must be determined to prescribe the spacetime. They play the role of a ten-component potential in Einstein's theory, applied to the phenomenon of gravity, analogous to the one-component potential in Newton's theory of universal gravitation. The metric tensor solutions follow as the solutions of Einstein's field equations, which relate the matter distribution of a closed system to its geometry.

It will be shown in this chapter that Einstein's field equations have the form

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \kappa T_{\mu\nu}. \tag{2.3}$$

The left-hand side is a complicated, nonlinear functional form in $g^{\mu\nu}$. $T^{\mu\nu}$ is the symmetric second-rank tensor which, in the flat spacetime limit, is the energy-momentum tensor of the matter distribution of the closed system. The field equation (2.3) is a system of 10 nonlinear differential equations.¹ The other tensors in (2.3) are as follows:

$$R_{\mu\nu} = \partial_\nu \Gamma_{\mu\lambda}^\lambda - \partial_\lambda \Gamma_{\mu\nu}^\lambda + \Gamma_{\alpha\nu}^\lambda \Gamma_{\mu\lambda}^\alpha - \Gamma_{\alpha\lambda}^\lambda \Gamma_{\mu\nu}^\alpha. \tag{2.4}$$

This is called the Ricci tensor of the spacetime.

¹ As identities, Einstein's field equations (2.3) are 10 unique nonlinear differential equations in 10 unknowns, without ambiguity. Thus there are 10 independent relations that yield the solutions of (2.3), the 10 components of the metric tensor $g_{\mu\nu}$, given the non-zero energy-momentum matter source, the 10-component symmetric tensor $T_{\mu\nu}$. If the covariant divergence of (2.3) should be taken, it would yield the identity $0 = 0$. In the standard interpretation of the field equations (2.3), wherein they are not identities, one may have physical meaning for the vacuum situation, the right-hand side being identically zero, everywhere. In this case, the vanishing of the four-divergence of the left-hand side of (2.3) introduces 4 extra constraints, reducing the number of independent components of $g_{\mu\nu}$ from 10 to 6. But this is not the physical interpretation here. Indeed, because the Einstein group for general relativity is a 16-parameter Lie group, the number of independent components one expects for the metrical field should be 16, rather than 10, as we will see later on in the quaternion development of this field theory.

The affine connection is

$$\Gamma_{\mu\alpha}^{\rho} = \frac{1}{2}g^{\rho\lambda}(\partial_{\mu}g_{\lambda\alpha} + \partial_{\alpha}g_{\mu\lambda} - \partial_{\lambda}g_{\alpha\mu}) . \quad (2.5)$$

It is defined as the function that must be added to the ordinary derivative of a vector field in order to make it integrable, as follows:

$$A_{\mu;\lambda} = \partial_{\lambda}A_{\mu} - \Gamma_{\lambda\mu}^{\rho}A_{\rho} , \quad (2.6)$$

where the left-hand side of this equation is called the covariant derivative of the covariant vector field A_{μ} .

Finally, R in Einstein's equations (2.3), is the Riemann scalar curvature field,

$$R = g^{\mu\nu}R_{\mu\nu} ,$$

and it derives from $R_{\mu\nu\lambda}^{\rho}$ – the Riemann curvature tensor, as discussed below.

It should be noted at the outset that, in view of the interpretation of Einstein's field equations (2.3) as a way of expressing the matter of a closed system in terms of geometrical variables or vice versa, then given the tensor of the matter distribution $T_{\mu\nu}$ to derive the geometrical field $g^{\mu\nu}$ from (2.3), it is equally valid to invert these equations as a way of deriving the matter tensor $T_{\mu\nu}$, given the geometrical field $g^{\mu\nu}$.

2.4 Expression of Einstein's Field Equations

Einstein's derivation of his field equations (2.3) starts from the curvature of spacetime, which is to represent the existence of matter. The Riemann curvature tensor $R_{\mu\nu\lambda}^{\rho}$ is defined as follows: transporting a covariant vector A_{μ} on a closed path encircling the surface area ΔS , we have

$$\lim_{\Delta S \rightarrow 0} \frac{\Delta A_{\mu}}{\Delta S^{\nu\rho}} = \frac{1}{2}R_{\mu\nu\rho}^{\lambda}A_{\lambda} . \quad (2.7)$$

In a flat spacetime, the difference ΔA_{μ} , for a covariant vector transported around a closed path, is automatically zero, since it is a parallel transport. Thus, the Riemann curvature tensor is a measure of the curvature of the spacetime, i.e., a departure from 'flatness' of the spacetime.

This definition of the curvature tensor is equivalent to the expression of the difference of second covariant derivatives of the vector field, as follows:

$$A_{\mu;\nu;\rho} - A_{\mu;\rho;\nu} = R_{\mu\nu\rho}{}^{\lambda} A_{\lambda} , \quad (2.8)$$

where the covariant derivative of the vector field is given in (2.6). The added term that entails the affine connection term $\Gamma_{\mu\alpha}^{\rho}$ in (2.5), is required to make the vector field A_{μ} integrable in the curved spacetime as indicated in (2.6).

It follows from (2.6) and (2.8) that the form of the Riemann curvature tensor is

$$R_{\mu\nu\rho}{}^{\lambda} = \partial_{\nu}\Gamma_{\mu\rho}^{\lambda} - \partial_{\rho}\Gamma_{\mu\nu}^{\lambda} + \Gamma_{\alpha\nu}^{\lambda}\Gamma_{\mu\rho}^{\alpha} - \Gamma_{\alpha\rho}^{\lambda}\Gamma_{\mu\nu}^{\alpha} . \quad (2.9)$$

The Ricci tensor in the field equation (2.3) is a contraction of the curvature tensor with the metric tensor, viz., $R_{\mu\nu} = g^{\lambda\alpha} R_{\alpha\mu\nu\lambda} = R_{\mu\nu\lambda}{}^{\lambda}$.

It is readily verified that $R_{\mu\nu} = R_{\nu\mu}$, i.e., that it is a symmetric second-rank tensor. This is seen as follows:

$$\begin{aligned} R_{\mu\nu} &= R_{\mu\nu\lambda}{}^{\lambda} = g^{\lambda\alpha} R_{\alpha\mu\nu\lambda} = g^{\lambda\alpha} R_{\nu\lambda\alpha\mu} \\ &= -g^{\lambda\alpha} R_{\lambda\nu\alpha\mu} = g^{\lambda\alpha} R_{\lambda\nu\mu\alpha} = R_{\nu\mu\alpha}{}^{\alpha} = R_{\nu\mu} . \end{aligned}$$

Thus, $R_{\mu\nu}$, like $g_{\mu\nu}$, has 10 independent components. The Einstein field equations (2.3) thus constitute a symmetric second-rank tensor equation corresponding to 10 independent equations in the curved spacetime.

2.5 Einstein's Equations from the Variational Principle

The Einstein field equations follow from Hamilton's principle of least action applied to the field formalism. It is based on the vanishing of the variation $\delta(S_G + S_M) = 0$, where $S_G = \int L_G \sqrt{-g} d^4x$ is the action function and $L_G = R$ is the Lagrangian density for the metrical field, leading to the laws of gravitation. $S_M = \int L_M \sqrt{-g} d^4x$ is the action function for the matter part of the field formalism and L_M is the corresponding matter field Lagrangian density, leading to the source term $\kappa T_{\mu\nu}$.

Thus, the variational problem is

$$\delta S = \delta(S_G + S_M) = 0 , \quad (2.10)$$

where

$$\delta S_G = \delta \int R \sqrt{-g} d^4x = \delta \int g^{\mu\nu} R_{\mu\nu} \sqrt{-g} d^4x . \quad (2.11)$$

The invariant volume element in the spacetime is

$$\sqrt{-g}d^4x = \sqrt{-g'}d^4x' ,$$

where $-g = |g_{\mu\nu}|$ is the metric density, and the vertical bars denote the determinant. Thus,

$$\delta S_G = \int (R_{\mu\nu} \sqrt{-g} \delta g^{\mu\nu} + R_{\mu\nu} g^{\mu\nu} \delta \sqrt{-g} + g^{\mu\nu} \sqrt{-g} \delta R_{\mu\nu}) d^4x . \quad (2.12)$$

The second term in this integrand contains the term

$$\delta \sqrt{-g} = -\frac{1}{2\sqrt{-g}} \delta(-g) = -\frac{1}{2} \sqrt{-g} g_{\mu\nu} \delta g^{\mu\nu} .$$

The latter equation comes from the following:

$$\frac{\delta(-g)}{\delta g_{\mu\nu}} = \text{cofactor } g_{\mu\nu} = (-g) g^{\mu\nu} ,$$

where the cofactor of $g_{\mu\nu}$ is the 3×3 determinant of the matrix of $g_{\mu\nu}$ with the μ th row and the ν th column removed.

It then follows that the first two terms in the integrand above are

$$\left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right) \sqrt{-g} \delta g^{\mu\nu} .$$

To show that the third term in the integrand in (2.12) vanishes, we will use the technique of assuming the flat spacetime limit, and then take the global extension of the result in the curved spacetime. In the flat spacetime, $\partial_\lambda g_{\mu\nu} = 0$, $\Gamma_{\mu\nu}^\rho = 0$. In this limit, the third term in the integrand of (2.12) is

$$\begin{aligned} g^{\mu\nu} \delta R_{\mu\nu} &= -g^{\mu\nu} \left[\delta(\partial_\rho \Gamma_{\mu\nu}^\rho) - \delta(\partial_\nu \Gamma_{\mu\rho}^\rho) \right] \\ &= -g^{\mu\nu} \left[\partial_\rho (\delta \Gamma_{\mu\nu}^\rho) - \partial_\nu (\delta \Gamma_{\mu\rho}^\rho) \right] \\ &= -g^{\mu\nu} \partial_\rho (\delta \Gamma_{\mu\nu}^\rho) + g^{\mu\rho} \partial_\rho (\delta \Gamma_{\mu\nu}^\nu) . \end{aligned}$$

The indices ν, ρ have been interchanged in the last term above (since these are summed over all four of their values).

As we have discussed earlier, the term $\Gamma_{\mu\nu}^\rho A_\rho dx^\nu$ denotes the addition to the ordinary derivative of a vector under parallel transfer through a displacement dx^ν . Thus, the term entailing the variation of the affine connection, $\delta \Gamma_{\mu\nu}^\rho A_\rho dx^\nu$ denotes a difference between two vectors for the result of two parallel displacements, one with the affine connection Γ and the other with the affine connection $\Gamma + \delta \Gamma$:

$$A + (\Gamma + \delta\Gamma)Adx - (A + \Gamma Adx) = \delta\Gamma Adx .$$

This entails a variation $\delta\Gamma$ at the same point in spacetime. Thus, $\delta\Gamma_{\mu\nu}^{\rho} A_{\rho} dx^{\nu}$ is a covariant entity – it is a four-vector. It then follows that, while the affine connection $\Gamma_{\mu\nu}^{\rho}$ is not a tensor, the variation of the affine connection $\delta\Gamma_{\mu\nu}^{\rho}$ does transform as a third rank tensor. Thus,

$$w^{\rho} = g^{\mu\nu} \delta\Gamma_{\mu\nu}^{\rho} - g^{\mu\rho} \delta\Gamma_{\mu\nu}^{\nu} \quad (2.13)$$

is a four-vector.

It then follows that in the limit of a flat spacetime, where $\partial_{\rho} g^{\mu\nu} = 0$ and $\Gamma_{\mu\nu}^{\rho} = 0$,

$$g^{\mu\nu} \delta R_{\mu\nu} = \partial_{\rho} w^{\rho} = g^{\mu\nu} \partial_{\rho} \delta\Gamma_{\mu\nu}^{\rho} - g^{\mu\rho} \partial_{\rho} \delta\Gamma_{\mu\nu}^{\nu} .$$

Since $\partial_{\rho} w^{\rho} = w^{\rho}_{;\rho}$, because this is a scalar, and with

$$w^{\rho}_{;\rho} = \frac{1}{\sqrt{-g}} \partial_{\rho} (\sqrt{-g} w^{\rho}) ,$$

it follows that in its global extension of the ordinary divergence to the covariant divergence (from the flat spacetime to the curved spacetime),

$$\int g^{\mu\nu} \delta R_{\mu\nu} \sqrt{-g} d^4x = - \int \partial_{\rho} (\sqrt{-g} w^{\rho}) d^4x . \quad (2.14)$$

The details of the Riemannian calculus are spelled out in [117], Chaps. 2 and 6.

Using Gauss' theorem, the integral on the right-hand side of (2.14) in 4-space, over the four-dimensional volume V , is equal to the surface integral over the 3-dimensional surface S that contains V :

$$\int_S (\sqrt{-g} w^{\rho}) n_{\rho} dS = 0 ,$$

where n_{ρ} is the unit normal vector to the surface element dS . The reason that the latter surface integral vanishes is that w^{ρ} (2.13) depends on the variation $\delta\Gamma$ on the surface S that encloses the volume V , and according to the variational calculus, $\delta\Gamma = 0$ on S .

It then follows that the third term in the integral in (2.12) vanishes identically. Thus we are left with the variation

$$\begin{aligned} \delta S_G &= \int \left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right) \delta g^{\mu\nu} \sqrt{-g} d^4x \\ &= -\delta S_M = \int \kappa T_{\mu\nu} \delta g^{\mu\nu} \sqrt{-g} d^4x . \end{aligned}$$

Equating the integrands in all spacetime, we arrive at the Einstein field equation (2.3), viz.,

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \kappa T_{\mu\nu} , \quad (2.15)$$

as indicated earlier in the text.

Since the field equation (2.15) has the form of a second-rank symmetric tensor equation, it corresponds to 10 independent (component) equations in the curved spacetime, with the Riemannian curvature tensor defined in (2.9). It is noted here that the underlying group of general relativity is a 16 parameter Lie group. Thus one should expect to have 16 relations at each spacetime point to define the spacetime. The second-rank tensor formalism of general relativity (2.15), which entails only 10 relations, is then not a complete expression of the theory. We will see in the next chapter that the reason for this is that the formalism (2.15) is reducible – that it is symmetric under the discrete reflections in space and time, which is not required by the group of general relativity, as well as the continuous transformations, which is the only requirement of the principle of general covariance. The procedure in the next chapter will then be to remove the discrete transformations, leading to a factorization of Einstein's tensor equations to a quaternion form, wherein there are indeed 16 independent relations at each spacetime point, to define the spacetime. The extra degrees of freedom in the metrical field will then be shown to lead to a natural incorporation of the field equations of electromagnetism with those of gravity in the quaternion formalism for the theory of general relativity.

2.6 The Vacuum Equation

In a matterless universe, the energy–momentum tensor $T_{\mu\nu} = 0$ everywhere. This corresponds to a reduction of the Einstein field equation (2.15) to the vacuum equation

$$R_{\mu\nu} = 0 \quad (2.16)$$

everywhere. In view of the meaning of the Einstein field equations as a way of interpreting the geometry of spacetime in terms of the matter content of a closed system, and vice versa, it follows that the only exact, valid solution of the vacuum equation (2.16) is the flat

spacetime constant solution, $g^{00} = 1$, $g^{kk} = -1$ for $k = 1, 2, 3$ the three spatial directions, and $g^{\mu\nu} = 0$ for $\mu \neq \nu$.

Yet, there is a solution of the vacuum equation (2.16) – the Schwarzschild solution – that does not correspond with the flat spacetime and leads to predictions of this equation that are in agreement with empirical facts about gravitation. These are the crucial tests of general relativity referred to earlier: the perihelion precession of Mercury's orbit, the gravitational redshift and the bending of a beam of starlight as it glances the rim of the Sun. In view of the success of the Schwarzschild solution to predict these empirical facts, one must look upon these exact solutions of the vacuum equation (2.16) as an approximation for the actual solutions of the Einstein equation (2.15), rather than an exact solution of Einstein's theory itself.

These conclusions are due, in part, to the nonlinear nature of the field equations. According to this field theory of matter, there is no actual 'outside' separated from an 'inside' of a matter distribution, such as a star. There is simply a single continuum, which 'peaks' in particular spatial regions of the universe – exhibiting the distinguishable modes of the single continuum that is, in principle, the universe. This is unlike a linear theory, such as Newton's theory of gravitation, where the solutions of the vacuum, outside of a star (represented by Laplace's equation $\nabla^2\phi = 0$), and inside of the star (represented by Poisson's equation $\nabla^2\phi = 4\pi\rho$, where ρ is the matter density of the star). What is then done is to equate the inside solution to the outside solution at the boundary of the star, to determine the integration constants in the problem. This is based on the feature of linear differential equations that the solutions of the equation with sources are a linear superposition of the solutions of the equation without sources. But this is not the case with nonlinear equations, where there is no 'inside' and 'outside' in an exact form! An example of an explicitly linear formalism is Maxwell's equations for electromagnetism. Here the source-free radiation solutions are used to make up a linear superposition of solutions to express the solutions of the field equations with sources – as in quantum electrodynamics – where one expresses the solutions of the equation with sources as a perturbation expansion in terms of the radiation solutions of the source-free equations.

A nonlinear field theory does not allow this procedure because a solution of one nonlinear equation (say with sources) is not a linear superposition of the solutions of another nonlinear equation (say one without sources).

Nevertheless, at times it could be useful to regard the solutions of one nonlinear differential equation, such as one without sources, as an approximation for the solutions of another, different nonlinear differential equation, such as one with sources. This is a good example to demonstrate the correspondence between Newton's equations of gravitation and the solutions of the vacuum equation as a linear approximation for Einstein's equations. This will be shown in the next section in our discussion of the geodesic equation in general relativity, and the use of the Schwarzschild solution for the vacuum equation (2.16) and its comparison with Newton's equation for gravitation.

2.7 The Geodesic Equation

The solutions of the geodesic equation determine the paths of minimal separation between any points of spacetime. A feature of general relativity theory is that a 'free body' must move along such a path. That is, it would take external energy to do the work to remove the body from the geodesic path. The geodesics of Euclidean geometry are straight lines, those of Riemannian geometry are curves. Thus, the idea that the 'natural path' of a free body in a Riemannian spacetime is a particular curve is a generalization of Galileo's principle of inertia, prescribing the path of the free body to be a straight line – the geodesic of a Euclidean spacetime. Of course, in his day, Galileo had no idea that there could be any geometry other than that of Euclid!

The family of geodesic curves of a Riemannian spacetime defines a 'curved spacetime'. The minimal path in this spacetime is determined by the vanishing of the variation of an arbitrary line integral:

$$\delta \int ds = 0 ,$$

where the arbitrary limits of the line integral are s_a and s_b . The argument is that, since

$$\delta \int ds = 0 \implies \delta \int ds^2 = 0 ,$$

it follows from $ds^2(\text{Riemann})$ in (2.1) that

$$0 = \delta \int g_{\mu\nu} dx^\mu dx^\nu = \int (\delta g_{\mu\nu}) dx^\mu dx^\nu + 2 \int g_{\mu\nu} dx^\mu \delta(dx^\nu) . \quad (2.17)$$

The factor of 2 appears above because both μ and ν are summed over while the variation of the differentials dx^μ and dx^ν each yield the same result.

Using the fact that the differential and the variation operations commute, viz.,

$$\delta(dx^\mu) = d(x^\mu + \delta x^\mu) - dx^\mu = d(\delta x^\mu) ,$$

and the definition of the variation of $g_{\mu\nu}$ with respect to the coordinate x^λ as

$$\delta g_{\mu\nu} = (\partial_\lambda g_{\mu\nu}) \delta x^\lambda ,$$

equation (2.17) may be expressed in the form

$$\int \left(\frac{1}{2} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} \partial_\lambda g_{\mu\nu} \delta x^\lambda + g_{\mu\nu} \frac{dx^\mu}{ds} \frac{d(\delta x^\nu)}{ds} \right) ds = 0 . \quad (2.18)$$

(Henceforth all line integrals will assume arbitrary limits s_a and s_b .) The second part of the left-hand side of (2.18) may be integrated by parts to give

$$\begin{aligned} g_{\mu\nu} \frac{dx^\mu}{ds} \delta x^\nu (\text{between endpoints } s_a \text{ and } s_b) - \int \frac{d}{ds} \left(g_{\mu\nu} \frac{dx^\mu}{ds} \right) \delta x^\nu ds \\ = 0 - \int \left(g_{\nu\lambda} \frac{d^2 x^\lambda}{ds^2} + \partial_\nu g_{\mu\lambda} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} \right) \delta x^\nu ds . \end{aligned} \quad (2.19)$$

The first term on the left-hand side becomes zero on the right-hand side because, by definition, the variation of the global coordinates δx^ν vanishes at the endpoints. Substituting (2.19) into (2.17) and expressing the last term in (2.19) in the equivalent form

$$\frac{1}{2} (\partial_\nu g_{\mu\lambda} + \partial_\mu g_{\nu\lambda}) \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} ,$$

the vanishing of the variation of the line integral over ds^2 implies the vanishing of the integrand. Since $\delta x^\nu \neq 0$ and with the second kind of affine connection defined by $\Gamma_{\lambda,\mu\nu} = g_{\rho\lambda} \Gamma_{\mu\nu}^\rho$, we have

$$g_{\mu\lambda} \frac{d^2 x^\lambda}{ds^2} + \Gamma_{\lambda,\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} = 0 . \quad (2.20)$$

This is one form of the geodesic equation. The second form is obtained by multiplying (2.20) by the reciprocal of the metric tensor, that is, $g^{\rho\lambda}$, and summing over λ to give

$$\frac{d^2 x^\rho}{ds^2} + \Gamma_{\mu\nu}^\rho \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} = 0 . \quad (2.21)$$

The solutions of the geodesic equation (2.21) are a mapping of a family of paths of minimal separation between any two points of spacetime. It defines the curved spacetime. This equation is a generalization of Newton's first law of motion. Expressed with Euclidean geometry, Newton's first law of motion states that, when there is no external influence (i.e., 'force') on a material body then its acceleration in space is zero:

$$\frac{d^2 \mathbf{r}}{ds^2} = 0 .$$

The solution of this second order, ordinary differential equation is the (vector) displacement

$$\mathbf{r}(t) = \mathbf{v}t + \mathbf{r}_0 ,$$

where $\mathbf{v}$ is the body's constant velocity (a special case is rest, $\mathbf{v} = 0$) and $\mathbf{r}_0$ is the body's position at the time $t = 0$. The latter are the two required boundary conditions to specify a unique trajectory for the body.

In the tensor form of general relativity, should we attach a clock to the origin (the location of the moving body), then the differential metric ds entails no spatial part. Its square is $ds^2 = g_{00}c^2dt^2$.

The interval ds is called the proper time of the body. The geodesic equation in the spatial sector is then

$$\frac{d^2 x^k}{ds^2} = -\Gamma_{\mu\nu}^k \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} \quad (k = 1, 2, 3) . \quad (2.22)$$

Note that both equations of motion of a body – Newton's law and the geodesic equation in general relativity – are independent of the inertial mass of the body, as Galileo anticipated in his derivation of $g = \text{constant}$ for the acceleration of a body in free fall.

The equation of motion (2.22) in general relativity is along a natural curved path, by virtue of $\Gamma_{\mu\nu}^k \neq 0$ (anywhere). Thus it predicts that the curved path of starlight, as its trajectory glances the edge of the Sun, has a particular shape. Eddington and his group observed it, in agreement with the theory, in 1918 (both qualitatively and quantitatively). This equation also predicts two of the other crucial tests of general relativity – the perihelion precession of a planetary orbit and the gravitational redshift, as discussed earlier.

All three of these tests of general relativity used the geodesic equation combined with the Schwarzschild solution of the vacuum equation. The details of the three predictions are spelled out in [117].

2.8 The Schwarzschild Solution and the Newtonian Limit

Assuming that the solutions of the vacuum equation (2.16) are spatially spherically symmetric, the time dependence of these solutions automatically transforms away (Birkhoff's theorem). Schwarzschild then found that a solution of the vacuum equation (2.16) yields the Riemannian (squared) metric as

$$ds^2 = \left(1 - \frac{2\alpha}{r}\right) (dx^0)^2 - \frac{dr^2}{1 - 2\alpha/r} - r^2 d\Omega^2, \quad (2.23)$$

where α is a constant of integration with the dimension of length. We see here that, when $\alpha/r = 0$, this solution reduces to the flat spacetime metric. The coordinates here are $x^0 = ct$, the time measure, and spatial coordinates expressed in the spherical coordinate system [r is the radial coordinate, θ bounded by $(0, \pi)$ is the polar angle, and ϕ bounded by $(0, 2\pi)$ is the azimuthal angle]. The solid differential angle is $d\Omega = \sin^2 \theta d\theta d\phi$.

In the co-moving coordinate system, where the test body is at the spatial origin, this metric reduces to

$$ds = \left(1 - \frac{2\alpha}{r}\right)^{1/2} cdt.$$

The quantity ds/c is the proper time measure of the body. The flat space limit corresponds to being at an infinite distance from the center of the force, such as the effect of the Sun on a test body, as $r \rightarrow \infty$, and $ds \rightarrow cdt$.

In the next approximation, we assume that the ratio $2\alpha/r \ll 1$, where it follows from the binomial expansion that

$$\frac{ds}{dt} \approx \left(1 - \frac{\alpha}{r}\right) c.$$

In this case,

$$\frac{dx^\mu}{ds} = \frac{dx^\mu}{dt} \frac{dt}{ds} = \frac{1}{c} \frac{dx^\mu}{dt} \left(1 - \frac{\alpha}{r}\right)^{-1}.$$

The individual terms are then

$$\frac{dx^0}{ds} = \left(1 - \frac{\alpha}{r}\right)^{-1}, \quad \frac{dx^k}{ds} = \frac{v^k}{c} \left(1 - \frac{\alpha}{r}\right)^{-1} \quad (k = 1, 2, 3). \quad (2.24)$$

When $v^k/c \ll 1$, then $dx^0/ds \gg dx^k/ds$ and it follows that

$$\Gamma_{\mu\nu}^{\rho} \frac{dx^{\mu}}{ds} \frac{dx^{\nu}}{ds} \approx \Gamma_{00}^{\rho} .$$

Also, with $\alpha/r \ll 1$, $ds/dt \approx c$ and

$$\frac{d^2 x^{\rho}}{ds^2} = \frac{d}{ds} \frac{dx^{\rho}}{dt} \frac{dt}{ds} = \frac{d}{dt} \left(\frac{dx^{\rho}}{dt} \frac{dt}{ds} \right) = \frac{1}{c^2} \frac{d^2 x^{\rho}}{dt^2} .$$

In this case, the geodesic equation (2.21) becomes

$$\frac{d^2 x^{\rho}}{dt^2} = c^2 \Gamma_{00}^{\rho} .$$

With the approximation $\alpha/r \ll 1$, it follows that

$$g_{00} = \left(1 - \frac{\alpha}{r}\right) , \quad \Gamma_{00}^{\rho} = g^{\rho\lambda} \Gamma_{\lambda,00} = \frac{1}{2} g^{\rho\lambda}(\mathbf{L}) \partial_{\lambda}(\alpha/r) ,$$

where $g^{\rho\lambda}(\mathbf{L})$ is the flat space (Lorentz) limit (2.1) of the metric tensor. With $\rho = 0$, and r being independent of the time measure, $g^{00}(\mathbf{L}) = 1$, $\partial_0(\alpha/r) = 0$, so that $\Gamma_{00}^0 = 0$.

Thus, with these conclusions in this approximation, $dx^0/dt = 0 \Rightarrow x^0 = K_1 t + K_2$. The integration constants here are $K_1 = c$, $K_2 = 0$, and thus $x^0 = ct$.

With $k = 1, 2, 3$, $g^{kk}(\mathbf{L}) = -1$, $\partial_k(\alpha/r) \equiv \partial_r(\alpha/r) = -\alpha/r^2$. It then follows that

$$\Gamma_{00}^1 = \frac{\alpha}{2} \partial_r(1/r) = -\frac{\alpha}{r^2} .$$

The geodesic equation (2.22) in the radial coordinate r is then

$$\frac{d^2 r}{dt^2} = -\frac{c^2 \alpha}{r^2} . \tag{2.25}$$

With Einstein's interpretation of the geodesic equation as the equation of motion of a test body in the curved space, a comparison with Newton's theory for the acceleration produced by an external force yields

$$\frac{d^2 \mathbf{r}}{dt^2} = \frac{\mathbf{F}}{m} = -\nabla \Phi = -\frac{GM}{r^3} \mathbf{r} , \tag{2.26}$$

where Φ is Newton's gravitational potential and M is the mass of the body that influences the test body with mass m . G is Newton's universal gravitational constant.

It then follows, in this approximation, that $-c^2\alpha/r^2 = -GM/r^2$. Thus, the integration constant in the Schwarzschild solution is GM/c^2 . Inserting this into the Schwarzschild metric, we have

$$ds^2 = \left(1 - \frac{2GM}{c^2 r}\right) (dx^0)^2 - \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2 - r^2 d\Omega^2. \quad (2.27)$$

Thus, when $v/c \ll 1$ and $\alpha/r \ll 1$, the form of the geodesic equation (2.25) matches Newton's equation of motion (2.26). This is a Newtonian approximation for the result of general relativity, under the conditions that lead to the Schwarzschild solution.

The integration constant α is called the Schwarzschild radius. Substituting the universal constant of Newton, $G = 6.67 \times 10^{-8} \text{ dyn cm}^2 \text{ g}^{-2}$ and the value of M for the Sun, α has a value that is a few kilometers. Thus the assumption used above, that α is much less than the radius r of the Sun, is valid. It is in this limit that we see that Newton's theory of universal gravitation corresponds with Einstein's theory when it is applied to the vacuum equation $R_{\mu\nu}(r) = 0$.

The reader should take note again that the vacuum equation for general relativity is not an exact equation for the spacetime associated with, say, the existence of the Sun; it is only an approximation for the full expression of Einstein's equations (2.3). Thus, it cannot be said that the Schwarzschild metric (2.27) is based on an exact solution of Einstein's field equations. As discussed earlier, this is because a massive body, such as the Sun, is in reality represented by a continuous field everywhere, without actual boundaries – there is no absolute 'inside' or absolute 'outside' of the Sun, or of any mass distribution. Of course, we have the illusion that it is a discrete entity because the mass field of the Sun 'peaks' in a particular region of the space of our vision. But in this theory there is no actual 'vacuum', as there is in the model of Newton's theory of gravitation.

In the latter case, one obtains the gravitational potential that would act on a test body by solving the vacuum equation, Laplace's equation $\nabla^2\phi = 0$, and the equation for the inside of the star, Poisson's equation $\nabla^2\phi = 4\pi\rho$, where ρ is the mass density of the star, and then equating these solutions at the boundary of the star to evaluate the integration constants – to give a unique solution for the gravitational potential of the star. But in the continuous field theory of general relativity, there is no sharp boundary of the star, and therefore no inside and outside of the star, in objective terms!

The Schwarzschild solution has been used in astrophysics to describe a black hole – a star with such a great density that the bound-

ary of the star is at the Schwarzschild radius $r = r_S = 2\alpha = 2GM/c^2$, where M is the mass of this black hole and r_S is its radius. We see in the solution (2.27) that this corresponds to a singularity in space. But in general relativity theory, as we will discuss in Chap. 8, the solutions must be regular – without singularities anywhere. The problem of the black hole will be discussed further in Chap. 8, as well as other problems from astrophysics that relate to the topic of this book, quantum mechanics and gravity.

In the next chapter, we will show that a unified theory develops when one fully exploits the algebraic logic of general relativity theory, giving rise to a factorization of the tensor field equations (10 relations) into a quaternion field representation (16 relations), thereby demonstrating a unification of the gravitational and the electromagnetic manifestations of matter.

3. A Unified Field Theory

A ‘unified field theory’ refers to the explanation of all possible forms of force exerted by matter on matter as well as the inertial response of the matter that is acted upon, based on a single continuous field formalism. The appearance in nature of one type of force or another and the inertia of matter depend on the type of physical circumstances experienced by matter.

The principle of covariance – the underlying idea of the theory of general relativity – asserts that the symmetry that specifies the objectivity of all of the laws of nature entails a continuous set of changes of the space and time measures that leave the laws covariant, i.e., unchanged in form. This is summarized explicitly in terms of the underlying differential metric of spacetime, denoted ds , which stays unchanged (invariant) under the same space and time transformations that leave the forms of the laws unchanged.

The logic of the spacetime language has both geometrical and algebraic features. As we go from classical physics to general relativity physics, the type of geometry changes from Euclidean geometry to Riemannian geometry. The latter is a change from the flat spacetime to a curved spacetime. The algebra that underlies the logic of relativity theory is expressed in terms of a symmetry group, the Einstein group in general relativity and the Poincaré group in special relativity. This is a set of continuous, analytic transformations (without reflections) that preserve the forms of the laws of nature. This requirement leads to the basis of the laws in terms of spinor and quaternion field variables.

The rejection of reflections in the space and time symmetry of the physical laws leads to a factorization of the 10-component Einstein tensor equations to a 16-component quaternion field equation, wherein 10 components prescribe the laws of gravity and 6 components prescribe the laws of electromagnetism. The basic field in this generalization is the four-vector $q^\mu(x)$, wherein each of the four components ($\mu = 0, 1, 2, 3$) is a 4-element quaternion, rather than a single-component real number.

The geodesic equation in quaternion form then implies a generalization of the time measure of a body that moves along a trajectory in space. In the standard description of motion, the ‘time’ is a single parameter t , whose continuous change specifies the motion of a body in space from point to point. The quaternion generalization of the time measure requires the specification of four numbers at each point of a spatial trajectory to four other numbers, in continuously extending the body along the spatial trajectory. Thus this is a generalization of the use of the time measure in the description of the motion of a body in space.

3.1 Einstein’s Field Theory in Quaternion Form

The underlying axiom of Einstein’s theory of general relativity is the principle of covariance, also called the principle of relativity. It is the assertion that all of the laws of nature must preserve their forms under transformations between any frames of reference in which the laws are expressed, from the view of any one of them. The forms of the laws, in turn, are expressed in terms of a language of space and time parameters. Although not the only possible language, it must be in one-to-one correspondence with any other language that is used. The space and time (or other language) parameters are a continuous set. The functions of the spacetime parameters that are the field variables – the solutions of the laws of nature – are a continuous set of functions of the space and time parameters.

A further restriction on the field solutions is that they must be analytic. This follows from Noether’s theorem – asserting that a necessary and sufficient condition for the existence of laws of conservation of energy and linear and angular momentum is that the spacetime parameters and the fields that solve the laws of nature are analytic – that they are differentiable to all orders. Such functions are called regular.

In a letter to D. Bohm in 1953, Einstein said the following: “When one is not starting from the correct elementary concepts, if, for example, it is not correct that reality can be described as a continuous field, then all my efforts are futile, even though the constructed laws are the greatest simplicity thinkable. The fact that logically simple field equations are necessarily nonlinear, and that a consistent field theory cannot permit singularities, makes it impossible, for the time being, to draw from the theory any conclusions which would allow it to be tested empirically. This does not, however, convince me of the incorrectness of the theory. It only shows our present mathematical methods

are insufficient.” (Einstein Archives, Jewish National and University Library, Jerusalem, Call No. 4 1576:8-053.)

Einstein's requirement is that the field solutions of the laws of nature are regular functions – that is, singularity-free – everywhere. A strong reason for this requirement is the insistence on the inclusion of the laws of the global extension of the conservation of energy, linear and angular momentum, such as $T_{\mu\nu}{}^{;\nu} = 0$. These equations take the form of the conventional conservation equation $\partial^\nu T_{\mu\nu} = 0$ in the flat spacetime limit.

The set of continuous, analytic transformations that underlie the covariance of the theory of special relativity is the 10-parameter Lie group, called the Poincaré group. The transformation group of the theory of general relativity is the 16-parameter Lie group called the Einstein group. A Lie group is a set of continuous, analytic transformations obeying the algebra of a group. The 16 essential parameters of the Einstein group are the 16 derivatives $\partial x^{\mu'}/\partial x^\nu$ in the curved spacetime, where $\mu, \nu = 0$ denotes the time coordinate and $\mu, \nu = 1, 2, 3$ are the spatial coordinates. The Einstein field equations (2.3) are covariant under the transformations of this Lie group, which is a requirement of the theory. But they are also covariant under extra transformations – the discrete reflections in space and time – which is not required of the theory. Thus, the covariance group of Einstein's tensor equations (2.3) is reducible. When the reflections are removed from the underlying group, we arrive at the irreducible form of the theory. To fully exploit the irreducible representations of the Einstein group, one must have the same number of relations at each spacetime point as the number of essential parameters of the Lie group. We see, then, that since Einstein's tensor field equations (2.3) entail 10, rather than 16 relations at each spacetime point, they must be ‘degenerate’, thus factorizable to 16 equations – the correct number of relations consistent with a 16-parameter Lie group.

It will be shown that it is this generalization to the irreducible symmetry group underlying general relativity theory that leads to a unified field theory, wherein the 16-component (factorized) field entails 10 components that correspond with gravity, according to Einstein's formulation, and the remaining 6 components correspond with the 3 components of the electric field and 3 components of the magnetic field, in accordance with Maxwell's field equations. Thus fully exploiting the Lie group that underlies the theory of general relativity yields a genuine unified field theory.

In his researches on a unified field theory, Einstein said [3]: “Not for a moment, of course, did I doubt that this formulation [his field equation (2.3)] was merely a makeshift in order to give the general principle of relativity a preliminary closed expression. For it was nothing more than a theory of the gravitational field, which was somewhat artificially isolated from a total field of as yet unknown structure.” He also said: “To remain with the narrower group and at the same time to have the relativity theory of gravitation based upon the more complicated tensor structure implies a naive consequence.”

Indeed, if the (unwanted) reflection symmetry elements should be removed from the reducible group of symmetry of his equations (2.3), these equations do factorize to an irreducible form in terms of 16 relations at each spacetime point. The resulting irreducible representation of the Einstein group obeys the algebra of quaternions and their basis functions are spinors in a curved spacetime.

We may exploit these ideas of generalizing Einstein’s tensor formalism by starting out with the invariant differential metric in a Riemannian spacetime:

$$ds = q^\mu(x)dx_\mu, \quad (3.1)$$

where q^μ is a four-vector field (dependent on all of the coordinates x of the spacetime) in which all four vector components are quaternion-valued, rather than being real-number-valued. Thus the generalized metrical field q^μ is a 16-component variable. The invariant metric ds of (3.1) is a factorization of the Riemannian metric,

$$ds^2 = g^{\mu\nu}dx_\mu dx_\nu = ds ds^*, \quad (3.2)$$

where the asterisk denotes the quaternion conjugate of ds (corresponding to time reversal or space reflection).

The correspondence between the metric tensor and the quaternion metric is then, from (3.1) and (3.2),

$$g^{\mu\nu} \Longleftrightarrow -\frac{1}{2}(q^\mu q^{*\nu} + q^\nu q^{*\mu}), \quad (3.3)$$

where the minus sign is chosen because of normalization.

Note that (3.3) is not a definition of the metric tensor $g^{\mu\nu}$. Rather, it is an identification between the more fundamental quaternion metric and the tensor metric of the older formulation. In this generalization of the Riemannian metric, the metric tensor is replaced by the quaternion metrical field, as the more fundamental expression of the geometry of the spacetime.

An important feature of the quaternion-valued variable, as distinguished from the real-number-valued variable, is that the quaternions form a nonabelian algebra, i.e., wherein the variables do not commute under multiplication. This feature plays an important role in the formulation of quantum mechanics from a generally covariant field theory of the inertia of matter, to be demonstrated in Chap. 4.

The flat spacetime limit of the metric tensor (2.2) corresponds to the following limits on the quaternion components

$$q^\mu(x) \longrightarrow \sigma^\mu = (\sigma^0, \sigma^k) \quad (k = 1, 2, 3), \quad (3.4)$$

where σ^0 is the unit 2-dimensional matrix and σ^k are the three Pauli matrices:

$$\sigma^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

This quaternion basis, in terms of the four elements, the unit two-dimensional matrix and the three Pauli matrices, is a four-dimensional generalization of the two-dimensional complex number system, with the basis elements $(1, i = \sqrt{-1})$. The algebra of quaternions follows from the commutation rules

$$\sigma^0 \sigma^k = \sigma^k \sigma^0, \quad \sigma^k \sigma^j = -\sigma^j \sigma^k = \varepsilon_m^{kj} \sigma^m,$$

where $k \neq j \neq m = 1, 2, 3$, and the Levi-Civita symbol is $\varepsilon_m^{kj} = +1$ for $k = 1, j = 2, m = 3$, and any even permutation of these indices, otherwise it is equal to -1 .

3.2 Spin Affine Connection

Just as a vector field is not integrable in a curved spacetime without the addition to its ordinary derivative of an affine connection term, as in (2.6), so the spinor field is not integrable in a curved spacetime without the addition to its ordinary derivative of a spin affine connection term. The covariant derivative of a spinor variable in the curved spacetime is

$$\eta_{;\mu} = \partial_\mu \eta + \Omega_\mu \eta, \quad (3.5)$$

where the spin affine connection is

$$\Omega_\mu = \frac{1}{4}(\partial_\mu q^\rho + \Gamma_{\tau\mu}^\rho q^\tau)q_\rho^* = -\frac{1}{4}q_\rho(\partial_\mu q^{*\rho} + \Gamma_{\tau\mu}^\rho q^{*\tau}) . \quad (3.6)$$

This is derived in [117], Chap. 3.

The two-dimensional matrix field Ω_μ is covariant as a four-vector, but it is not covariant in spinor space. Instead, it transforms in spinor space as

$$x_\mu \longrightarrow x'_\mu \implies \Omega_\mu \longrightarrow \Omega'_\mu = S\Omega_\mu S^{-1} - (\partial_\mu S)S^{-1} , \quad (3.7)$$

where $\{S(x)\}$ are the spinor transformations of the Einstein group:

$$S(\theta_\nu^\mu(x)) = \exp \frac{q_\mu q_\nu \theta_\nu^\mu(x)}{2} , \quad (3.8)$$

and θ_ν^μ are the x -dependent parameters that define the 16 continuous transformations of the Einstein group of general relativity [not summed in (3.8)]. Equation (3.8) is a global extension in the curved spacetime of the representations of the Poincaré group of special relativity, wherein the parameters θ_ν^μ are constants, rather than the continuous, analytic functions of the space and time parameters x .

3.3 The Quaternion Variables in a Riemannian Spacetime

Within the quaternion formalism, we have the normalization

$$q^\mu q_\mu^* = \text{invariant} = 4\sigma^0 . \quad (3.9)$$

It follows from this invariance that the covariant derivatives of the quaternion metrical fields must vanish identically. (This is analogous to the condition of the vanishing of the covariant derivatives of the metrical field $g_{\mu\nu}$ of Einstein's tensor formulation.) The quaternion metrical field itself is a second-rank spinor of the type

$$q^\mu \sim (\eta \otimes \bar{\eta})^\mu , \quad (3.10)$$

where the overline above denotes the complex conjugate of the spinor variable η .

It then follows that the difference between the second covariant derivatives of the quaternion fields must also vanish. Expressing these as the sum of changes in 'spinor space' and in 'configuration space', we have

$$\begin{aligned}
 0 &= q_{\mu;\rho;\lambda} - q_{\mu;\lambda;\rho} \\
 &= [(\eta_{\alpha;\rho;\lambda} - \eta_{\alpha;\lambda;\rho})\bar{\eta}_{\beta} + \eta_{\alpha}(\bar{\eta}_{\beta;\rho;\lambda} - \bar{\eta}_{\beta;\lambda;\rho})] + ([q_{\mu;\rho;\lambda}] - [q_{\mu;\lambda;\rho}]) .
 \end{aligned} \tag{3.11}$$

The first (bracketed) term above refers to the second-rank spinor behavior of the quaternion field, where $\alpha, \beta = 1, 2$ are the components of the first-rank spinor field η . The last (bracketed) term above entails the behavior of q_{μ} as a four-vector in configuration space. The latter difference is

$$[q_{\mu;\rho;\lambda}] - [q_{\mu;\lambda;\rho}] = R_{\kappa\mu\rho\lambda}q^{\kappa} , \tag{3.12}$$

by definition of the Riemann curvature tensor $R_{\kappa\mu\rho\lambda}$.

With the definition (3.6) of the spin affine connection, the spin curvature tensor $K_{\rho\lambda}$ is defined by

$$\begin{aligned}
 \eta_{;\rho;\lambda} - \eta_{;\lambda;\rho} &\equiv K_{\rho\lambda}\eta = (\partial_{\lambda}\Omega_{\rho} + \Omega_{\lambda}\Omega_{\rho} - \partial_{\rho}\Omega_{\lambda} - \Omega_{\rho}\Omega_{\lambda})\eta \\
 &= (\Omega_{\lambda;\rho} - \Omega_{\rho;\lambda})\eta .
 \end{aligned} \tag{3.13}$$

Substituting (3.12) and (3.13) into (3.11), we have the following relation (suppressing the spinor indices α, β):

$$K_{\rho\lambda}q_{\mu} + q_{\mu}K_{\rho\lambda}^{\dagger} = -R_{\kappa\mu\rho\lambda}q^{\kappa} , \tag{3.14}$$

where the symbol $\dagger$ denotes the Hermitian adjoint of the function.

In a similar fashion, the vanishing of the covariant derivatives of the conjugated quaternion variables yields the following relation that accompanies (3.14):

$$K_{\rho\lambda}^{\dagger}q_{\mu}^{*} + q_{\mu}^{*}K_{\rho\lambda} = R_{\kappa\mu\rho\lambda}q^{*\kappa} . \tag{3.15}$$

Multiplying (3.14) on the right with the conjugated quaternion q_{γ}^{*} and (3.15) on the left with q_{γ} , adding the resulting equations and using the orthogonality identity,

$$q_{\gamma}q^{*\kappa} + q^{\kappa}q_{\gamma}^{*} = -2\sigma_0\delta_{\gamma}^{\kappa} ,$$

the correspondence with the Riemann curvature tensor is as follows:

$$\sigma_0 R_{\kappa\mu\rho\lambda} = \frac{1}{2}(K_{\rho\lambda}q_{\mu}q_{\kappa}^{*} - q_{\kappa}q_{\mu}^{*}K_{\rho\lambda} + q_{\mu}K_{\rho\lambda}^{\dagger}q_{\kappa}^{*} - q_{\kappa}K_{\rho\lambda}^{\dagger}q_{\mu}^{*}) . \tag{3.16}$$

The correspondence with the Ricci tensor is obtained by contracting the Riemann curvature tensor (3.16) with the metric tensor as follows:

$$\begin{aligned}
 \sigma_0 g^{\mu\lambda}R_{\mu\kappa\rho\lambda} &\equiv \sigma_0 R_{\kappa\rho} \\
 &= \frac{1}{2}(K_{\rho\lambda}q^{\lambda}q_{\kappa}^{*} - q_{\kappa}q^{*\lambda}K_{\rho\lambda} + q^{\lambda}K_{\rho\lambda}^{\dagger}q_{\kappa}^{*} - q_{\kappa}K_{\rho\lambda}^{\dagger}q^{*\lambda}) .
 \end{aligned} \tag{3.17}$$

Finally, the scalar curvature field follows from a further contraction of the Ricci tensor (3.17) with the metric tensor, $g^{\kappa\rho}R_{\kappa\rho} = R$, giving

$$\sigma_0 R = \frac{1}{2}(K_{\rho\lambda}q^\lambda q^{*\rho} - q^\rho q^{*\lambda}K_{\rho\lambda} + q^\lambda K_{\rho\lambda}^\dagger q^{*\rho} - q^\rho K_{\rho\lambda}^\dagger q^{*\lambda}). \quad (3.18)$$

3.4 Derivation of the Quaternion Metrical Field Equations from the Principle of Least Action

The Lagrangian density that leads to the metrical field equation is taken to be the trace of the scalar curvature (3.18) (with respect to the spinor indices) as follows:

$$L_E = \text{Tr}(R)\sqrt{-g} = \frac{1}{2}\text{Tr}(q^{*\rho}K_{\rho\lambda}q^\lambda + \text{h.c.})\sqrt{-g}, \quad (3.19)$$

where h.c. stands for the hermitian conjugate of the preceding term.

To derive the quaternion form of the metrical field equations, we take the Lagrangian density L_E (3.19) to be dependent on the variational parameters that are the quaternion and conjugated quaternion variables, as well as the components of the spin affine connection and their covariant derivatives. The subsequent minimization of the action functional with respect to these variables then yields two field equations – one that transforms as the quaternion q^μ and the other as its quaternion conjugate $q^{*\mu}$, along with the correct relation (3.6) between the quaternion metrical field variables and the spin affine connection. Such a separation into conjugated field equations appears because of the lack of reflection symmetry in the underlying Einstein group.

The variational calculation is then based on the extremum condition

$$\delta \int \left\{ \text{Tr}[R(q^\mu, q^{*\mu}), \Omega_\mu, \Omega_\mu^\dagger, \Omega_{\mu;\lambda}, \Omega_{\mu;\lambda}^\dagger] + L_M \right\} \sqrt{-g} d^4x = 0. \quad (3.20)$$

Recalling that $\sqrt{-g} = \sqrt{-\det g_{\mu\nu}}$, it follows that the quaternion derivatives of the metric density are

$$\begin{aligned} \frac{\partial \sqrt{-g}}{\partial q^{*\lambda}} &= \frac{\partial \sqrt{-g}}{\partial g^{\rho\lambda}} \frac{\partial \text{Tr} g^{\rho\lambda}}{\partial q^{*\lambda}} \\ &= -\frac{1}{2}\sqrt{-g}g_{\rho\lambda} \frac{\partial}{\partial q^{*\lambda}} \left[-\frac{1}{4}\text{Tr}(q^\rho q^{*\lambda} + q^\lambda q^{*\rho}) \right]. \end{aligned} \quad (3.21)$$

Since these derivatives refer to matrices, they are a matrix of derivatives with respect to the components of the quaternion variables.

If A and B are any two matrix fields, then,

$$\frac{\partial \text{Tr}(AB)}{\partial B_{ij}} = A_{ji} \quad \text{or} \quad \frac{\partial \text{Tr}(AB)}{\partial B} = A^{\text{tr}},$$

where A^{tr} denotes the transposed matrix. Inserting these relations into (3.21), we have

$$\frac{\partial \sqrt{-g}}{\partial q^{*\lambda}} = \left[\frac{q_\lambda}{8} \sqrt{-g} \right]^* . \quad (3.22)$$

Finally, with (3.19) and (3.22), and the antisymmetric feature of the spin curvature tensor $K_{\rho\lambda} = -K_{\lambda\rho}$, the following relation is obtained:

$$\frac{\partial L_E}{\partial q^\rho} = \overline{\left[-\frac{1}{4}(K_{\rho\lambda}^\dagger q^{*\lambda} + q^{*\lambda} K_{\rho\lambda}) + \frac{1}{8} R q_\rho^* \right]} \sqrt{-g} . \quad (3.23)$$

The conjugated relation that accompanies (3.23) is

$$\frac{\partial L_E}{\partial q^{*\rho}} = \overline{\left[\frac{1}{4}(q^\lambda K_{\rho\lambda}^\dagger + K_{\rho\lambda} q^\lambda) + \frac{1}{8} R q_\rho \right]} \sqrt{-g} , \quad (3.24)$$

where the overline above denotes the complex conjugate.

Since the Lagrangian density L_E does not have any explicit dependence on the derivatives of the quaternion variables, substitution of (3.23) and (3.24) into the Lagrange–Euler equation for any field variables Λ , viz.,

$$\frac{\partial L}{\partial \Lambda} = \left[\frac{\partial L}{\partial \Lambda_{;\mu}} \right]^{;\mu} ,$$

where $L = L_E + L_M$ is the total Lagrangian density, we have the following field equations:

$$\frac{1}{4}(K_{\rho\lambda} q^\lambda + q^\lambda K_{\rho\lambda}^\dagger) + \frac{1}{8} R q_\rho = k T_\rho , \quad (3.25a)$$

$$-\frac{1}{4}(K_{\rho\lambda}^\dagger q^{*\lambda} + q^{*\lambda} K_{\rho\lambda}) + \frac{1}{8} R q_\rho^* = k T_\rho^* , \quad (3.25b)$$

where the asterisk here denotes the quaternion conjugate. The field equations (3.25a) and (3.25b) – a factorization of Einstein’s tensor equations (2.3) – are the expression of the unified field theory. These are 16 relations at each spacetime point, rather than 10 – a number that is compatible with the 16 essential parameters of the underlying Lie group, the Einstein group of general relativity theory. All that

remains to be done is to demonstrate the relation between the quaternion fields, treated so far as variational parameters, and the spin affine connection, also treated as variational parameters. This will now be shown.

The right-hand side of (3.25a) – the source term for this metrical field equation – is the (complex conjugate of) the field that follows an extremum of the matter part of the action function S_M . The corresponding Lagrangian density L_M relates to the matter field equations, to be discussed in Chap. 4 and the electromagnetic field equations, to be discussed in Chap. 6, when varied with respect to the quaternion variables.

Finally, following Palatini's method, we can determine the relationship between the spin affine connection matrix field Ω_μ and the quaternion variables q_μ – by varying L_E with respect to the variational parameters Ω_μ and $\Omega_{\mu;\nu}$. We assume that the matter Lagrangian L_M is independent of Ω_μ . A relaxation of this assumption would lead to an even more general expression for the affine connection Ω_μ , indicating derivative coupling in the matter field Lagrangian. This is something that may be explored in the future.

Since the Lagrangian density L_E depends on Ω_μ through the spin curvature tensor $K_{\mu\nu}$, which in turn depends on $\Omega_{\mu;\nu}$ (3.13), the variational procedure yields the following result (using the chain rule):

$$\begin{aligned} \delta \int L_E \sqrt{-g} d^4x &= \int \left[\frac{\partial L_E}{\partial K_{\mu\nu}} \frac{\partial K_{\mu\nu}}{\partial \Omega_{\sigma;\tau}} \right]_{;\tau} \delta \Omega_\sigma \sqrt{-g} d^4x = 0 \\ &\Rightarrow \left(\frac{\partial L_E}{\partial K_{\mu\nu}} \right)_{;\mu} = (q^\nu q^{*\mu} - q^\mu q^{*\nu})_{;\mu} = 0. \end{aligned} \quad (3.26)$$

The only generally valid solution of (3.26) corresponds to $q^\nu_{;\mu} = 0$. This result in turn yields the connection (3.6) between the quaternion variables q^μ and the spin affine connection matrix field Ω_μ , which was to be shown.

Thus the combination of the field equations (3.25a) and (3.25b) and the restriction on Ω_μ (3.6) forms the field relationship between the metrical quaternion fields and the matter field variables – in its irreducible form, according to the underlying Einstein group of the theory of general relativity. It is the factorization of Einstein's tensor field equations (2.3) that we sought.

It will be shown next, by iteration, that one may recover from the equations (3.25a) and (3.25b) the equivalent of Einstein's tensor field equations for gravity (10 relations) and the remaining 6 field equations which lead, by taking its covariant divergence, to Maxwell's equations

for electromagnetism. Thus we see that the factorization (3.25a) and (3.25b) of the Einstein field equations (2.3) is indeed a unified field theory in the sense that Einstein and Schrödinger sought. It is a theory that unifies the phenomena of gravity and electromagnetism in a single, covariant metrical field – the 16-component quaternion variable $q^\mu(x)$. It is interesting to note at this juncture that the factorization (3.25a) and (3.25b) of Einstein’s field equations (2.3) follows for the same reason that the Klein–Gordon equation in special relativity factorizes into the space and time reflected spinor equations:

$$\square\psi = m^2\psi \quad (3.27)$$

$$\implies \sigma_\mu \partial^\mu \eta = -m\chi \quad \text{and} \quad \sigma_\mu^* \partial^\mu \chi = -m\eta, \quad (3.28)$$

where $\square = (\partial^0)^2 - \nabla^2$ is the d’Alembertian operator, and $\chi = \varepsilon\bar{\eta}$ is the time reversal (or space reflection) (sometimes called the Wigner time reversal) of the two-component spinor η . It follows by removing the time and space reflection elements from the underlying covariance group of special relativity, to yield the Poincaré group – the irreducible, continuous group of special relativity theory. The spinors η and χ are the reflections of each other, just as the quaternion solutions q^μ and $q^{*\mu}$ are the reflections of each other. Note again that the overline on the spinor fields above denotes the complex conjugate while the asterisk on the quaternions is the quaternion conjugate – denoting the time (or space) reflection directly.

Because of the extra degrees of freedom, the factorized spinor equations (3.28) give a more general solution and extra predictions (e.g., the anomalous Zeeman effect) than (3.27), just as the factorized quaternion equations (3.25a) and (3.25b) give a more general solution and extra predictions, compared with the Einstein equations (2.3).

3.5 A Symmetric Tensor–Antisymmetric Tensor Representation of General Relativity

The factorized form of the field equations (3.25a) [or (3.25b)] transform in the curved spacetime as the quaternion vector field $q^\mu(x)$. They therefore constitute 16 independent relations at each spacetime point, in accordance with the irreducible representations of the Einstein group.

The reason that there are more equations here – 16 relations – than in Einstein’s tensor formulation (2.3) – 10 relations – is that the tensor formulation is reducible, since its covariance includes reflections

in space and time, while the quaternion formulation – 16 relations – is based on an irreducible representation of the symmetry group of general relativity. The latter are covariant only under the continuous transformations in space and time that leave the form of the laws unchanged, which is all that is required by general relativity theory.

Let us now extract from the non-reflection-symmetric quaternion formalism the part that is even and the part that is odd under space and time reflections. We will see that the even part is in one-to-one correspondence with Einstein's symmetric tensor field equations (2.3). These correspond to 10 out of the 16 relations of the quaternion metric field equations. The remaining 6 equations, which are odd under spatial and temporal reflections, transform as an antisymmetric second-rank tensor formalism. Taking the covariant divergence of the latter then yields a formalism that is in one-to-one correspondence with the Maxwell field formalism for electromagnetism.

To carry out this program, we iterate (3.25a) and (3.25b). Multiplying (3.25a) on the right with the conjugated quaternion solution q_γ^* and (3.25b) on the left with the quaternion solution q_γ , and then adding and subtracting the two resulting equations (a procedure that is unique up to the constant k or k' on the right), the following relations are obtained:

$$\begin{aligned} \frac{1}{2} [K_{\rho\lambda} q^\lambda q_\gamma^* - (\pm) q_\gamma q^{*\lambda} K_{\rho\lambda} + q^\lambda K_{\rho\lambda}^\dagger q_\gamma^* - (\pm) q_\gamma K_{\rho\lambda}^\dagger q^{*\lambda}] \\ + \frac{1}{4} (q_\rho q_\gamma^* \pm q_\gamma q_\rho^*) R = 2(k \text{ or } k') (T_\rho q_\gamma^* \pm q_\gamma T_\rho^*) . \end{aligned} \quad (3.29\pm)$$

Since both equations (3.29 $\pm$) come from a single quaternion field equation (3.25a) [or (3.25b)], it follows that the constant k' does not entail any new universal constants, other than those associated with the matter field (h, c, m) and the gravitational field G . The latter is identified with the asymptotic limit of the theory where Newton's formalism for universal gravitation emerges as an approximation.

3.5.1 Einstein's Field Equation from the Symmetric Tensor Part

Comparing terms on the left hand side of (3.29+) and the correspondences (3.17) and (3.3) for the respective Ricci tensor, scalar curvature and metric tensor, we see that the left-hand side of (3.29+) corresponds precisely with the Einstein tensor, viz., $R_{\rho\gamma} - (1/2)g_{\rho\gamma}R$. Thus, the right-hand side of (3.29+) corresponds with the symmetric tensor of the source field $T_{\rho\lambda}$ of Einstein's equations (2.3).

To demonstrate that the left-hand side of (3.29+) transforms as a symmetric, second-rank tensor, let us take the trace of both sides of this equation, with respect to the spinor indices:

$$\begin{aligned} \frac{1}{4}\text{Tr}[K_{\rho\lambda}(q^\lambda q_\gamma^* - q_\gamma q^{*\lambda}) + \text{h.c.}] + \frac{1}{8}\text{Tr}(q_\rho q_\gamma^* + q_\gamma q_\rho^*)R \\ = k\text{Tr}(T_\rho q_\gamma^* + q_\gamma T_\rho^*) . \end{aligned} \quad (3.30)$$

Since, according to (3.13), $K_{\rho\lambda} = -K_{\lambda\rho}$, it follows from the definition of the Riemann curvature tensor (3.16) that

$$R_{\mu\rho\gamma\lambda} = -R_{\rho\mu\gamma\lambda} = -R_{\mu\rho\lambda\gamma} = R_{\gamma\lambda\mu\rho} .$$

The proof that the Ricci tensor is symmetric then follows:

$$\begin{aligned} R_{\rho\gamma} &\equiv R_{\rho\gamma\lambda}{}^\lambda = -R_{\gamma\rho\lambda}{}^\lambda = -g^{\lambda\alpha} R_{\gamma\rho\lambda\alpha} \\ &= g^{\alpha\lambda} R_{\gamma\rho\alpha\lambda} = R_{\gamma\rho\alpha}{}^\alpha = R_{\gamma\rho} . \end{aligned}$$

At this stage of the analysis, the right hand side of (3.29+) could, in itself, be either symmetric or antisymmetric in its tensor indices. However, a second-rank tensor may always be expressed as a sum of a symmetric part and an antisymmetric part. Since the left-hand side of this equation has now been shown to be symmetric, only the symmetric part of the source term on the right could play a role.

It is noted further that the matter source field on the right-hand side of (3.29+) is even with respect to reflections in space and time. This follows from the property of a product of quaternions when conjugated, that it is equal to the conjugated quaternions separately, multiplied in the reverse order. Thus, the right-hand side of (3.29+) is the sum of a quaternion and its time reversal (or space reflection) quaternion – that is, it is invariant (even with respect to reflections) in space and time.

Einstein's field equations (2.3) are then in one-to-one correspondence with the quaternion field equation (3.29+). These are 10 out of the 16 equations of the quaternion factorization of the Einstein formalism.

3.5.2 The Maxwell Field Equations from the Antisymmetric Part

The remaining 6 equations (3.29–) will now be seen to correspond to an antisymmetric tensor formalism of the second rank. Let us proceed by taking the trace (in spinor indices) of both sides of (3.29–), as follows:

$$R_{\rho\gamma}^\dagger + \frac{1}{8}\text{Tr}(q_\rho q_\gamma^* - q_\gamma q_\rho^*)R = k'\text{Tr}(T_\rho q_\gamma^* - q_\gamma T_\rho^*) , \quad (3.31)$$

where, by definition, the anti-Ricci tensor is

$$R_{\rho\gamma}^\dagger \equiv R_{\rho\gamma\lambda}^\dagger{}^\lambda = g^{\lambda\alpha} R_{\rho\gamma\lambda\alpha}^\dagger = \frac{1}{4}g^{\lambda\alpha}\text{Tr}[K_{\rho\lambda}(q_\alpha q_\gamma^* + q_\gamma q_\alpha^*) + \text{h.c.}] .$$

Since the spin curvature tensor $K_{\rho\lambda}$ is an antisymmetric tensor, it follows that the Riemann curvature tensor in (3.16) is now altered to the anti-Riemann curvature tensor by replacing the minus signs with plus signs, giving

$$R_{\gamma\lambda\mu\rho}^\dagger = -R_{\lambda\gamma\mu\rho}^\dagger = R_{\gamma\lambda\rho\mu}^\dagger .$$

Thus, the corresponding anti-Ricci tensor is antisymmetric, as follows:

$$\begin{aligned} R_{\rho\gamma}^\dagger &\equiv R_{\rho\gamma\lambda}^\dagger{}^\lambda = -R_{\gamma\rho\lambda}^\dagger{}^\lambda = -g^{\lambda\alpha} R_{\gamma\rho\lambda\alpha}^\dagger \\ &= -g^{\lambda\alpha} R_{\gamma\rho\alpha\lambda}^\dagger = -R_{\gamma\rho\alpha}^\dagger{}^\alpha = -R_{\gamma\rho}^\dagger . \end{aligned}$$

Since the trace of (3.29–) transforms as an antisymmetric tensor in spacetime, so must the full equation (3.29–) transform in this way.

Because the left-hand side of (3.31) transforms as a second-rank antisymmetric tensor in spacetime, it follows that the only part of the right-hand side (the source term) that contributes to the field equation (3.29–) is $2k'(T_\rho q_\gamma^* - q_\gamma T_\rho^*)$, its antisymmetric tensor part.

The foregoing results of the analysis lead to the exact form of the Maxwell field equations for electromagnetism. This is accomplished by converting (3.29–) into a vector field equation, consistent with Maxwell's equations. To do this, we take the covariant divergence of both sides of (3.29–) to obtain the formal expression of the four of Maxwell's equations with sources:

$$F_{\rho\gamma}{}^{;\rho} = \frac{4\pi}{c}j_\gamma , \quad (3.32)$$

where the electromagnetic field intensity solutions correspond with the geometrical fields:

$$\begin{aligned} F_{\rho\lambda} \Longleftrightarrow Q \left[\frac{1}{4}(K_{\rho\lambda}q^\lambda q_\gamma^* + q_\gamma q^{*\lambda}K_{\rho\lambda} + q^\lambda K_{\rho\lambda}^\dagger q_\gamma^* + q_\gamma K_{\rho\lambda}^\dagger q^{*\lambda}) \right. \\ \left. + \frac{1}{8}(q_\rho q_\gamma^* - q_\gamma q_\rho^*)R \right] , \end{aligned} \quad (3.33)$$

where Q is a constant of proportionality with the dimension of electric charge, applied to both sides of the equation. The current density source term is

$$j_\gamma \Longleftrightarrow \frac{cQk'}{4\pi} (T_\rho{}^{;\rho} q_\gamma^* - q_\gamma T_\rho{}^{*;\rho}) . \quad (3.34)$$

The four of Maxwell's equations without sources,

$$F_{[\rho\gamma}{}^{;\nu]} = 0 , \quad (3.35)$$

where the brackets denote a cyclic sum, then follow from the definition (3.33) of the electromagnetic field tensor in terms of the antisymmetric spin curvature tensor $K_{\mu\nu}$ in both parts of the right-hand side of (3.33). That is, $F_{[\mu\nu;\lambda]} \propto K_{[\mu\nu;\lambda]}$. But since $K_{\mu\nu} = \Omega_{\mu;\nu} - \Omega_{\nu;\mu}$ according to (3.13), and the spin affine connection Ω_μ is a four-vector in configuration space, it follows that the cyclic sum $K_{[\mu\nu;\lambda]} = 0$, automatically, and therefore that $F_{[\mu\nu;\lambda]} = 0$. This shows that magnetic monopoles are not predicted by this generalized theory of electromagnetism, in the quaternion formalism, in agreement with the empirical facts. This is similar to the conclusion that in the standard formalism the four of Maxwell's equations without sources, $F_{[\mu\nu;\lambda]} = 0$, because of the fact that the antisymmetric tensor $F_{\mu\nu}$ is equivalent to a four dimensional curl of a four-vector, $\partial_\mu A_\nu - \partial_\nu A_\mu$.

In deriving the source term j_γ above, use was made of the fact that the covariant derivatives of the quaternion fields q_γ vanish identically.

Since $F_{\rho\gamma}$ is an antisymmetric tensor in spacetime, it follows that the four-divergence of the current density four-vector vanishes:

$$j_\gamma{}^{;\gamma} = \frac{cQ}{4\pi} F_{\rho\gamma}{}^{;\rho;\gamma} = 0 . \quad (3.36)$$

With the assumption of a flat spacetime limit, this global conservation equation reduces to the law of conservation of electric charge, from the equation of continuity:

$$\partial^\gamma j_\gamma = 0 .$$

Using Gauss' theorem, this leads to the conservation of electric charge $Q = \int j^0 d^3x = \text{constant}$, coming from the integral expression of the continuity equation:

$$\partial^0 \int j_0 d^3x = \int \partial^k j_k d^3x = \int j_k n^k dS = 0 ,$$

where n^k is the normal vector to the surface S bounding the charge Q , contained in the volume V . The last integral vanishes because of the boundary condition that the current density in space $j_k(S)$ vanishes on

the spatial boundary S . Thus, the total charge $Q = \int j_0 d^3x$ contained in an arbitrary volume is conserved in time.

Summing up, we have seen that the antisymmetric, second-rank tensor equation (3.29–), derived from the quaternion formulation of general relativity theory (3.25a) and (3.25b), leads to the precise structure of Maxwell's field equations for electromagnetism. The solutions $F_{\rho\gamma}$ of Maxwell's equations are given in the correspondence (3.33), whilst the current density source terms are given in the correspondence (3.34). The conclusion is that, by going to the irreducible representations of the Einstein group of general relativity theory, we arrive at a totally unified field theory, in terms of the 16-component metrical field equation (3.25a) or its conjugate (3.25b), expressing the laws of gravitation and electromagnetism in terms of a single covariant field.

This means of reaching a unified field theory of gravity and electromagnetism was originally suggested by Einstein, in general terms. He advised that the most general symmetry group of general relativity – the Einstein group – must be a set of continuous (and I have argued analytic) transformations – a Lie group – at each spacetime point, that leave the laws of nature covariant. In his article [5], Einstein said: “Every attempt to establish a unified field theory must start, in my opinion, from the group of transformations which is no less general than that of the continuous transformations of the four coordinates. [...] It is further reasonable to attempt the establishment of a unified theory by the generalization of the relativistic theory of gravitation.”

3.6 The Geodesic Equation in Quaternion Form

The 16 equations (3.29) that are the symmetric tensor–antisymmetric tensor expression of the basic quaternion field equations (3.25a) and (3.25b) are another way of expressing the latter. The question then arises: does the geodesic equation that follows from a minimization of the line integral over the quaternion differential metric $\delta \int ds = \delta \int q^\mu dx_\mu = 0$ differ from the geodesic equation (2.22) of Einstein's tensor formulation? We will see in this section that the geodesic equation in the quaternion formulation of general relativity has the same functional form as it does in the tensor formulation. But indeed there is a difference when they are expressed in terms of rates of change of the global coordinates, with respect to the proper time of a moving body.

We will see that indeed, in this natural generalization the quaternion-valued time parameter entails a geodesic equation that is a set of four

simultaneous differential equations while the standard formulation entails a real-number-valued time parameter, determined by a single differential equation. Thus, the ‘time measure’ in the quaternion formulation is more general than it is in the tensor formulation of general relativity.

It is interesting to note what William Hamilton said in the 19th century, after his discovery of quaternion algebra [12]: “It appeared to me to regard algebra as being no mere art, nor language, nor primarily as Science of Quality; but rather as the Science of Order in Progression. It was, however, a part of this conception, that the progression here spoken of was understood to be continuous and unidimensional: extending indefinitely forward and backward, but not in any lateral direction. And although the successive states of such progression might (no doubt) be represented by a point upon a line, yet I thought that their simple successiveness was better conceived by comparing them with moments of time, divested, however, of all reference to cause and effect; so that ‘time’ here considered might be said to be abstract, ideal or pure, like the ‘space’ which is the object of geometry. In this manner I would regard algebra as the Science of Pure Time. [...] And with respect to anything unusual in the interpretation thus proposed, it is my wish to be as preparatory to the study of quaternions.”

In order to predict the motion of a test body in the field of influence of the other mass of a closed system, it is assumed, as in Einstein and Galileo, that the test body moves along a geodesic – the extremum path in space and time. This follows from the correspondence of this path with the minimal energy of the body. That is to say, it would take external energy to do the work to move the body off of this geodesic path.

In the quaternion calculus, the extremum trajectory is determined from the vanishing of the variation of the line integral over the invariant differential ds , leading to the geodesic equation (2.19) in the standard formulation. However, the differential ‘proper time’ ds is in this case quaternion-valued. Thus it entails four real-number-valued parameters rather than one. That is, the geodesic equation may be expressed in the form:

$$\left[\frac{d^2 x^\mu}{ds^2} + \Gamma_{\nu\lambda}^\mu \frac{dx^\nu}{ds} \frac{dx^\lambda}{ds} = 0 \right]_{\alpha\beta}, \quad (3.37)$$

where $\alpha, \beta = (1,1), (1,2), (2,1), (2,2)$ refer to the four equations of motion that must be solved simultaneously here in order to predict the trajectory of the test body.

It is interesting then that the theory of general relativity, in its irreducible form, entails a ‘proper time’ of a moving body that is more general than it is in the mechanics of the classical Newtonian theory, or in Einstein’s tensor formalism. For, here, one needs four numbers, rather than one, to signify the time parameter at each spatial coordinate on the trajectory, before proceeding to the other (continuously connected) coordinates of the trajectory, to describe the motion of any massive body – an electron, a planet or a galaxy.

Let us now parameterize the variation with respect to the quaternionic line element. In the conventional real number theory, one has the differential operator:

$$\frac{d}{ds} = \lim_{\Delta x^\alpha, \Delta x^\beta \rightarrow 0} \frac{\Delta}{\Delta(g_{\alpha\beta}x^\alpha x^\beta)^{1/2}} = \frac{d}{(g_{\alpha\beta}dx^\alpha dx^\beta)^{1/2}},$$

where, by definition of the derivatives, $g_{\alpha\beta}$ is evaluated where the derivatives are taken.

In the present quaternion formulation, we must instead utilize the differential operator

$$\frac{d}{ds} = \lim_{\Delta x^\alpha \rightarrow 0} \frac{\Delta}{\Delta(q_\alpha x^\alpha)} = q_\alpha^{-1} \frac{d}{dx^\alpha}, \quad (3.38)$$

where q_α^{-1} is the inverse of q_α , evaluated at the spacetime point where the derivative is taken. According to the algebra of quaternions, the inverse of a quaternion is the ratio of the quaternion conjugate to its norm:

$$q_\alpha^{-1} = \frac{q_\alpha^*}{q_\alpha q_\alpha^*}. \quad (3.39)$$

The second order derivative of the quaternion variable is then

$$\frac{d^2}{ds^2} = q_\alpha^{-1} \frac{d}{dx^\alpha} \left[q_\alpha^{-1} \frac{d}{dx^\alpha} \right]. \quad (3.40)$$

The vanishing variation $\delta \int ds = \delta \int q_\alpha dx^\alpha = 0$ then gives rise, in the same manner as in the derivation of the geodesic equation (2.19), to the same fundamental form of the geodesic equation (3.37).

The particular problem of determining the path of a test body need not entail the noncommutative feature of the quaternions, e.g., when polarization does not play a role in the application at hand. Since the determinant of the product of matrices is the product of the determinants of the factor matrices, the determinant with respect to the quaternion differential is

$$\int \left[\frac{1}{2} \partial_\lambda g_{\alpha\beta} \frac{dx^\alpha}{dS} \frac{dx^\beta}{dS} \delta x^\lambda + g_{\alpha\beta} \frac{dx^\alpha}{dS} \frac{d(\delta x^\beta)}{dS} \right] dS = 0, \quad (3.41)$$

where

$$dS = |q_\alpha| dx^\alpha, \quad \frac{d}{dS} = \left| q_\alpha^{-1} \frac{d}{dx^\alpha} \right| = |q_\alpha|^{-1} \frac{d}{dx^\alpha},$$

and the vertical bars denote the determinant with respect to the spinor indices.

Integrating the second term on the left-hand side of (3.41) by parts, using the fact that the variation $\delta(dx^\beta)$ vanishes at the endpoints of the line of integration, and the relation between the components of the affine connection and the metric tensor, it follows that

$$\frac{d^2 x^\rho}{dS^2} + \Gamma_{\alpha\beta}^\rho \frac{dx^\alpha}{dS} \frac{dx^\beta}{dS} = 0. \quad (3.42)$$

This equation is the determinant of the geodesic equation (3.37). This real number form of the geodesic equation then has the conventional form as in Einstein's tensor formalism. Nevertheless, they are different because the latter is in terms of derivatives with respect to the differential $ds = (g_{\alpha\beta} dx^\alpha dx^\beta)^{1/2}$, while the former is in terms of derivatives with respect to $dS = |q_\alpha| dx^\alpha$. The geodesic equation (3.42), applied to stationary state problems (such as orbital motion), is not the same as the conventional one because of the time-dependence that is implicit in q_α as compared with the time-independence of $g_{\alpha\beta}$ in the planetary motion problem, as we will discuss in Chap. 8.

3.7 Summary

It has been shown in this chapter that as soon as one removes the space and time reflection elements from the underlying covariance group of Einstein's tensor field equations (2.3), one takes the reducible representations that describe the symmetry of that formalism in general relativity to the irreducible representations of the theory of general relativity. This leads to a factorization of the 10-component tensor formalism to a 16-component quaternion formalism. This is consistent with the fact that the underlying symmetry group of general relativity – the Einstein group – is a 16-parameter Lie group, that is, it entails 16 essential parameters.

These 16 equations of the quaternion formalism, that is a vector formalism in which each of the vector components is quaternion-valued,

were then shown, by iteration to a second-rank tensor form, to break up into 10 symmetric tensor equations and 6 antisymmetric tensor equations. The 10 equations are in one-to-one correspondence with the original form of Einstein's symmetric tensor field equations, thus explaining gravity. The remaining 6 equations are re-expressed in a form that is in one-to-one correspondence with the Maxwell field equations in terms of its antisymmetric tensor solutions, $F_{\rho\gamma}$ – the components of the electric and the magnetic fields of charged matter – thus explaining electromagnetism.

We have seen, then, that the use of the irreducible Einstein group to underlie the symmetry of the laws of nature leads in a natural way to the unification of gravitational and electromagnetic phenomena, in terms of a single, covariant (16 component) quaternion metrical field $q^\mu(x)$ – a four-vector wherein each of the four components is quaternion-valued.

A second new feature of this analysis is that the geodesic equation, which serves as the equation of motion of a test body, entails a generalization of the time parameter for its trajectory. The proper time of a 'freely moving' body in a curved spacetime is now a generalization whereby the time measure entails four parameters at each point along the body's trajectory – its geodesic – rather than one, as in the real-number-valued theories in mechanics. This conclusion corroborates the intuition of the discoverer of quaternion algebra, William Hamilton, who, in the 19th century, believed that the quaternion variables would be essential to our understanding of the time measure in physics. These features of the quaternion geodesic will be elaborated and applied to physical problems in astrophysics in Chap. 8. It is also interesting to note, at this point of the discussion, that in the 19th century, Maxwell had the following to say about the use of the quaternions in physics: "... the virtue of the quaternions lies not so much as yet in solving hard questions as in enabling us to see the meaning of the question and its solution." (Extract of a letter from Maxwell to P.G. Tait, Archives, Cavendish Laboratory, University of Cambridge.)

4. Quantum Mechanics from a Theory of Inertial Mass in Relativity

Quantum mechanics has been eminently successful in twentieth century physics in describing the properties of matter in the atomic domain. It demonstrates that, under some physical circumstances of experimentation, particles of matter behave like continuous waves, while under other conditions of experimentation they behave like discrete particles. The doctrine of this theory is then that of ‘wave–particle dualism’, specifying that the nature of the particle of matter is not ‘predetermined’ – it depends on the way in which it is measured.

Quantum mechanics is conventionally interpreted as a theory of matter that is rooted, foundationally, in probability, thus indicating that the basic variables of matter are the probabilities of locating the particle at one place or another. But the question arises: where does this probabilistic theory of matter originate? It was Einstein’s idea that the formal expression of the probability calculus that is called quantum mechanics is an incomplete theory of matter that originates in a complete continuous field theory in general relativity. The latter would be a theory of matter wherein all of the variables of matter are ‘predetermined’, without the possibility of any probability interpretation.

In this chapter, it is shown that a field theory of the inertia of matter – not derived before in physics – based on general relativity, yields as an approximation the entire formal expression of quantum mechanics. Thus the incomplete probability theory that correctly describes atomic matter at low energies is only a mathematical approximation for a complete field theory of the inertia of matter. In the latter, a local quantity of matter is a mode of a continuous field, rather than the discrete singularity of the particle view.

The concept of the elementarity of the interaction replaces the concept of the elementarity of the particle of matter. This implies that the fundamental theory of matter in general relativity is holistic, wherein there are no separable, singular things of matter. Rather, we have a continuum with an infinite distribution of correlated modes, analogous to the multiple ripples of a continuous, though disturbed pond.

The theory proves that in the Newtonian limit of general relativity, the force of gravity can only be attractive (an empirical fact). But when we are not in this approximation, the force of gravity can be either attractive or repulsive, thereby leading to the dynamics of the oscillating universe cosmology.

4.1 Introduction

The aim of this treatise is to formulate a unified field theory wherein all fields of force originate in a single covering field of force. The term ‘force’ per se then represents the effect of matter on ‘other matter’, in terms of a single, continuous, nonsingular field theory in general relativity. But one must ‘close the circle’ for a closed system, by including the reaction of the ‘other matter’ that is acted upon by the unified force field. This is in accordance with the spirit of Newton’s third law of motion – for every action (force exerted) by one body A on another body B there must be an equal and oppositely directed reaction (force exerted) by B on A.

To determine the reaction of matter that is acted upon we must incorporate in the total closed system the inertial mass of this matter. The inertial mass is, by definition, the resistance of matter to a change of its inertial state of motion of rest or constant speed in a straight line, relative to the force-exerting matter. Newton introduced this concept and quantified the inertial mass with the parameter m , as it appears in his second law of motion: $\mathbf{F} = m\mathbf{a}$. We must, however, continually keep in mind that in general relativity, there are no discrete particles of matter – there are only correlated matter fields. In limits, however, these appear as discrete particles with their own intrinsic masses. But these, according to the field theory of general relativity are, in reality, distributed peaks of a continuous mass distribution.

The question then arises: with the continuous field model posed in general relativity, what is the origin of the manifestations of matter that is its inertial mass? In the 17th century, Newton assumed the atomistic model wherein the inertial mass is one of the intrinsic qualities of a bit of matter, quantified by the parameter m . In the 19th century, Mach took the view that a body’s mass is a measure of its coupling to all of the other matter of a closed system – in principle the universe. Einstein called this view the Mach principle. It is a view that, like Newton’s third law of motion, implies a closed system at the outset. This is in contrast with the open system of interacting ‘things’,

as assumed in Newton's classical view and in the quantum mechanical 'atomistic' model of matter.

I have argued that in the continuum field theory of matter according to the theory of general relativity, the Mach principle necessarily follows in a generalized form. Here, not only the inertial property of matter, but all of its other (earlier claimed) intrinsic qualities, such as electric charge and magnetic moment, are also measures of coupling between observed matter and the rest of the closed system of which it is a constituent. I have called this idea the generalized Mach principle. It is a view that exorcises all remnants of the atomistic model of matter.

According to Newton's theory, for a constant mass m , the effect of this body, that is the acceleration $\mathbf{a}$, is caused by an external force $\mathbf{F}$. This follows from the empirical fact that, at low energy, the ratio of the magnitudes of two external forces exerted on a single body with mass m is the corresponding ratio of the magnitudes of the caused accelerations: $F_1/F_2 = a_1/a_2$. Thus, the parameter m is the constant of proportionality between the external (vector) force exerted on a body and its caused (vector) acceleration, $\mathbf{F} = m\mathbf{a}$. In the ratio F_1/F_2 for a single body, the mass cancels. This linear 'law' of motion implies that an 'atom of matter', with intrinsic mass m , is an individual, separable entity.

In Mach's approach, on the other hand, one may say that (still consistent with Newton's second law of motion), for constant acceleration of two different bodies, the ratio of their masses is equal to the ratio of the forces exerted to produce their constant acceleration.

An example is the free fall of two different massive bodies to earth, at the same constant acceleration of gravity g . One may express this empirical fact, discovered by Galileo, with the law $F_1/F_2 = m_1/m_2$. This is equivalent to the linear law $F = km$, where $k = F_0/m_0$ is the constant of proportionality, and (F_0 and m_0) are the external force and mass of a test body, chosen as a standard. This formula implies that the entire external force that acts on this body is the cause of the inertial mass of the given body. That is to say, it assumes a closed system at the outset, rather than individual, separable bodies. This closed system is, in principle, the entire universe. Nevertheless, the far-away contribution of, say, the distant stars, to the mass of a local electron is likely much less significant than its coupling to other nearby massive bodies, such as the particle-antiparticle pairs of the 'physical vacuum' in its vicinity.

The theory of general relativity, which implies a holistic, nonsingular, continuum field theory, is compatible with the generalized Mach principle. In this chapter, we will see how the formal expression of general relativity theory, in its irreducible form, gives rise to a derivation of a field theory of the inertial mass of matter, including empirical features not derived previously. Examples of the latter, to be discussed in Chap. 7, are

- mass doublets,
- a mass spectrum in the linear limit of the theory,

and in this chapter we will discuss

- the predicted attractive feature of the gravitational force in the Newtonian limit of the theory.

A generally covariant field theory of the inertia of matter so-derived will be shown in a linear limit, i.e., in the asymptotically flat spacetime to have the formal (Hilbert space) structure of quantum mechanics. Thus, in accordance with Einstein's expectations, we will find that quantum mechanics is formalizable in terms of a probability calculus because it is an incomplete representation for the matter of the microdomain. It is incomplete because it is a linear approximation, in special relativity, and in nonrelativistic quantum mechanics, for a complete, generally covariant, nonlinear field theory of the inertia of matter. The nonlinearity of the formalism in general relativity in itself prohibits the interpretation in terms of a probability calculus as foundational, because the latter is necessarily a linear theory. The probability interpretation is a cornerstone of quantum mechanics – it is the reason for the claim that a fundamental law of nature is the principle of linear superposition.

The quantum theory and the theory of general relativity are fundamentally incompatible, in terms of their respective foundations. Some of the basic conflicts are:

- the particle model of matter in quantum mechanics, versus the holistic, continuum view in general relativity,
- the linear nature of the quantum mechanical formalism versus the nonlinearity of the theory in general relativity,

even though one of these formalisms may be an accurate approximation for the other, under the appropriate circumstances. But if, in principle, one of these theories of matter is scientifically true, the other must be scientifically false!

Thus we see that there can be no ‘quantization’ of Einstein’s field theory of gravitation. The particles called gravitons – the quanta of the Einstein field – cannot then exist in reality. Nevertheless, in accordance with the field theory in general relativity, forces between matter do propagate at a finite speed, though they cannot be identified with gravitons. The latter view is a topic of contemporary research in quantum field theory, to see if one can incorporate gravitation with the other forces in nature in the context of a quantum field theory. What I have argued here is that this goal is unachievable because of the natural (mathematical and conceptual) incompatibility of the quantum and relativity theories. The opposite view, pursued in this book, is that it is the context of general relativity that leads to the unified field theory that is sought.

4.2 Discovery of Quantum Mechanics

After the experimental discovery of the wave nature of matter, in the electron diffraction studies in the 1920s (Davisson and Germer in the US and G.P. Thomson in the UK), Erwin Schrödinger set about to find the wave equation whose solutions are the matter waves.

Schrödinger’s original intention was to demonstrate a relativistic wave equation, whose solutions are the (scalar) complex functions that would successfully describe the matter waves in a relativistically covariant fashion. He was not successful because his relativistic wave equation did not incorporate an equation of continuity that, in turn, would yield a law of conservation of electric charge. It was Schrödinger’s idea that the discovered nature of electrons, described as (complex function) matter waves, must be ‘buried in’ the real number expression of charge density and current density sources of Maxwell’s equations for electromagnetism. The natural relativistic covariance of Maxwell’s equations very likely led Schrödinger to require a relativistically covariant expression of his wave equation, leading to the source terms of the Maxwell equations.

Although he intended eventually to find a relativistic expression for his wave equation, for these reasons, he did initially find a non-relativistic form of the wave equation that incorporated a continuity equation, thus leading (in the nonrelativistic limit) to an equation of continuity and the required conservation of electric charge.

Nevertheless, the rest of the physics community, led by Niels Bohr, Werner Heisenberg and Max Born, interpreted Schrödinger’s nonrelativistic wave equation of continuity that accompanied the Schrödinger

wave equation as a law of conservation of probability. The interpretation of quantum mechanics then came to the positivistic view of the Copenhagen school, wherein the matter wave – the complex function probability amplitude (the solution of the Schrödinger wave equation) became central. It was to be tied to the macro-observer's measurement of the properties of matter, in the context of a probability calculus. Thus, contrary to Schrödinger's own interpretation, quantum mechanics became a theory of measurement – asserting it to be the foundational feature of the laws of elementary matter.

Schrödinger and Einstein, as well as Planck and de Broglie, were unwilling to accept the 'new physics' – a positivistic view wherein the laws of nature are laws of chance, and where the only reality about the electrical properties of elementary matter is the subjective measurement of its properties by macro-observers, expressed in the context of a probability calculus.

Not long after the experimental and theoretical discoveries of the wave nature of matter, Paul Dirac discovered a way to express the Schrödinger wave equation that was compatible with the symmetry requirements of special relativity theory. This required going from the (single component, complex) scalar form of the wave equation to the (two or four component complex) spinor form of the wave equation. The Dirac spinor wave equation was then covariant under the Poincaré group of special relativity, in addition to incorporating an equation of continuity to accompany his wave equation.

As we discussed in the preceding chapter, when the reflection symmetry transformations are removed from the covariance group of the Klein–Gordon equation (3.27) (which is in turn constructed by using Schrödinger's operator equivalents of energy and momentum in the 'classical' relativistic equation $E^2 = p^2c^2 + m^2c^4$), it factorizes into the following pair of two-component spinor equations:

$$(\sigma^\mu \partial_\mu + I)\eta = -m\chi, \quad (4.1a)$$

$$(\sigma^{*\mu} \partial_\mu + I^*)\chi = -m\eta, \quad (4.1b)$$

where $\chi = \varepsilon \bar{\eta}$ (henceforth the overline stands for the complex conjugate of the function) is the time-reversed spinor variable, or in a different convention, the space reflection, where one multiplies the right-hand sides of (4.1a) and (4.1b) by i . Units are chosen so that $\hbar/2\pi = 1$ and $c = 1$, and m is the inertial mass of the particle described by the spinor field (η, χ) . Conventionally, the value of the mass term is determined by solving the equations (4.1a) and (4.1b) and then fitting a term dependent on mass, such as mc^2 , with the data. One then inserts this

mass value in the spinor wave equations (4.1a) and (4.1b). This is in contrast with deriving the value of m from first principles. The function I in the wave equations (4.1a) and (4.1b) is the ‘interaction’ term that is due to all of the ‘other matter’ of the closed system that couples to the matter represented by the solution (η, χ) . If the latter spinor field is labeled by the superscript (i) , as in $(\eta^{(i)}, \chi^{(i)})$, then

$$I_i = I(\eta^{(1)}, \chi^{(1)}, \eta^{(2)}, \chi^{(2)}, \eta^{(3)}, \chi^{(3)}, \dots, \eta^{(i-1)}, \chi^{(i-1)}, \eta^{(i+1)}, \chi^{(i+1)}, \dots).$$

This symbolizes the dependence of the given matter field on all of the other matter fields of the closed system, in accordance with the Mach principle.

What Dirac did next was to bring back reflection symmetry by combining the two-component spinor equations (4.1a) and (4.1b) into a single four-component (bispinor) wave equation

$$(\gamma^\mu \partial_\mu + I)\psi = -m\psi, \quad (4.2)$$

where γ^μ (the Dirac matrices) are 4×4 matrices that are particular combinations of the Pauli matrices σ^0, σ^k ($k = 1, 2, 3$). But the bispinor Dirac equation (4.2) is again based on the reducible representation of the Poincaré group, since it entails reflection symmetry, which is not required of the symmetry of special relativistic covariance, as exhibited in the Lie group that is the Poincaré group in terms of its irreducible representations. Indeed, the most general irreducible expression of relativistic quantum mechanics, in special relativity, is obtained by removing the reflection elements once again, thus returning to the irreducible form (4.1a) and (4.1b) of wave mechanics, which is called the Majorana equation. More of the details of this discussion can be found in [117].

Thus, (4.2) – the Dirac equation – is a reducible form of relativistic quantum mechanics. The two-component spinor field equations (4.1a) and (4.1b) are the irreducible form of relativistic wave mechanics. As such, it is the form (4.1a) and (4.1b) of the wave equation that contains the maximum predictions of this field theory. In reducible forms such as (4.2), some of the predictions of the theory are masked, such as parity violating interactions in the microscopic domain.

Since this is a linear formalism, the principle of linear superposition applies, implying that any sum of solutions of the equations is another possible solution. This feature is conducive to the formulation of a probability calculus, since this is a required feature of probabilities. It was then Max Born’s prime motivation to interpret

Schrödinger's nonrelativistic wave mechanics or relativistic wave mechanics of Dirac/Majorana as a probability calculus. He found that as such, quantum mechanics could be couched in the formal expression of a linear Hilbert function space as soon as one assumed that the solutions are square integrable, with the square being the product of a probability amplitude (a complex function) and its complex conjugate, thus having only positive values. Then, with normalization of these squares for each state of the microsystem, they could be represented in terms of combinations of probabilities, whose sum must be equal to unity.

It was then Bohr and Heisenberg who interpreted quantum mechanics as a subjective theory of measurement, in the context of probabilities, taken as the final form of a fundamental explanation of matter in the microscopic domain.

From a topological point of view of the most primitive (irreducible) formulation of quantum mechanics (4.1a) and (4.1b) in special relativity, one may interpret these two-component spinor equations as a mapping of one type of spinor η onto its time-reversed spinor χ . The mass parameter m is then taken to be a measure of this mapping. The aim now is to derive (rather than insert) the mass associated with the matter wave by re-expressing the relativistic wave equations (4.1a) and (4.1b) in general relativity, without the insertion of the parameter m . That is, the plan is to identify the most primitive expression of the inertial mass of matter with a mapping between the time-reflected spinor fields in a curved spacetime and thereby to recover a field that plays the role of the inertial mass m .

When this is done, it will be seen that the mass field, so derived, is a positive definite function of the space and time coordinates, since it is the modulus of a complex variable. This result implies that in the Newtonian limit of the theory, in the flat spacetime, the gravitational force can only be attractive. This is an empirically correct result that has never been predicted from first principles, either in Newton's theory of universal gravitation or in Einstein's tensor formulation of gravitation in the reducible form of general relativity. It will also be seen that, in accordance with the Mach principle, the inertial mass of observed matter vanishes as the total environment of this matter correspondingly vanishes. This follows as a consequence of the approach of the curvature of spacetime (in representing the existence of matter) to a flat spacetime (representing, in principle, an empty spacetime – a perfect vacuum, everywhere).

4.3 Inertial Mass from General Relativity

To carry out the program of deriving the inertial mass of matter from general relativity, we set $m = 0$ on the right-hand sides of (4.1a) and (4.1b) and express the left-hand sides of these equations in terms of operators in a curved spacetime. Thus, the global extension is

$$\sigma^\mu \partial_\mu \eta \longrightarrow q^\mu \eta_{;\mu} \equiv q^\mu (\partial_\mu + \Omega_\mu) \eta .$$

This extension of the spinor equations (4.1a) and (4.1b) to the curved spacetime, with $m = 0$, then yields the spinor wave equation and its time-reversed equation

$$[q^\mu (\partial_\mu + \Omega_\mu) + I] \eta = 0 , \quad (4.3a)$$

$$[q^{*\mu} (\partial_\mu + \Omega_\mu^\dagger) + I^*] \chi = 0 , \quad (4.3b)$$

respectively. If we are to derive the mass field that plays the role of m in the special relativistic equations (4.1a) and (4.1b), it must then reside in the spin affine connection field Ω_μ .

Consider the following Hermitian and anti-Hermitian matrix fields, that are the contraction of the quaternion q^μ and the spin affine connection fields in a Riemannian spacetime:

$$\Lambda_\pm = q^\mu \Omega_\mu \pm \text{h.c.} . \quad (4.4)$$

It follows from the time-reversal τ of the quaternion fields

$$\tau q^\mu = q^{*\mu} = \varepsilon \bar{q}^\mu \varepsilon , \quad (4.5)$$

that the time reversal of the spin affine connection is

$$\tau \Omega_\mu = -\Omega_\mu^\dagger . \quad (4.6)$$

It then follows that

$$\tau \Lambda_\pm = \pm \varepsilon \bar{\Lambda}_\pm \varepsilon \quad (4.7)$$

and

$$(\tau \Lambda_\pm) \Lambda_\pm = \pm |\det \Lambda_\pm| \exp(i\delta) \sigma^0 , \quad (4.8)$$

where the vertical bars denote the absolute value and

$$\delta = 0 \quad \text{if} \quad \det \Lambda_\pm < 0 , \quad \delta = \pi \quad \text{if} \quad \det \Lambda_\pm > 0 .$$

Let us now set up (4.8) in terms of the matrix equation

$$[(\tau \Lambda_{\pm}) \Lambda_{\pm}] \eta = \pm (2\alpha_{\pm})^2 \eta \exp(i\delta) . \quad (4.9)$$

This matrix equation may now be factorized as follows:

$$\Lambda_+ [\eta \exp(-i\delta/2)] = 2\alpha_+ [\chi \exp(i\delta_+)] , \quad (4.10a)$$

$$\tau \Lambda_+ [\chi \exp(i\delta_+)] = 2\alpha_+ [\eta \exp(i\delta/2)] , \quad (4.10b)$$

and

$$\Lambda_- [\eta \exp(-i\delta/2)] = 2i\alpha_- [\chi \exp(i\delta_-)] , \quad (4.11a)$$

$$\tau \Lambda_- [\chi \exp(i\delta_-)] = 2i\alpha_- [\eta \exp(i\delta/2)] , \quad (4.11b)$$

where

$$(2\alpha_{\pm})^2 = |\det \Lambda_{\pm}| . \quad (4.12)$$

The factorization (4.10a), (4.10b), (4.11a) and (4.11b) is unique only up to the arbitrary relative phase between the spinors η and χ . Thus, this factorization is independent of the arbitrary phase factors above, $\exp(i\delta_+)$ and $\exp(i\delta_-)$ that are incorporated in χ . Since these phases may be adjusted continuously without altering the forms (4.10a), (4.10b), (4.11a) and (4.11b), the latter equations may be grouped so that the relative phase between the time-reversed spinors in (4.10a) and (4.10b) are the same as in (4.11a) and (4.11b). In this way, (4.10a) [or (4.10b)] may be added to (4.11a) [or (4.11b)] without the need to explicitly specify their relative phases.

Combining the definition of the matrix fields $\Lambda_{\pm}$ shown in (4.4) with the sum of (4.10a) and (4.11a), we obtain the following relation between the time-reversed spinor variables:

$$q^{\mu} \Omega_{\mu} \eta = (\alpha_+ + i\alpha_-) \chi \equiv \lambda \exp(i\gamma) \chi , \quad (4.13)$$

where [with (4.12)],

$$\lambda = \text{mod}(\alpha_+ + i\alpha_-) = \frac{1}{2} (|\det \Lambda_+| + |\det \Lambda_-|)^{1/2} \quad (4.14)$$

is the modulus of a complex variable. λ plays the role of the inertial mass in the quantum mechanical equations. The argument of the complex variable $(\alpha_+ + i\alpha_-)$ is the function

$$\gamma(x) = \tan^{-1} \frac{\alpha_-}{\alpha_+} = \tan^{-1} \left(\frac{|\det \Lambda_-|}{|\det \Lambda_+|} \right)^{1/2} .$$

The geometrical mapping between the time-reversed spinors η and χ that we sought is then given in (4.13).

Finally, the time-reversed equation that accompanies (4.13) is obtained by taking the complex conjugate of the latter equation and applying the Levi–Civita symbol from the left and the right, to give

$$\varepsilon \bar{q}^\mu \varepsilon \Omega_\mu^* \varepsilon \varepsilon \eta = \lambda \exp(-i\gamma) \varepsilon \bar{\chi}.$$

Combining this with (4.5), (3.6) and (4.6), and the time-reversed spinor $\chi = \varepsilon \bar{\eta}$, it follows that

$$-q^{\mu*} \Omega_\mu^\dagger \chi = \exp(-i\gamma) \eta. \quad (4.15)$$

4.4 The Matter Field Equations in General Relativity

The generally covariant extension for the matter field is given in (4.3a) and (4.3b). With (4.13) these equations take the form

$$q^\mu \eta_{;\mu} = q^\mu \partial_\mu \eta + \lambda \exp(i\gamma) \chi = -I\eta. \quad (4.16)$$

With (4.15), the time-reversal of the spinor equation (4.16) is

$$q^{\mu*} \chi_{;\mu} = q^{\mu*} \partial_\mu \chi + \lambda \exp(-i\gamma) \eta = -I^* \chi. \quad (4.17)$$

We see, then, that if the phase factors $\exp(\pm i\gamma)$ could be ‘transformed away’ from (4.16) and (4.17), they would become, precisely the generally covariant form of the two-component spinor Dirac equation (the Majorana form) for quantum mechanics in a curved spacetime. We will find in the next section that this does happen when we add gauge covariance to the underlying symmetry of the theory.

4.5 Gauge Covariance

It was discovered at an early stage of field theories that the incorporation of a continuity equation with the rest of the laws of nature is a necessary and sufficient consequence of the addition of gauge covariance to the relativistic spacetime covariance. The gauge symmetry, to be defined below, is then equivalent to the incorporation with the laws of nature of the continuity equation in special relativity, and its global extension in general relativity:

$$\partial_\mu (\eta^\dagger \sigma^\mu \eta) = 0 \longrightarrow (\eta^\dagger q^\mu \eta)_{;\mu} = 0. \quad (4.18)$$

The explicit gauge transformations follow from this analysis: using the product rule for covariant differentiation as well as ordinary differentiation, we have

$$[\eta \exp(-i\gamma/2)]_{;\mu} = \eta_{;\mu} \exp(-i\gamma/2) - \frac{i}{2} \eta \gamma_{;\mu} \exp(-i\gamma/2) , \quad (4.19a)$$

$$\partial_\mu [\eta \exp(-i\gamma/2)] = (\partial_\mu \eta) \exp(-i\gamma/2) - \frac{i}{2} \eta (\partial_\mu \gamma) \exp(-i\gamma/2). \quad (4.19b)$$

Multiplying (4.16) on the left with $\exp(-i\gamma/2)$, using (4.19a) and (4.19b) and subtracting the term

$$\frac{1}{2} i q^\mu (\partial_\mu \gamma) \eta \exp(-i\gamma/2)$$

from both sides of the equation, it follows that

$$\begin{aligned} q^\mu [\exp(-i\gamma/2)]_{;\mu} + \frac{1}{2} i q^\mu \eta \exp(-i\gamma/2) (\gamma_{;\mu} - \partial_\mu \gamma) \\ = q^\mu \partial_\mu [\eta \exp(-i\gamma/2)] + \lambda \exp(-i\gamma/2) \chi \\ = - \left[I \eta \exp(-i\gamma/2) + \frac{1}{2} i q^\mu (\partial_\mu \gamma) \eta \exp(-i\gamma/2) \right] . \end{aligned}$$

Since the phase $\gamma(x)$ is a scalar field, $\gamma_{;\mu} = \partial_\mu \gamma$. Thus the preceding equation takes a form without the explicit appearance of the phase factor as follows:

$$q^\mu \eta'_{;\mu} = q^\mu \partial_\mu \eta' + \lambda \chi' = -I' \eta' , \quad (4.20)$$

where

$$\eta' = \eta \exp(-i\gamma/2) , \quad \chi' = \chi \exp(i\gamma/2) , \quad (4.21a)$$

$$I' = I + \frac{1}{2} i q^\mu \partial_\mu \gamma . \quad (4.21b)$$

The invariance of the formalism with respect to the transformations (4.21a) and (4.21b) is called gauge covariance. Equation (4.21a) is called a gauge transformation of the first kind, and (4.21b) is called a gauge transformation of the second kind. The latter corresponds to the gauge invariance of classical electromagnetic theory. The former transformation under phase change (4.21a) corresponds to the necessary incorporation of the equation of continuity (4.18) in the field theory in accordance with Noether's theorem.

A similar result to (4.20) is obtained for the time-reversed equation. Thus we have the following set of two coupled two-component spinor equations in the curved spacetime:

$$(q^\mu \partial_\mu + I)\eta = -\lambda\chi, \quad (4.22a)$$

$$(q^{*\mu} \partial_\mu + I^*)\chi = -\lambda\eta. \quad (4.22b)$$

Thus we have derived from general relativity the fundamental structure of the equations of quantum mechanics (4.22a) and (4.22b). In the special relativistic limit of these equations we arrive at equations (4.1a) and (4.1b) in the flat spacetime – corresponding to the Majorana two-component spinor equations or Dirac's four-component bispinor equation (4.2). As $v/c \rightarrow 0$, these equations, in turn, take the Schrödinger form of nonrelativistic quantum mechanics, as we have discussed previously. Thus we have accomplished our task – showing that the formal expression of a covariant field theory of the inertia of matter takes the form of quantum mechanics in the curved spacetime, and then in approximation, in the linear limit of special relativity, the Dirac equation, and in the nonrelativistic approximation, Schrödinger's wave equation.

In the limit, wherein there is a very weak coupling to the i th component matter field, $\eta^{(i)}$ of a closed system of matter, the interaction functional $I(\eta, \chi)$ 'loses sight' of its matter field dependence and may be approximated with the ('smeared') function in the space and time coordinates, $I(x)$. In the final (in principle unachievable) limit, where the interaction of the i th matter field and the rest of the matter of the closed system vanishes, i.e., $I(x) \rightarrow 0$, one is left with the vacuum equation for the i th matter field.

In both of the above limits, $I(\eta) \rightarrow I(x)$ and $I \rightarrow 0$, the nonlinear spinor matter field equations (4.22a) and (4.22b) take on the linear approximation, as in ordinary quantum mechanics. In this limit only is the principle of linear superposition applicable. It is here that one may use the Hilbert function space, as we discussed earlier in terms of a probability calculus, as a linear approximation for the function space of nonlinear fields of general relativity.

4.6 The Elementary Interaction

The physical reason for the necessary incorporation of the equation of continuity in the generally covariant field theory of matter, and thus the requirement of gauge symmetry, is the interpretation of the spinor matter field η as a weighting amplitude for the interaction within the closed system of matter. The elementarity of 'interaction' here replaces the elementarity of 'particle' in the atomistic theories, such as the quantum theory. The conservation of interaction then replaces the conservation of particle in the holistic view of the field theory of

matter in general relativity. This feature will be stressed in Chap. 5, where we will concentrate on the many-body matter system and how this nonlinear field theory, even in the special relativity limit, differs from the linear quantum theory for many body systems. A particular derivation in Chap. 5 will focus on the origin of the Pauli exclusion principle.

4.7 Proof of the Attractive Gravitational Force in the Newtonian Limit and the Oscillating Universe Cosmology

The geometrical relationship between the spinor matter fields (η, χ) is in terms of the continuous field variable $\lambda(x)$ – the modulus of a complex function (4.14). Thus the variable λ , which plays the role in this theory of the inertial mass of matter m , is positive definite – it can only be positive under all circumstances.

In the Newtonian limit of this theory applied to gravitation, the force of gravity between the i th and j th matter components of a closed system is proportional to the product of ‘mass’ components λ_i and λ_j . Since they each have positive value, the Newtonian force of gravity, $F_G = G\lambda_i\lambda_j/r_{ij}^2$ can have only one polarity. As we structure a macro-quantity of matter, such as the Sun, its mass must be the sum of micromasses: $m_S = \sum_i \lambda_i$, and the mass of an interacting body, such as the Earth, is the sum over its micromasses: $m_E = \sum_j \lambda_j$. The magnitude of the gravitational force between the Sun and the Earth is $Gm_S m_E / r_{SE}^2$, where r_{SE} is the separation between the centers of mass of the Sun and the Earth. G entails the sign of the gravitational force: attractive if it is negative and repulsive if it is positive. Because the micromasses that make up the mass of the Sun and those that make up the mass of the Earth are always positive, macromasses of the Sun and the Earth, m_S and m_E , can only be positive. Thus, the sign of G determines the sign of the gravitational force. If it is seen to be negative (attractive force) in one case it must be attractive in all cases. Because of the empirical observation of the attractive nature of the Newtonian gravitational force in, say, the interaction between the Sun and the Earth, it must be attractive in all cases.

This is a proof of the attractiveness of the Newtonian gravitational force, as a limit of the gravitational force in general relativity. This result has never been proven before, either in Newton’s theory of universal gravitation or in Einstein’s tensor formulation of general relativity.

Nevertheless, when we are not in the Newtonian approximation, the gravitational force in general relativity relates to the affine connection field $\Gamma_{\mu\nu}^\rho$, as it appears in the geodesic equation (2.21), interpreted as an equation of motion of a material body. Since the affine connection is not a positive-definite function, it could be positive (denoting a repulsive force) under one set of physical conditions, and negative (denoting an attractive force) under other physical conditions

If the matter of the closed system is sufficiently dense and the relative speeds of interacting matter are close to the speed of light, the dominant force between massive bodies would be repulsive. This could explain the ‘expanding universe’ – the galaxies of the universe are seen to be moving away from each other – in a state of ‘explosion’. Then, as the matter of the universe becomes sufficiently rarefied, an inflection point is reached where the dominant repulsive force changes to a dominant attractive force. The matter of the universe would then implode until matter has become sufficiently dense, when there would be another inflection point, where the implosion changes to an explosion – the universe then expands once again, and so on ad infinitum. This is the oscillating universe cosmology. These ideas will be gone into further in Chap. 8, where we will discuss ideas of cosmology in the context of this field theory.

4.8 From the Mach Principle to the Generalized Mach Principle

As remarked earlier, the results of this analysis are in accordance with the requirement of the Mach principle. For as the environment of any ‘observed matter’ should be depleted toward a vanishing quantity of matter, the curvature of spacetime, representing the existence of the matter in the environment of the ‘observed body’ would correspondingly vanish. In accordance with the relationship (4.14), the matrix fields Ω_μ and $\Lambda_\pm$ would both vanish and the mass of the ‘observed matter’ vanishes, $\lambda \rightarrow 0$.

Of course, one cannot observe any manifestations of the mass of a single particle in a vacuum universe. But in principle one can observe the effect of depleting a material medium toward the vacuum state, to see if this causes a corresponding diminishing of the inertial mass of some observed matter – say an elementary particle in a high energy collision event.

In Chap. 3, on electromagnetic phenomena in general relativity, we saw that the electromagnetic field of an isolated charged body simi-

larly vanishes. This result supports the generalized Mach principle – the assertion that all previously considered ‘intrinsic’ properties of elementary matter are instead measures of coupling of this matter to its total environment. Thus if the environment of an observed element of matter vanishes, the coupling vanishes and the previously considered ‘intrinsic’ qualities of this matter correspondingly vanish. With this view, then, the atomistic model is rejected – all intrinsic properties of the atom of matter, such as inertial mass, electric charge, magnetic moment, etc., are ‘exorcised’, thereby supporting the continuum, holistic model of matter, according to the theory of general relativity.

5. Electromagnetism

After the eighteenth and nineteenth century discoveries of the relation between the electric charge and its motion on the one hand, and the electric and magnetic fields of force on the other, Faraday interpreted the fields of force, rather than singular sources of these fields, as the basic way of representing charged matter. Maxwell then formulated these relations between charged matter sources and the associated fields of force in terms of partial differential equations, now called Maxwell's equations. These were then discovered to be the laws of electromagnetism, including the phenomenon of optics.

As in the case of Einstein's field equations, Maxwell's equations are interpreted in the theory proposed in this monograph as identities – the changes in field intensities on the left-hand sides of these equations are another way of talking about the charge and its motion on the right-hand side, and vice versa. The source-free solutions of these equations (free radiation and the 'photon') are then rejected in this view as unphysical. They are replaced here with the solutions for charged matter, very far from their locations.

It is argued that the symmetry of the theory of relativity requires that when charged matter (an emitter) sends a signal to other charged matter (the absorber), R cm away, the absorber must simultaneously send a signal to the emitter. (This is called delayed-action-at-a-distance.) Both signals must then arrive simultaneously at the matter called the emitter and that called the absorber.

The underlying symmetry of relativity theory – without the reflections in space and time – implies that the expressions of the electromagnetic field equations must be in terms of spinor variables and quaternion operators, as in the case of gravity, discussed in the previous chapter. Thus, the factorization of Maxwell's equations from the standard vector form to a spinor form is demonstrated as the most general expression of the laws of electricity and magnetism in accordance with the symmetry group of relativity theory. Generalizations

then appear that reveal features of the weak interaction along with the standard electromagnetic interaction.

At this stage, we have demonstrated that a unification of the gravitational, inertial and electromagnetic manifestations of matter is in terms of a spinor/quaternion formalism that is prescribed by the symmetry group of the theory of relativity. Such unification implies that this is a single field for a closed system in which the gravitational and electromagnetic forces of matter on matter are implicit, as is the inertial manifestation of matter that is acted upon. There are also further hints of the implicit inclusion of the short-range weak and strong (nuclear) interactions in the nuclear domain.

5.1 Introduction

In the preceding chapters we have shown that the inertial, electromagnetic and gravitational features of matter emerge from the theory of general relativity as a truly unified field theory. We saw that the inertial feature of matter in particular underlies all of the force manifestations, including gravitation and electromagnetism as well as, in principle, other forces not yet explored in detail, such as the short range forces called weak and strong forces in the nuclear domain. In addition, as a bonus, the generally covariant field theory of inertia leads, in a linear approximation, to the Hilbert space structure of quantum mechanics. In this chapter, we will explore in detail the electromagnetic force manifestation of matter, exploiting the ideas discussed thus far:

- the generalized Mach principle,
- the implications of the irreducible symmetry group of the theory of relativity.

It was discovered by Michael Faraday and James Clerk Maxwell, in the 19th century, that a general expression for the laws of electromagnetism that incorporates all of the discovered empirical relations between the electric and the magnetic field intensities and the charge density and current density sources of these fields, has the form of Maxwell's field equations in special relativity:

$$\partial^\nu F_{\mu\nu} = 4\pi j_\mu , \quad (5.1a)$$

$$\partial^{[\rho} F_{\mu\nu]} = 0 , \quad (5.1b)$$

where the square bracket in (5.1b) denotes the cyclic sum over the indices, from 0 to 3. The zero on the right hand side of (5.1b) denotes

the absence of magnetic monopoles in this field theory, as we have shown in Sect. 3.5.2.

The global extension of the Maxwell equations (5.1a) and (5.1b) to the curved spacetime is as follows:

$$F_{\mu\nu}{}^{;\nu} = 4\pi j_\mu, \quad F_{[\mu\nu}{}^{;\rho]} = 0,$$

where the covariant derivative of the second-rank tensor is

$$F_{\alpha\beta}{}^{;\gamma} = \partial^\gamma F_{\alpha\beta} + \Gamma_{\alpha\gamma}{}^\rho F_{\rho\beta} + \Gamma_{\beta\gamma}{}^\rho F_{\alpha\rho},$$

and $\Gamma_{\alpha\gamma}{}^\rho$ are the ordinary affine connection coefficients.

The electromagnetic (antisymmetric) tensor in (5.1a) and (5.1b) is the following combination of electric and magnetic field components:

$$\begin{aligned} F_{12} = -F_{21} = H_3, \quad F_{13} = -F_{31} = -H_2, \quad F_{23} = -F_{32} = H_1, \\ F_{10} = -F_{01} = E_1, \quad F_{20} = -F_{02} = E_2, \quad F_{30} = -F_{03} = E_3, \\ F_{\mu\mu} = 0. \end{aligned} \quad (5.2)$$

Thus, in terms of the electric and magnetic variables $\mathbf{E}$ and $\mathbf{H}$, Maxwell's equations (5.1a) and (5.1b) take the vector form

$$\nabla \times \mathbf{H} - \partial^0 \mathbf{E} = 4\pi \mathbf{j}, \quad \nabla \cdot \mathbf{E} = 4\pi \rho, \quad (5.3a)$$

$$\nabla \times \mathbf{E} + \partial^0 \mathbf{H} = 0, \quad \nabla \cdot \mathbf{H} = 0. \quad (5.3b)$$

A manipulation of (5.3a) and (5.3b) yields the equations of conservation of energy ($E^2 + H^2$) called Poynting's equation:

$$\frac{1}{8\pi} \partial^0 (E^2 + H^2) + \frac{1}{4\pi} \nabla \cdot (\mathbf{E} \times \mathbf{H}) = -\mathbf{E} \cdot \mathbf{j}, \quad (5.4)$$

and the conservation of momentum ($\mathbf{E} \times \mathbf{H}$) equation:

$$\frac{1}{4\pi} \partial^0 (\mathbf{E} \times \mathbf{H}) = \rho \mathbf{E} + \mathbf{j} \times \mathbf{H}. \quad (5.5)$$

The right hand side of (5.5) is the Lorentz force density. It predicts how the charge density ρ and the current density $\mathbf{j}$ couple to the electric and magnetic fields. It is interesting that this equation had to be added to Maxwell's equations to predict the equation of motion of charged matter in an electromagnetic field. On the other hand, the equation of motion of matter in general relativity theory, under the influence of the gravitational field is inherent in the Einstein metrical field equations for gravity, i.e., no extra equations of motion need be added. The main reason for this is the nonlinearity of Einstein's field

equations in contrast with the linearity of Maxwell's field equations for electromagnetism.

Another form of Maxwell's equations (5.1a) and (5.1b) is in terms of the electromagnetic four-potential, as follows:

$$\square A_\mu = 4\pi j_\mu , \quad (5.6)$$

where $\square = (\partial^0)^2 - \nabla^2$ is the d'Alembertian (wave) operator. Equation (5.6) is subject to the Lorentz gauge:

$$\partial^\mu A_\mu = 0 . \quad (5.7)$$

In terms of this formulation, the electromagnetic tensor $F_{\mu\nu}$ is, by definition, the four-dimensional curl of the vector potential:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu . \quad (5.8)$$

It is the left-hand side of (5.8) that normally relates to the observable electric and magnetic fields, rather than the four potential. However, researches of Bohm and Aharanov¹ indicate an experimental situation in the quantum domain whereby there are non-zero observable components of the four-potential at locations where the electric and magnetic field intensities are zero.

The standard field equations for electromagnetism (5.1a) and (5.1b) constitute a vector representation of the theory because these equations transform as a four-vector representation of the symmetry group of relativity theory. These equations are covariant under the continuous transformations of space and time (the Poincaré group in special relativity) – which is the requirement of the theory of relativity – as well as reflections in space and time – which is not a requirement of the theory. Because of the latter inclusion of discrete transformations in the symmetry of the field equations (5.1a) and (5.1b), the symmetry group of these equations is reducible, with respect to the Poincaré group of special relativity theory. We will see later on that the irreducible group, without the reflections, leads to a factorization of the field equations (5.1a) and (5.1b) to a two-component spinor formalism for the laws of electromagnetism.

The invariants of the vector formalism for the Maxwell theory are

$$I_1(\text{scalar}) = F_{\mu\nu} F^{\mu\nu} \iff E^2 - H^2 , \quad (5.9a)$$

$$I_2(\text{pseudoscalar}) = \varepsilon_{\mu\nu\lambda\rho} F^{\mu\nu} F^{\lambda\rho} \iff \mathbf{E} \cdot \mathbf{H} , \quad (5.9b)$$

$$I_3 = j^\mu j_\mu \iff \rho^2 - j^2 , \quad (5.9c)$$

$$I'_3 = j^\mu A_\mu \iff \rho A_0 - \mathbf{j} \cdot \mathbf{A} , \quad (5.9d)$$

¹ Y. Aharanov and D. Bohm, Phys. Rev. **115**, 485 (1959). Also see: C.A. Mead, *Collective Electrodynamics* (MIT, 2000).

where the latter relates to the structure of I_3 since the solution A_μ transforms in the same way as the current density j_μ . That is to say, the particular solution of (5.6) is

$$A_\mu(x) = \int j_\mu(x') G(x - x') d^4 x' , \quad (5.10)$$

where $G(x - x')$ is the Green's function for d'Alembert's equation (5.6). This solution will be discussed further in the context of delayed action-at-a-distance in the concluding section of this chapter.

5.2 Interpretation of Maxwell's Equations in the Holistic Field Theory

This brings us to the interpretation of Maxwell's equations in the context of the holistic field theory for the closed system. Their logical meaning is that the field intensity solutions on the left-hand side of the equations (5.1a) and (5.1b) [or (5.3a) and (5.3b)] is another way of talking about the source fields, on the right-hand sides of these equations. That is to say, these equations are interpreted as identities. Thus, in the limiting case of the vacuum, where there are no sources anywhere, there could not be any meaningful electric and magnetic field solutions anywhere. The radiation solutions of Maxwell's equations – the homogeneous solutions – are then excluded as not being physically pertinent in this holistic field theory of a closed system. The logical implication is that the quantum of the source-free radiation field – the photon – is also excluded as a bona fide elementary particle in nature.

It follows, then, that in this model, the only meaningful solutions of Maxwell's equations for electromagnetism are the particular solutions. In terms of the vector potential form of these equations (5.6), the particular solutions are given in (5.10). The invariant (scalar) interaction – the interaction Lagrangian density – that leads to the Lorentz force density, has the structure:

$$I_3 = L_{\text{em}}(x) = j^\mu(x) A_\mu(x) = j^\mu(x) \int j_\mu(x') G(x - x') d^4 x' . \quad (5.11)$$

The total interaction Lagrangian then takes the form of a current-current coupling term:

$$\mathcal{L}_{\text{em}} = \int L_{\text{em}} d^4 x = \int j^\mu(x) j_\mu(x') G(x - x') d^4 x' d^4 x . \quad (5.12)$$

This portion of the electromagnetic interaction Lagrangian is then proportional to the universal constant e^2 , since each of the current density factors is proportional to the charge of the electron e . With the units we have chosen, $\hbar/2\pi = 1 = c$, the universal electromagnetic coupling constant is the fine structure constant $2\pi e^2/\hbar c \approx 1/137$.

5.3 The Elementary Interaction Formulation

In accordance with our preceding discussion wherein we interpret Maxwell's equations as identities, it follows that for each interacting component of the closed system, there must be a separate set of Maxwell equations. That is, in the form (5.3a) and (5.3b), we must insert a superscript on each of the fields of the interacting components, as follows:

$$\begin{aligned}\nabla \times \mathbf{H}^{(u)} - \partial^0 \mathbf{E}^{(u)} &= 4\pi \mathbf{j}^{(u)} , & \nabla \cdot \mathbf{E}^{(u)} &= 4\pi \rho^{(u)} , \\ \nabla \times \mathbf{E}^{(u)} + \partial^0 \mathbf{H}^{(u)} &= 0 , & \nabla \cdot \mathbf{H}^{(u)} &= 0 .\end{aligned}\quad (5.13)$$

The conservation laws (5.4) and (5.5) then take the form

$$\frac{1}{8\pi} \partial^0 \sum_v \sum_{u \neq v} \mathbf{E}^{(u)} \times \mathbf{H}^{(v)} = \sum_v \sum_{u \neq v} [\rho^{(u)} \mathbf{E}^{(v)} + \mathbf{j}^{(u)} \times \mathbf{H}^{(v)}], \quad (5.14a)$$

$$\begin{aligned}- \sum_v \sum_{u \neq v} \mathbf{E}^{(u)} \cdot \mathbf{j}^{(v)} &= \frac{1}{8\pi} \partial^0 \sum_v \sum_{u \neq v} [\mathbf{E}^{(u)} \cdot \mathbf{E}^{(v)} + \mathbf{H}^{(u)} \cdot \mathbf{H}^{(v)}] \\ &+ \frac{1}{4\pi} \nabla \cdot \sum_v \sum_{u \neq v} \mathbf{E}^{(u)} \times \mathbf{H}^{(v)} .\end{aligned}\quad (5.14b)$$

The restriction above that $u \neq v$ indicates that a field cannot interact with itself, i.e., there is no self-energy defined in this theory. This view, originally attributed to Faraday, eliminates the problem of the infinite self-energy of the electron, as Lorentz explored it in the context of the model of the discrete, singular electron.

Because of the linearity of the Maxwell field formalism, the net electromagnetic fields for the n -component closed system of charged matter, are the vector sums

$$\mathbf{E} = \sum_u \mathbf{E}^{(u)} , \quad \mathbf{H} = \sum_u \mathbf{H}^{(u)} \quad (u = 1, \dots, n) . \quad (5.15)$$

5.4 A Spinor Formulation of Electromagnetic Theory in Special Relativity

As we discussed earlier, the underlying symmetry group of the theory of special relativity is the 10-parameter Lie group – the Poincaré group. The Maxwell field formalism (5.1a) and (5.1b) [or equivalently (5.3a), (5.3b) and (5.6)], is a reducible representation of the group because it includes the discrete transformations in space and time – which is not a requirement of the theory – as well as the continuous transformations in space and time – which is a requirement of the theory.

Just as in Chap. 3, where we factorized Einstein's tensor formulation of general relativity (10 relations) by dropping the discrete reflections in space and time, to achieve the irreducible quaternion structure of the theory (16 relations), so we may factorize the vector representation of the Maxwell field theory (5.1a) and (5.1b) to achieve the spinor representation of electromagnetic theory. In relativistic quantum mechanics, this occurs for the same reason, wherein the Klein–Gordon equation factorizes to a pair of two-component spinor equations, as we saw in Chap. 4, revealing the (relativistic) Dirac electron equation, in the Majorana form of two coupled two-component spinor equations.

Consider the following structure of the electromagnetic field intensity in terms of the complex components

$$G_k = (\mathbf{H} + i\mathbf{E})_k, \quad G_0 = 0 \quad (k = 1, 2, 3),$$

where $\mathbf{H}$ is the magnetic field intensity and $\mathbf{E}$ is the electric field intensity of charged matter.

Let us now structure the two-component spinor variables as follows, in accordance with the columns of a quaternion Q whose zeroth component is zero:

$$Q(x^0 = 0, x^1, x^2, x^3) = \begin{pmatrix} x^3 & x^1 - ix^2 \\ x^1 + ix^2 & -x^3 \end{pmatrix} = \boldsymbol{\sigma} \cdot \mathbf{r}. \quad (5.16)$$

We will consider the identification between the vector G_k and the two-component spinor fields ϕ_1 and ϕ_2 , and the identification between the current density four-vector j_μ and the two-component spinor fields γ_1 and γ_2 as follows:

$$\begin{aligned} \phi_1 &= \begin{pmatrix} G_3 \\ G_1 + iG_2 \end{pmatrix}, & \gamma_1 &= -4\pi i \begin{pmatrix} \rho + j_3 \\ j_1 + ij_2 \end{pmatrix}, \\ \phi_2 &= \begin{pmatrix} G_1 - iG_2 \\ -G_3 \end{pmatrix}, & \gamma_2 &= -4\pi i \begin{pmatrix} j_1 - ij_2 \\ \rho - j_3 \end{pmatrix}. \end{aligned} \quad (5.17)$$

It is readily shown that the two-component spinor field equations

$$\sigma^\mu \partial_\mu \phi_\alpha(x) = \Upsilon_\alpha(x) , \quad (5.18)$$

where $\alpha = 1, 2$ are two independent two-component spinor equations, are in one-to-one correspondence with the Maxwell field equations (5.3a) and (5.3b), where the first order quaternion differential operator is

$$\sigma^\mu \partial_\mu = \sigma^0 \partial_0 - \boldsymbol{\sigma} \cdot \boldsymbol{\nabla} . \quad (5.19)$$

σ^0 is the unit two-dimensional matrix and σ^k ($k = 1, 2, 3$) are the three Pauli matrices. These are the four basis elements of a quaternion, as we saw in Chap. 3.

It is important to note at this juncture that the identification (5.17) between the two-component variables $(\phi_\alpha, \Upsilon_\alpha)$ of the spinor representation of electromagnetism and the variables $(\mathbf{E}, \mathbf{H}, \rho, \mathbf{j})$ of the vector representation are not form-invariant under the transformations of the Poincaré group. That is to say,

$$x \longrightarrow x' \implies \phi_\alpha(x) \longrightarrow \phi'_\alpha(x') \neq \phi'_\alpha(E', H') ,$$

where $(E', H') \equiv F'_{\mu\nu} = \alpha_\mu^\lambda \alpha_\nu^\rho F_{\lambda\rho}$ are the transformations of a second-rank (covariant) tensor and α_μ^λ are the vector transformations under the Poincaré group. This lack of form-invariance is, of course, due to the fact that the latter are transformations under the vector representations of the group and the former are inequivalent to these – they are the transformations under the spinor representations of the group, as we discussed in more detail in Chap. 3.

This does not raise any difficulties because the physical requirement is that it is the observables, in terms of the conservation laws and the invariants of the field theories, that must be duplicated in the different reference frames, but not the field equations themselves. This will be illustrated below.

5.5 Invariants and Conservation Equations in the Spinor Formalism

The spinor formulation of the field equations for electromagnetism (5.18) in special relativity is a nonabelian theory, since it entails non-commuting variables. It contains all of the physical predictions of the standard vector formulation. However, the spinor formulation (5.18)

entails additional conservation laws and invariants that have no counterpart in the usual (abelian) vector expression of electromagnetism. Thus this spinor formulation is a true generalization of the usual vector-tensor field theory of Maxwell.

In special relativity, covariance of the spinor field equations is maintained if and only if ϕ_α transforms under the Poincaré group as the matter spinor field η (Chap. 4) and the source field $\mathcal{Y}_\alpha$ transforms as its time-reversal χ . That is,

$$x \longrightarrow x' \implies \begin{cases} \phi_\alpha(x) \longrightarrow \phi'_\alpha(x') = S\phi_\alpha(x) , \\ \mathcal{Y}_\alpha(x) \longrightarrow \mathcal{Y}'_\alpha(x') = (S^\dagger)^{-1}\mathcal{Y}_\alpha(x) , \end{cases} \quad (5.20)$$

where the spin-1/2 representations $\{S\}$ of the Poincaré group are related to the Lorentz transformations $\{\alpha^\nu_\mu\}$ according to the (two-to-one) relation

$$S^\dagger \boldsymbol{\sigma}^\mu S = \alpha^\mu_\nu \boldsymbol{\sigma}^\nu , \quad (5.21)$$

with the (double-valued) solutions

$$S(\theta_{\mu\nu}) = \exp \left[\frac{1}{2} \boldsymbol{\sigma}^\mu \boldsymbol{\sigma}^\nu \theta_{\mu\nu} \right] . \quad (5.22)$$

In (5.22), the indices (μ, ν) are not summed and $\theta_{\mu\nu}$ are the parameters that characterize the continuous transformations of the group.

It follows from the algebra of spinors that the following forms are invariants:

$$J_1 = \phi_2^{\text{tr}} \varepsilon \phi_1 \quad \text{and} \quad J_2 = \mathcal{Y}_2^{\text{tr}} \varepsilon \mathcal{Y}_1 , \quad (5.23)$$

where ε is the two-dimensional Levi-Civita symbol. Substitution of (5.17) into (5.23) gives the following identification:

$$J_1 = E^2 + H^2 + 2i\mathbf{E} \cdot \mathbf{H} , \quad J_2 = j^2 - \rho^2 . \quad (5.24)$$

These terms correspond with the standard invariants of electromagnetic theory – the real part of J_1 is the scalar invariant and the imaginary part is the pseudoscalar invariant of the conventional formalism. They appear here as a single complex invariant because of the rejection of reflection symmetry in this formalism. The second invariant J_2 is the modulus of the current density four-vector of the standard Maxwell formalism.

In addition to the invariants J_1 and J_2 , the transformations (5.20) indicate the following four complex number (eight real number) invariants:

$$I_{\alpha\beta} = \phi_\alpha^\dagger \Upsilon_\beta \quad (\alpha, \beta = 1, 2) . \quad (5.25)$$

These have no counterpart in the standard Maxwell formalism.

Let us now investigate the conservation equations that emerge from these invariants. Multiplying (5.18) on the left with the Hermitian adjoint spinor $\phi_\beta^\dagger$, we have the scalar equation

$$\phi_\beta^\dagger \boldsymbol{\sigma}^\mu \partial_\mu \phi_\alpha = \phi_\beta^\dagger \Upsilon_\alpha . \quad (5.26)$$

Taking the Hermitian adjoint of this equation, interchanging the labels α and β and then adding the resulting equations yields the following four complex-number-valued conservation equations:

$$\partial_\mu (\phi_\beta^\dagger \boldsymbol{\sigma}^\mu \phi_\alpha) = \phi_\beta^\dagger \Upsilon_\alpha + \Upsilon_\beta^\dagger \phi_\alpha . \quad (5.27)$$

The right-hand side of (5.27) (eight real number conservation equations) has the dimension of a force density. This should be compared with the single force density term – the four-vector source – on the right-hand side of (5.4) and (5.5) in the standard formulation.

To exhibit the numerical correspondence between the conservation equations (5.27) and (5.4) and (5.5), consider the sum of equations (5.27) with $\alpha = \beta = 1$ and with $\alpha = \beta = 2$. It is readily verified that this sum yields the energy conservation equation (5.4). Other linear combinations yield the momentum conservation equation (5.5).

It should be re-emphasized that the eight real number force density equations on the right-hand side of (5.27) are separately scalars. That is, they have the same form, individually, in all Lorentz frames, rather than transforming together, as in the standard vector formalism, as a four-vector.

The generalization from the vector formalism of the Maxwell theory to the spinor formalism will lead to extra predictions of the theory. These will be discussed further in Chap. 7, in regard to atomic and elementary particle physics applications.

5.6 Lagrangian for the Spinor Formulation of Electromagnetism

The principle of least action for the electromagnetic laws is as follows:

$$\delta(S_{\text{em}} + S_{\text{m}}) = 0 ,$$

where S_m is the matter part of the action function, depending on the gravitational and inertial variables, and S_{em} is the electromagnetic part of the action, with the form

$$S_{em} = \int L_{em} d^4x = ig_M \int \sum_{\alpha} [(-1)^{\alpha} \phi_{\alpha}^{\dagger} (\sigma^{\mu} \partial_{\mu} \phi_{\alpha} - 2\Upsilon_{\alpha}) + \text{h.c.}] d^4x, \quad (5.28)$$

where the sum is taken over the two terms $\alpha = 1, 2$.

The variation of L_{em} with respect to the spinor variables $\phi_{\alpha}, \phi_{\alpha}^{\dagger}$ and their derivatives, as the variational parameters, then yields the spinor field equations (5.18) for electromagnetism.

The fundamental constant g_M does not appear explicitly in the latter equations since it occurs on both sides of the equation (5.18) so derived. But this constant does appear in the matter field equations in terms that have no counterpart in the standard unification of the matter field equations and those of electromagnetism. When the total Lagrangian containing the part in (5.18) is varied so as to yield the matter field equations (4.22a) and (4.22b), by variation with respect to the matter variables (η, χ) and their derivatives as the variational parameters, these variables implicit in the electromagnetic source fields Υ_{α} , that is, in the part of the Lagrangian

$$2ig_M \sum_{\alpha} (-1)^{\alpha} \phi_{\alpha}^{\dagger} \Upsilon_{\alpha} + \text{h.c.}, \quad (5.29)$$

contribute a term in the matter field equations depending on g_M that we will explore in detail later on. It is shown in Chap. 7 to predict the Lamb shift in hydrogenic atoms, in numerical agreement with the experimental data, as well as other physical implications.

The field in (5.29) to the right of the summation sign has the dimension of energy density/length. Thus, to give this term the correct dimension of energy density, the universal constant g_M must have the dimension of length. We will see in Chap. 7 that its numerical value is

$$g_M = (2.087 \pm 0.001) \times 10^{-14} \text{ cm}. \quad (5.30)$$

Thus, a fundamental length appears in the theory as a consequence of the factorization of the vector-tensor formalism of the Maxwell theory to the spinor-quaternion formalism. This is a generalization that is indicated by the irreducible representations of the symmetry group of the theory of relativity. This is not an ad hoc generalization. It is a feature that was present at the outset in the irreducible form of the algebraic structure of the theory of relativity.

5.7 Faraday's Approach and the Mach Principle

According to Faraday's interpretation of the field as primary for each constituent of a closed system of charged matter, as well as Mach's rejection of any implicit qualities of the 'things' of matter, each of the electromagnetic field equations must have its own 'label'. Thus the spinor field equations (5.18), for the u th component of the closed system must be expressed as

$$\sigma^\mu \partial_\mu \phi_\alpha^{(u)} = \Upsilon_\alpha^{(u)} \quad (u = 1, 2, \dots, n). \quad (5.31)$$

The conservation equations that accompany this spinor field equation are

$$\partial_\mu \sum_v \sum_{u \neq v} \phi_\beta^{(u)\dagger} \sigma^\mu \phi_\alpha^{(v)} = \sum_v \sum_{u \neq v} [\phi_\alpha^{(u)} \Upsilon_\beta^{(v)} + \Upsilon_\beta^{(u)\dagger} \phi_\alpha^{(v)}]. \quad (5.32)$$

It is interesting to note that four of the eight real-number-valued conservation equations (5.27) – without the indexing (u, v) – reduce to $0 = 0$. This is because of the particular representation chosen, where the imaginary part of the four complex equations [with $\alpha, \beta = (1, 1), (1, 2), (2, 1), (2, 2)$] are all identically zero. However, invoking the Faraday interpretation with the (u, v) indexing, as in (5.31), the full set of eight real-number-valued conservation equations is recovered. This is discussed further in [117], Sect. 5.7.

5.8 Spinor Formulation of Electromagnetism in General Relativity

One may globally extend the preceding spinor formulation in special relativity to the global domain of general relativity by:

- replacing the constant quaternion basis elements – the unit two-dimensional matrix and the three Pauli matrices – by the quaternion fields $\{q^0(x), q^k(x)\}$ ($k = 1, 2, 3$),
- replacing the ordinary derivatives in the flat spacetime by covariant derivatives in the curved spacetime.

That is, in the global extension from the flat spacetime,

$$\sigma^\mu \partial_\mu \phi_\alpha \longrightarrow q^\mu(x) \phi_{\alpha;\mu} \equiv q^\mu(x) (\partial_\mu + \Omega_\mu) \phi_\alpha = \Upsilon_\alpha. \quad (5.33)$$

The covariance of these field equations is prescribed by the spin-1/2 representations $\{S(x)\}$ of the Einstein group, as follows:

$$x \longrightarrow x' \implies \begin{cases} \phi_\alpha(x) \longrightarrow \phi'_\alpha(x') = S(x)\phi_\alpha(x) , \\ \gamma_\alpha(x) \longrightarrow \gamma'_\alpha(x') = [S(x)^\dagger]^{-1}\gamma_\alpha(x) , \end{cases} \quad (5.34)$$

where $S(x)$ is a solution of the global equation in the curved spacetime:

$$[S(x)]^\dagger q^{\nu'}(x') \frac{\partial x^\mu}{\partial x^{\nu'}} S(x) = q^\mu(x) . \quad (5.35)$$

The solutions $S(x)$ are the global extension of the spin-1/2 representations of the Poincaré group (5.22), as follows:

$$S_{\mu\nu} = \exp \left[\frac{1}{2} q^\mu(x) q^\nu(x) \theta_{\mu\nu}(x) \right] . \quad (5.36)$$

In this equation, $\mu, \nu = 0, 1, 2, 3$ are not summed.

The 16-parameters of these spin-1/2 representations of the Einstein group $\theta_{\mu\nu}(x)$ are spacetime dependent, whereas in the flat spacetime limit (special relativity), they are constants, $\Omega_\nu \rightarrow 0$ and $q^\mu(x) \rightarrow \sigma^\mu$.

We will see in Chap. 7 that the expansion of the irreducible representations $S(x)$, (5.36), about those of the Poincaré group, (5.22), as the zeroth approximation, leads in the first approximation to the ‘quantization’ of electric charge.

5.9 Extension of the Spinor Conservation Laws of Electromagnetism in General Relativity

Consider the linear combinations of the four complex invariants $I_{\alpha\beta} = \phi_\alpha^\dagger \gamma_\beta$, with

$$\Phi(0) = -\frac{1}{8\pi} \sum_v \sum_{u \neq v} [\phi_1^{(u)\dagger} \gamma_1^{(v)} + \phi_2^{(u)} \gamma_2^{(v)}] , \quad (5.37a)$$

$$\Phi(1) = -\frac{1}{8\pi} \sum_v \sum_{u \neq v} [\phi_2^{(u)\dagger} \gamma_1^{(v)} + \phi_1^{(u)\dagger} \gamma_2^{(v)}] , \quad (5.37b)$$

$$\Phi(2) = \frac{i}{8\pi} \sum_v \sum_{u \neq v} [\phi_2^{(u)\dagger} \gamma_1^{(v)} - \phi_1^{(u)\dagger} \gamma_2^{(v)}] , \quad (5.37c)$$

$$\Phi(3) = \frac{1}{8\pi} \sum_v \sum_{u \neq v} [\phi_1^{(u)\dagger} \gamma_1^{(v)} - \phi_2^{(u)\dagger} \gamma_2^{(v)}] . \quad (5.37d)$$

We can now make up the following four scalar complex invariants:

$$\Theta(\mu) = \frac{1}{2} \left\{ \Phi(\mu) + \Phi(\mu)^\dagger + i\xi [\Phi(\mu) - \Phi(\mu)^\dagger] \right\} , \quad (5.38)$$

where ξ is a coefficient to be determined by experiment. It measures the relative strength of the real (scalar) and imaginary (pseudoscalar) components of the complex force invariants.

The generalized conservation equation in general relativity then has the form

$$\partial_\nu \zeta(\nu, \mu) + \Pi(\mu) + i\xi \Xi(\mu) - \Theta(\mu) = 0, \quad (5.39)$$

where ζ, Π, Ξ are functions of the electromagnetic spinor and quaternion metrical field variables. These terms are spelled out in detail in [117], Sect. 5.12.

A typical set of terms in (5.38) is for $\mu = 0$, as follows:

$$\zeta(\nu, 0) = -\frac{1}{16\pi} \sum_v \sum_{u \neq v} [\phi_1^{(u)\dagger} q^\nu \phi_1^{(v)} + \phi_2^{(u)\dagger} q^\nu \phi_2^{(v)}], \quad (5.40a)$$

$$\Pi(0) = -\frac{1}{16\pi} \sum_v \sum_{u \neq v} \sum_\alpha \phi_\alpha^{(u)\dagger} \Lambda_+ \phi_\alpha^{(v)}, \quad (5.40b)$$

$$\Xi(0) = -\frac{1}{16\pi} \sum_v \sum_{u \neq v} \sum_\alpha \left[[\phi_\alpha^{(u)} | \partial_\mu | q^\mu \phi_\alpha^{(v)}] + \phi_\alpha^{(u)\dagger} \Lambda_- \phi_\alpha^{(v)} \right], \quad (5.40c)$$

where $\Lambda_\pm = q^\mu \Omega_\mu \pm \text{h.c.}$ relates to the inertial mass field (Chap. 4) and

$$[\phi_\alpha^{(u)} | \partial_\mu | q^\mu \phi_\beta^{(v)}] \equiv \phi_\alpha^{(u)\dagger} q^\mu \partial_\mu \phi_\beta^{(v)} - \partial_\mu \phi_\alpha^{(u)\dagger} q^\mu \phi_\beta^{(v)},$$

where $\alpha, \beta = 1, 2$ and $\mu, \nu = 0, 1, 2, 3$ refer to the spacetime coordinates.

Under the conditions whereby $\xi \ll 1$, so that we may neglect the (pseudoscalar) Ξ term in (5.39), this equation reduces to the following conservation equation:

$$[\zeta(\nu, \mu)]_{;\nu} - \frac{1}{2} [\theta(\mu) + \theta(\mu)^\dagger] = 0. \quad (5.41)$$

This is to be compared with the standard electromagnetic conservation equation in special relativity, in the tensor formulation:

$$\partial_\mu T_\nu^\mu - F_\nu^{(+)} = 0, \quad (5.42)$$

where T_ν^μ is the energy-momentum tensor and $F_\nu^{(+)}$ is the four-Lorentz force density, with

$$F_0 = \mathbf{j} \cdot \mathbf{E}, \quad F_k = \rho E_k + (\mathbf{j} \times \mathbf{H})_k \quad (k = 1, 2, 3). \quad (5.43)$$

Substituting the identification (5.17) between the spinor and vector variables into (5.41), it is found that

$$\frac{1}{2}[\theta(\mu) + \theta(\mu)^\dagger] \Longleftrightarrow F_\mu^{(+)} ,$$

although of course the left-hand side of this correspondence is a scalar, for each of the four values of μ , while the right-hand side – the standard Lorentz force density – is a four-vector. The spatial part of this four-vector is a polar vector.

Substituting (5.17) into the imaginary part of the complex function θ , (5.38), it is found that

$$\frac{i}{2}\xi[\theta(\mu) - \theta(\mu)^\dagger] \Longleftrightarrow F_\mu^{(-)} , \quad (5.44)$$

where

$$F_0^{(-)} = -\mathbf{j} \cdot \mathbf{H} , \quad F_k^{(-)} = -\rho H_k + (\mathbf{j} \times \mathbf{E})_k \quad (k = 1, 2, 3) . \quad (5.45)$$

This is the anti-Lorentz force density. The spatial part of this four-vector is an axial vector. The four-vector densities $F^{(+)}$ and $F^{(-)}$ have opposite reflection properties in space and time. An experimental confirmation of the axial vector part of the total Lorentz force density (5.44) would mean that the electromagnetic interactions are not symmetric with respect to spatial and temporal reflections, i.e., parity and time reversal symmetries in electromagnetic interactions would be violated. This is in complete conformity with the underlying group of relativity theory, that it is a continuous group, without any reflection symmetry elements.

Note that the anti-Lorentz force replaces the electric field $\mathbf{E}$ in $F^{(+)}$ with $\mathbf{H}$ and the magnetic field $\mathbf{H}$ in $F^{(-)}$ with $-\mathbf{E}$. It predicts, for example, that charged matter would move along the lines of the magnetic field intensity and that an atom with a magnetic moment $\boldsymbol{\mu}$ would align with an external electric field with the interaction $\boldsymbol{\mu} \cdot \mathbf{E}$. The latter predicts that, in nuclear physics, the neutron would have an electric dipole moment, appearing in the Hamiltonian as the term proportional to $\boldsymbol{\sigma} \cdot \mathbf{E}$. However, experiment reveals that these terms are too small to see (i.e., $\xi \ll 1$).²

² N.F. Ramsey, Phys. Rep. **43**, 409 (1978).

5.10 The Electromagnetic Interaction Functional in the Matter Field Equations

To identify the electromagnetic source field $\mathcal{Y}_\alpha^{(u)}$ in the Lagrangian density (5.28) with the matter field spinor variables $\eta^{(u)}$, we start with the identification of the current density j^μ and the matter field variables. Since, in special relativity, the current density satisfies the continuity equation $\partial_\mu j^\mu = 0$, and the bilinear form $\eta^{(u)\dagger} q^\mu \eta^{(u)}$ satisfies the same equation, the identification sought is as follows:

$$j^{\mu(u)} \Longleftrightarrow e \left(\eta^{(u)\dagger} q^\mu \eta^{(u)} \right) , \quad (5.46)$$

as it is done in the Dirac theory. In the (more usual) Dirac bispinor form,

$$j^{\mu(u)} \Longleftrightarrow e \left(\psi^{(u)\dagger} \gamma^0 \gamma^\mu \psi^{(u)} \right) , \quad (5.47)$$

where, as discussed earlier, ψ is the four-component bispinor whose top two components are $\eta + \chi$ and whose bottom two components are $\eta - \chi$.

Using the identification (5.17) between the spinor and vector variables of electromagnetism and substituting the components of the four-current density (5.46) into the Lagrangian density,

$$L''_{\text{em}} = i g_M \sum_v \sum_{u \neq v} \sum_\alpha (-1)^\alpha \phi_\alpha^{(u)\dagger} \mathcal{Y}_\alpha + \text{h.c.} . \quad (5.48)$$

The variation of the latter with respect to the matter variables $\eta^{(u)\dagger}$ yields the following contribution to the matter field equation for the field $\eta^{(u)}$ in general relativity:

$$I''_u = \sum_{v \neq u} 32\pi g_M e \left[\phi_1^{(v)}(1)^* (-q^0 + q^3) + \phi_1^{(v)}(2)^* (q^1 + i q^2) - \phi_2^{(v)}(1)^* (-q^1 + i q^2) - \phi_2^{(v)}(2)^* (q^0 + q^3) \right] + \text{h.c.} , \quad (5.49)$$

where $\phi_\alpha(1)$ and $\phi_\alpha(2)$ are the separate two-components of the spinors.

To determine the special relativistic limit of this functional, we replace $q^\mu(x)$ with σ^μ . Note that I''_u depends on the constant $e^2 g_M$, since the electromagnetic field intensity spinor ϕ_α is proportional to the electric charge e . It is this functional that we will see (in Chap. 7) leads to the correct prediction of the Lamb shift in the states of hydrogen.

The most influential part of the electromagnetic interaction functional in the matter field equations comes from the contribution of the vector potential A_μ . The part of the Lagrangian that leads to the Lorentz force (5.43) is the vector coupling term

$$L'_{\text{em}} = \sum_{u \neq v} j^{\mu(u)} A_\mu^{(v)} . \quad (5.50)$$

As discussed in Chap. 4, the imposition of gauge invariance necessitates the inclusion of this part of the Lagrangian.

The four-vector potential $A_\mu^{(u)}$ of the electromagnetic field equations (5.6) is (5.10). As remarked earlier, these are the particular solutions for the Maxwell theory. The homogeneous solutions (the radiation solutions) are excluded in accordance with the interpretation of the electromagnetic field equations as identities for a closed system of charged matter, and the requirement of the generalized Mach principle.

With the vector potential in the Lagrangian density (5.50), and the identification of the current density with the matter fields according to (5.46) or (5.47), it follows that the variation of L'_{em} with respect to the matter spinor variables yields the interaction terms in the two-component spinor matter field equations as follows:

$$I'_u \eta^{(u)}(x) = e^2 \left[q^\mu \int \sum_{v \neq u} \eta^{(v)\dagger} q_\mu \eta^{(v)}(x') G(x - x') d^4 x' \right] \eta^{(u)}(x) . \quad (5.51)$$

In the bispinor form for the matter field, it has the form

$$I'_u \psi^{(u)}(x) = e^2 \left[\gamma^\mu \int \sum_{v \neq u} \psi^{(v)\dagger} \gamma^0 \gamma^\mu \psi^{(v)}(x') G(x - x') d^4 x' \right] \psi^{(u)}(x) . \quad (5.52)$$

5.11 Delayed Action at a Distance

The symmetric Green's function in the solution (5.10)

$$G(x - x') = \frac{1}{2} |\mathbf{r} - \mathbf{r}'| \left\{ \delta[(x^0 - x'^0) - |\mathbf{r} - \mathbf{r}'|] + \delta[(x^0 - x'^0) + |\mathbf{r} - \mathbf{r}'|] \right\} , \quad (5.53)$$

where δ is the Dirac delta function, is an average of two Green's functions – one that yields the retarded potential and one that yields the advanced potential. In classical radiation theory, one only selects the retarded potential. This would be consistent with the mechanism whereby an emitter sends a signal at the time t , and it is received by the absorber $R = |\mathbf{r} - \mathbf{r}'|$ cm away at the later time $R/c + t$. The argument is that, although the advanced potential is a solution of the field equations, it must be rejected because (in the particle model of emitter and absorber) it would describe an emitter responding to the absorber at the earlier time $R/c - t$, thus implying that the 'future' can affect the 'past'. On the other hand, the field theory represented in the theory of relativity entails a closed system where there are only coupled fields, all mapped in the same space and time. With this view, there are not separate times for the emission and absorption of signals between them, as is the case in the discrete particle model theories.

The reason for selecting the averaged Green's function above, as the unique solution for this problem, is the requirement of relativistic covariance. That is, the theory must be invariant under the interchange of the variables of the emitter and the absorber, each being mapped in a single space and time.

Summing up, the electromagnetic interaction functional I_u in the u th spinor matter field equation is the sum of two parts:

$$I_u = I'_u + I''_u, \quad (5.54)$$

where I'_u is given in (5.51) and I''_u is given in (5.49). The former is equivalent to the standard term $-ie\sigma^\mu A_\mu$ in the Dirac theory, except that in the present approach there is no self-energy present. We will see in Chap. 7 that the interaction functional I_u predicts the entire hydrogen spectrum, including the Lamb shift. The latter contribution to the interaction functional, I''_u , which has no counterpart in the standard Dirac theory, is responsible for this result, as well as other new results in this formulation of electrodynamics.

6. The Pauli Principle and Pair Creation/Annihilation

The field variable representing a single body in quantum mechanics is the wave function $\psi(x)$, whose absolute square is interpreted as the probability density for locating a particle at the place x . The extension to many particles of matter leads to the many-body wave function Ψ , depending on the wave functions of the individual bodies, $\psi_1(x_1), \psi_2(x_2), \dots, \psi_n(x_n)$, for an n -body system of matter, where x_n denotes the location of the n th body.

The generalization of the many-body wave function in this field theory of matter is the interaction field, a relational function of all of the matter fields of a closed system, each defined at the same spacetime point x . This is a solution of an equation interpreted as the conservation of interaction. It replaces the law of conservation of probability and conservation of particles of the standard quantum mechanical view. But the linear limit of the interaction field of this theory is the totally antisymmetrized many-body wave function of quantum mechanics.

It is proven rigorously in this chapter, without approximation, that the interaction field for the closed system vanishes identically if any two of the (spinor matter field) components are:

- in the same state of motion,
- have equal masses,

and

- have equal charges.

This is equivalent to the statement of the Pauli exclusion principle. It leads to the correct ordering of the Mendeleyev Periodic Table, the properties of metals, scattering phenomena, etc.

In the second part of this chapter, we pursue the implication of the conservation of interaction that matter cannot annihilate into a vacuum or be created out of a vacuum, at any time. Instead, a solution is demonstrated, corresponding to the data interpreted as pair annihilation, whereby the particle and antiparticle go into their true ground

state of maximum binding, but they do not annihilate. This predicted state is at zero energy – the ground state of the pair – relative to the unbound state that is at $2mc^2$ above this ground state. That is, in the ‘annihilation state’ the pair of particles does indeed exist, but they are invisible to an observer because they are so tightly bound that they do not impart energy and momentum to their surroundings, so as to create tracks, say in a cloud chamber. Delivering $2mc^2$ units of energy to this pair should ionize it. The particle and antiparticle would then be (almost) free of each other and could separately interact with their surroundings, creating tracks in the cloud chamber and giving rise to the data interpreted as pair creation. But matter here is not truly annihilated into a vacuum, or created out of a vacuum. These dynamical processes are fully deterministic.

The conclusion is then reached that the actual physical vacuum is a dense sea of such pairs in their ground states of null energy (and null linear and angular momentum). Such a background matter field was found to be responsible for the data on

- blackbody radiation,
- the inertial masses of observed matter,
- spontaneous decay of excited states of atoms and molecules.

6.1 Introduction

In the preceding chapters we have studied the implications of the curved spacetime of general relativity for the laws of matter – the unification of the gravitational and electromagnetic force manifestations of matter and its inertial manifestation. We saw that the naturally incorporated generally covariant field theory of the inertia of matter yields the quantum mechanical formalism in the linear limit.

As we asymptotically approach the linear limit of this field theory in special relativity, there are still features of the theory that remain incompatible with the formal expression and interpretation of quantum mechanics. A major difference is the natural nonlinear structure of the mathematical expression in relativity theory (special or general), compared with the natural linear structure of quantum mechanics. This field theory is nonlinear because it represents a closed system in terms of many correlated matter fields, all mapped in a single spacetime x . On the other hand, quantum mechanics is based on a linear model of an open system of n discrete particles of matter. The linearity is essential because of the interpretation of this theory in terms of a

probability calculus. The latter then entails a $4n$ -dimensional space in special relativity for the n coupled particle fields, rather than a single 4-dimensional spacetime in the approach of relativity theory, wherein the n matter fields are mapped. It is, generally, a holistic model in the theory of relativity (whether in its special or general form).

In the particle view, in quantum mechanics, as soon as one considers the interactions of the n coupled particle fields, each in its own 4-dimensional spacetime, one has the problem of ‘entanglement’ of the total $4n$ -dimensional system. However, with the holistic model of matter in special relativity, there is no ‘entanglement’ problem, since all matter fields are mapped in the same 4-dimensional spacetime x .

As we saw in Chap. 4, the two-component spinor matter fields $(\eta^{(u)}, \chi^{(u)})$ for the u th constituent of the closed material system solve the following coupled two-component spinor equations (the Dirac/Majorana form) in the special relativity limit:

$$\sigma^\mu \partial_\mu \eta^{(u)}(x) + \lambda^{(u)} \chi^{(u)}(x) = -I_u \eta^{(u)}(x) \quad (u = 1, 2, \dots, n), \quad (6.1)$$

where $\sigma^\mu = \{\sigma^0, \sigma^k\}$, σ^k are the Pauli matrices, and σ^0 is the unit two-dimensional matrix.

Its time-reversed (or space-reflected) equation is

$$\sigma^{*\mu} \partial_\mu \chi^{(u)}(x) + \lambda^{(u)} \eta^{(u)}(x) = -I_u^* \chi^{(u)}(x), \quad (6.2)$$

where $\sigma^{*0} = -\sigma^0$ and $\sigma^{*k} = \sigma^k$.

Equation (6.1) [or its time reversed form (6.2)] is the irreducible form of the u th matter field equation (since it is not reflection symmetric) out of the n coupled matter field equations of the closed system. The interaction field functional

$$I_u = I_u(\eta^{(1)}, \eta^{(2)}, \dots, \eta^{(u-1)}, \eta^{(u+1)}, \dots, \eta^{(n)})$$

is the essential interaction coupling to the u th matter field, depending on all other matter fields of the closed system, in accordance with the generalized Mach principle. I_u^* is the time reversal (or space reflection) of I_u .

If the matter field equations are space and time reflection symmetric, they can be expressed with the Dirac bispinor notation:

$$[\gamma^\mu \partial_\mu + \lambda^{(u)}] \psi^{(u)}(x) = -I_u \psi^{(u)}(x) \quad (u = 1, 2, \dots, n), \quad (6.3)$$

where γ^μ are the four-dimensional Dirac matrices (discussed in Chap. 4) and the bispinor solutions $\psi^{(u)}$ have the upper two components equal

to $(\eta + \chi)$ and the bottom two components equal to $(\eta - \chi)$. We will use this notation for the matter field equations in the remainder of this chapter.

The u th matter field equation (6.3) is fundamentally nonlinear because of the necessary appearance of the interaction functional I_u which in turn depends on the other matter fields $\psi^{(v \neq u)}(x)$ of the closed system. These solve equations that depend on $\psi^{(u)}(x)$ through the interaction terms $I_{v \neq u}$. Thus, the field operator acting on $\psi^{(u)}$ depends on $\psi^{(u)}$ itself. These equations are then nonlinear, by virtue of their being a part of a complete set of equations for the closed system. This is in contrast with the quantum mechanical equations that are fundamentally linear – so as to admit the principle of linear superposition as a basic rule to interpret this formalism as a probability calculus.

6.2 The Individual Particle Model

In particular limits, one may accurately approximate the nonlinear matter field equations (6.3) by linear differential equations. This is analogous to approximating Einstein's gravitational equations in general relativity by Newton's equations for his law of universal gravitation. The limit occurs when there is sufficiently small energy-momentum transfer between the interacting components of the closed material system. In this limit, one loses sight of the dependence of the interaction functional on the individual matter field components, wherein

$$I_u[\psi^{(1)}(x), \dots, \psi^{(u-1)}(x), \psi^{(u+1)}(x), \dots, \psi^{(n)}(x)]\psi^{(u)} \longrightarrow I_u(x)\psi^{(u)}.$$

In nuclear theory, this is called the individual particle model, wherein one views the constituent nucleons of a nucleus, one at a time, with the average background field $I_u(x)$ determining its behavior. In this limit, then, the matter field equations in $\psi^{(u)}$ linearize.

6.3 The Free Field Limit

In the asymptotic limit, when one may ignore the influence of the background field I_u altogether, i.e., when $I_u \rightarrow 0$, we have the free field approximation. In this case, there is a total uncoupling of the underlying matter fields of the closed system. The Dirac matter field equation then has the form

$$[\gamma^\mu \partial_\mu^{(u)} + \lambda^{(u)}]\psi^{(u)}(x_u) = 0, \quad (6.4)$$

where $\partial_\mu^{(u)} = \partial/\partial x^{\mu(u)}$ are the derivatives with respect to the four-vector components of the spacetime coordinates of each of the u matter components separately.

Since each of the n spinor equations in the approximation (6.4) is a linear differential equation, the sum of all of these n equations is also a solution, and we then have the single differential equation

$$\left[\sum_u (\gamma^\mu \partial_\mu^{(u)} + \lambda^{(u)}) \right] \Psi(\psi^{(1)}(x_1), \dots, \psi^{(n)}(x_n)) = 0. \quad (6.5)$$

The solution of (6.5) is a limiting form of a functional that is generally a connective relation between the elements of the set of solutions for the closed system $\{\psi^{(n)}\}$ of the coupled, nonlinear equations (6.3). In this linear limit, one solution is the product function

$$\Psi = \prod_u \psi^{(u)}(x_u) \quad (u = 1, \dots, n). \quad (6.6)$$

Other solutions are the $n! - 1$ other equivalent product functions, since the product is independent of the order of the factors, and there are $n!$ permutations of this order.

6.4 Conservation of Interaction

The generalized Mach principle and the holistic ontology of this field theory of matter in general relativity imply the idea of the conservation of interaction. This is a conceptual replacement of the conservation of particles of the classical Newtonian and the quantum theories of matter. The explicit function that solves the law of conservation of interaction is to represent a connective relation between the component spinor matter fields $\{\psi^{(u)}\}$ of the closed material system.

Taking the interaction field amplitude $\Psi(\psi^{(1)}, \dots, \psi^{(n)})$ for an n -fold holistic material system also to transform as a spinor variable, the differential form of the law of conservation of interaction is the following equation of continuity (in the Dirac bispinor notation):

$$\partial^\mu (\Psi^\dagger \gamma^0 \gamma_\mu \Psi) = 0. \quad (6.7)$$

This equation of continuity implies that, within the observer's frame of reference,

$$\int \Psi^\dagger \Psi d\mathbf{r} = \text{constant in time}. \quad (6.8)$$

With the normalization of the interaction field amplitude Ψ , the positive-definite function $\Psi^\dagger\Psi$ may then be interpreted as a weighting function. It is to weight the total interaction within the closed material system at all points of space, in any given observer's reference frame.

An important empirical implication of the law of conservation of interaction is that matter and antimatter (such as electron and positron or proton and antiproton) cannot mutually annihilate into a vacuum, or be created out of a vacuum. This problem will be studied further in this chapter, where it will be shown that there is an exact solution for the coupled matter field equations, for the bound particle and antiparticle pair, wherein they are in a state that exhibits all of the physical characteristics that are normally attributed to pair annihilation. But they are not annihilated into a vacuum (or created out of a vacuum). In the ground state of this pair, which has an energy value at $2mc^2$ below the state where they would be 'free' of each other, they simply do not interact with their surroundings as two oppositely charged particles would.

Because in the limit of no internal interaction, the n coupled differential equations uncouple into n independent equations (6.4), which are each identical, the complete set of solutions for each of these equations must cover the same function space. Further, the requirement of square integrability implied by the conservation of interaction (6.7) implies that the individual matter field solutions must have the same eigenfunction character as the solutions of ordinary quantum mechanics. Thus, the condition of square integrability applies to the individual matter fields themselves, so that

$$\int \psi^{(u)\dagger}\psi^{(u)}d\mathbf{r} = \text{finite number} . \quad (6.9)$$

Taking account of the fact that the order of products in the solution (6.6) does not affect the predicted properties (when this is a good approximation for the actual nonlinear field theory solutions), we may take the linear combination of all permutations of such product functions, assuming that each of the terms has equal weighting, in representing the total interaction of the closed system. The general solution of (6.3) would then take the form

$$\Psi = \left(\frac{1}{n!}\right)^{1/2} \sum_P \exp(i\alpha_P) \prod_u \psi^{(u)}(x_u) , \quad (6.10)$$

where α_P is the phase associated with the P th permutation. Its value will be determined below from the exact form of the function that fully exploits the nonlinear structure of the unapproximated equations.

6.5 The Pauli Exclusion Principle

The Pauli exclusion principle is an empirical rule in the quantum theory of a many-body system. It is the assertion that out of an entire ensemble of particles, no two spin-1/2 particles may have the same equation of motion (be in the same place at the same time, with the same constants of the motion) if they have the same mass and charge. This rule accounts for the classification of atoms in Mendeleyev's Periodic Table, the properties of metals, the proper scattering amplitudes of spin-1/2 particles, etc.

The following question then arises in the context of the field theory advocated in this book: how, within the framework of a holistic field representation of a closed material system, where there are in principle no discrete particles at all, can one interpret the Pauli exclusion principle, which appears to entail a system of discrete particles of matter? That is, in the holistic field theory the set of matter fields of the closed system $\{\psi^{(u)}(x)\}$ are all mapped in the same spacetime x , while in the discrete particle theory of quantum mechanics, the context of the Pauli exclusion principle entails the set of matter fields $\{\psi^{(u)}(x_u)\}$ that are each mapped in their own spacetime.

It will now be shown that, if any two components of a physically closed system, identified with the indices (u) and (v) , out of an n -component closed system, should have general electrodynamic coupling (discussed in detail in Chap. 5), and if they have:

- a repulsive interaction, i.e., $e^{(u)}e^{(v)} = +e^2$,
- the same inertial mass, i.e., $\lambda^{(u)} = \lambda^{(v)}$,
- the same state of motion everywhere, i.e.,

$$\psi^{(u)\dagger}(x)\gamma^0\gamma^\mu\psi^{(u)}(x) = \psi^{(v)\dagger}(x)\gamma^0\gamma^\mu\psi^{(v)}(x)$$

for all x ,

then the interaction weighting for the entire system of matter is identically equal to zero. In the particle theory in quantum mechanics, this corresponds to asserting that no two spin-1/2 particles out of an ensemble of particles can exist under the conditions above. This conclusion is physically equivalent to the assertion of the Pauli exclusion principle.

It will next be shown that when this exact feature of the theory is incorporated with the nonrelativistic approximation for the many-body interaction field amplitude, the latter function vanishes identically in the $4n$ -dimensional spacetime of the many-body particle description. The latter asymptotic form of the interaction field amplitude will be seen to be the totally antisymmetrized Slater determinant wave function for the many-body system of ordinary quantum mechanics. The latter is the form that underlies the description of the many-fermion system in terms of Fermi–Dirac statistics. Thus it will be seen that Fermi–Dirac statistics is only a linear approximation for the holistic, field theory of the many modes of a matter continuum, when it may be viewed as a system of discrete particles of matter.

To proceed, let Ψ_{uv} be the interaction weighting amplitude for a system with coupled spin-1/2 fields, excluding the u th and v th components of this system. The remaining amplitude will be denoted by ψ_{uv} , i.e., $\Psi = \Psi_{uv} + \psi_{uv}$. To clarify the notation further, Ψ_{uv} can depend on $\psi^{(u)}$ and $\psi^{(v)}$ separately, but only with regard to their separate coupling to the other matter fields of the closed system. The amplitude Ψ_{uv} then omits the contribution of the mutual coupling of (u) and (v) . If one then finds a field amplitude ψ_{uv} that vanishes identically under the three conditions posed above, one must conclude that the total interaction weighting amplitude $\Psi = \Psi_{uv}$. The physical consequences of the Pauli exclusion principle must then follow.

The preceding discussion assumed that for the entire closed system, the two-field amplitude ψ_{uv} exists. If this were not the case, then the total weighting amplitude Ψ could not be separated into the two additive parts, Ψ_{uv} and ψ_{uv} . However, it will be shown now that ψ_{uv} does indeed exist as a feature of the exact form of the coupled spinor matter field equations.

Note that if $\psi_{uv}^\dagger \psi_{uv}$ is an additive contribution to the weighting function for the closed system, then by itself it must satisfy the continuity equation

$$\partial_\mu (\psi_{uv}^\dagger \gamma^0 \gamma^\mu \psi_{uv}) = 0. \quad (6.11)$$

We must then seek a solution ψ_{uv} of the continuity equation (6.11) that depends on the two matter fields $\psi^{(u)}$ and $\psi^{(v)}$ of the coupled nonlinear matter field equations (6.3) and satisfies the continuity equation (6.11).

Multiplying the u th equation in the set (6.3) on the left by $\psi^{(v)\dagger} \gamma^0$ and the Hermitian conjugate of the v th equation on the right by $\gamma^0 \psi^{(u)}$, subtracting the two equations and repeating the operation with (u) and (v) interchanged, the sum of the resulting equations is:

$$\begin{aligned}
& \partial_\nu [\psi^{(v)\dagger} \gamma^0 \gamma^\nu \psi^{(u)} + \psi^{(u)\dagger} \gamma^0 \gamma^\nu \psi^{(v)}] \\
&= R_\nu [\psi^{(u)\dagger} \gamma^0 \gamma^\nu \psi^{(v)} - \psi^{(v)\dagger} \gamma^0 \gamma^\nu \psi^{(u)}] \\
&\quad - i g_M \left\{ \sum_\alpha \Phi_\alpha [\psi^{(v)\dagger} \gamma^0 \Gamma_\alpha \psi^{(u)} - \psi^{(u)\dagger} \gamma^0 \Gamma_\alpha \psi^{(v)}] - \text{h.c.} \right\} \\
&\quad + [\lambda^{(v)} - \lambda^{(u)}] [\psi^{(v)\dagger} \gamma^0 \psi^{(u)} - \psi^{(u)\dagger} \gamma^0 \psi^{(v)}] . \tag{6.12}
\end{aligned}$$

According to the development in Chap. 5, the symbols in (6.12) are as follows:

$$\begin{aligned}
R_\nu &= e^{(u)} e^{(v)} \int [\psi^{(v)\dagger} \gamma^0 \gamma_\nu \psi^{(v)} - \psi^{(u)\dagger} \gamma^0 \gamma_\nu \psi^{(u)}] G(x - x') d^4 x' \\
&\quad + [e^{(u)} - e^{(v)}] \sum_{w \neq u, v} e^{(w)} \int \psi^{(w)\dagger} \gamma^0 \gamma_\nu \psi^{(w)} G(x - x') d^4 x', \tag{6.13}
\end{aligned}$$

where $G(x - x')$ is the symmetric Green's function for d'Alembert's equation, shown in (5.53).

In the formulas above, $e^{(u)} e^{(v)} = +e^2$ for a repulsive interaction, $-e^2$ for an attractive interaction, and

$$\Phi_\alpha = (-1)^\alpha [e^{(u)} \phi_\alpha^{(v)\dagger} - e^{(v)} \phi_\alpha^{(u)\dagger}] + [e^{(u)} - e^{(v)}] \sum_{w \neq u, v} \phi_\alpha^{(w)\dagger} . \tag{6.14}$$

The right-hand side of (6.12) has three terms, each of them vanishing under special conditions. First, if the u th and v th matter fields have the same state of motion, i.e., when

$$\psi^{(u)\dagger} \gamma^0 \gamma^\mu \psi^{(u)}(x) = \psi^{(v)\dagger} \gamma^0 \gamma^\mu \psi^{(v)}(x) , \quad \forall x ,$$

it follows that

$$\psi^{(u)\dagger} \gamma^0 \Gamma_\alpha \psi^{(u)}(x) = \psi^{(v)\dagger} \gamma^0 \Gamma_\alpha \psi^{(v)}(x) , \quad \forall x . \tag{6.15}$$

The matrices Γ_α above are linear combinations of the Dirac matrices. They come from the spinor form of the electromagnetic field equations derived in Chap. 5:

$$\sigma^\mu \partial_\mu \phi_\alpha^{(u)} = \Upsilon_\alpha = e^{(u)} \psi^{(u)\dagger} \gamma^0 \Gamma_\alpha \psi^{(u)} . \tag{6.16}$$

According to (5.17), the correspondences are as follows:

$$\Upsilon_1(1) = \rho + j_3 \iff \psi^\dagger \gamma^0 (\gamma^0 + \gamma^3) \psi \equiv \psi^\dagger \gamma^0 \Gamma_1(1) \psi , \tag{6.17a}$$

$$\Upsilon_1(2) = j_1 + i j_2 \iff \psi^\dagger \gamma^0 (\gamma^1 + i \gamma^2) \psi \equiv \psi^\dagger \gamma^0 \Gamma_1(2) \psi , \tag{6.17b}$$

$$\Upsilon_2(1) = j_1 - i j_2 \iff \psi^\dagger \gamma^0 (\gamma^1 - i \gamma^2) \psi \equiv \psi^\dagger \gamma^0 \Gamma_2(1) \psi , \tag{6.17c}$$

$$\Upsilon_2(2) = \rho - j_3 \iff \psi^\dagger \gamma^0 (\gamma^0 - \gamma^3) \psi \equiv \psi^\dagger \gamma^0 \Gamma_2(2) \psi . \tag{6.17d}$$

It then follows that a form of the electromagnetic field equations (in special relativity) is

$$\sigma^\mu \partial_\mu [e^{(v)} \phi_\alpha^{(u)} - e^{(u)} \phi_\alpha^{(v)}] = e^{(u)} e^{(v)} [\psi^{(v)\dagger} \gamma^0 \Gamma_\alpha \psi^{(v)} - \psi^{(u)\dagger} \gamma^0 \Gamma_\alpha \psi^{(u)}] . \quad (6.18)$$

Recall that in the conceptual framework of this field theory, we may only accept the particular solutions of the electromagnetic equations. Thus substituting (6.15) into (6.18), the right-hand side of (6.18) vanishes, and the only solutions of (6.18) must correspond to the relation:

$$e^{(v)} \phi_\alpha^{(u)} = e^{(u)} \phi_\alpha^{(v)} .$$

The first part of (6.14) then vanishes. The second part of Φ_α vanishes as well under the conditions where $e^{(u)} e^{(w)} = e^{(v)} e^{(w)}$. If the mutual coupling between the u th and v th matter fields is repulsive, i.e., when $e^{(u)} e^{(v)} = +e^2$, then their separate coupling to the other fields of the closed system would have the same sign, i.e., $e^{(u)} e^{(w)} = e^{(v)} e^{(w)}$. In this case the total term $\Phi_\alpha = 0$. Further, if the inertial masses of the u th and v th matter fields are the same, $\lambda^{(u)} = \lambda^{(v)}$, the last part of the right-hand side of (6.12) vanishes as well.

Summarizing, if any two spinor matter fields out of an n -component closed system of spin-1/2 matter fields of a closed system:

- are in the same state of motion,
- have the same inertial mass,
- have a mutually repulsive electrodynamic interaction,

then $R_\nu = 0$, $\Phi_\alpha = 0$, $\lambda^{(u)} - \lambda^{(v)} = 0$, the entire right-hand side of (6.12) vanishes. It then follows that under these three conditions,

$$\partial_\nu [\psi^{(v)\dagger} \gamma^0 \gamma^\nu \psi^{(u)} + \psi^{(u)\dagger} \gamma^0 \gamma^\nu \psi^{(v)}] = 0 . \quad (6.19)$$

Combining (6.19) with the continuity equation for the separate matter fields,

$$\partial_\mu [\psi^{(u)\dagger} \gamma^0 \gamma^\mu \psi^{(u)}] = \partial_\mu [\psi^{(v)\dagger} \gamma^0 \gamma^\mu \psi^{(v)}] = 0 ,$$

and the general requirement of the closed system that it must be invariant under the interchange,

$$\psi^{(u)}(x) \longleftrightarrow \psi^{(v)}(x) ,$$

it follows that

$$\partial_\mu \left[(\psi^{(u)} \pm \psi^{(v)})^\dagger \gamma^0 \gamma^\mu (\psi^{(u)} \pm \psi^{(v)}) \right] = 0 .$$

Thus, the possible solutions ψ_{uv} of (6.11) are

$$\psi_{uv}(\pm) = \psi^{(u)} \pm \psi^{(v)} \quad \text{everywhere} . \quad (6.20)$$

The only remaining ambiguity, then, is the choice of the plus or the minus sign in this solution. Since the u th and the v th interacting matter fields correspond, in this case, to the same states of motion (6.11), and they have the same separate interaction with the rest of the closed system, it follows, because they have the same boundary conditions, that when the special case (6.20) is valid, each of the matter fields $\psi^{(u)}$ and $\psi^{(v)}$ solve identical differential equations, with identical boundary conditions. In this case, they must map into each other. That is to say,

$$\psi^{(u)}(x) = \psi^{(v)}(x) \quad \text{everywhere} . \quad (6.21)$$

It is important that there are not different valued constants in front of each of these functions. This is because they are subject to the same boundary conditions and solve the same nonlinear differential equations.

If the plus sign in (6.20) is correct, then we must solve the matter field equation for the solution $\psi_{uv} = 2\psi^{(u)}$. In this case, $\psi^{(u)}$ would be a solution of the matter field equation with the interaction term

$$I_u(\psi^{(1)}, \dots, \psi^{(v=u)}, \dots, \psi^{(n)})\psi^{(u)} . \quad (6.22)$$

But this would represent the matter field $\psi^{(u)}$ acting on itself, as well as with the other matter of the closed system. The acceptance of this solution is then unacceptable for the uv interaction field amplitude, $\psi_{uv}(+)$, because of the premise of this theory that rules out self-energy.

To be logically consistent, it then follows that the only solution in (6.20) is

$$\psi_{uv}(-) = \psi^{(u)} - \psi^{(v)} = 0 , \quad \forall x . \quad (6.23)$$

The zero on the right then follows from the three special conditions that led to the equality in (6.21) – the same conditions that lead to the statement of the Pauli exclusion principle.

Summing up, we have seen that under the three conditions imposed on a closed system of spinor field constituents:

- two of them, labeled (u) and (v) , are in the same state of motion everywhere,
- they have a mutually repulsive interaction (corresponding in the particle theory to each matter field having the same electric charge with the same polarity),
- they have equal mass,

the total interaction field amplitude Ψ for the closed, holistic system of matter reduces to Ψ_{uv} , which excludes the mutual coupling of the u th and the v th matter field constituents.

What this means is that, within the context of this field theory in relativity theory, no possible measurement could relate to the mutual coupling of the u th and the v th constituents. This conclusion is physically equivalent to the statement of the Pauli exclusion principle, though coming from a theory totally different from the quantum theory. Indeed, features of this theory that were responsible for the result are incompatible with the basis of the quantum theory! Examples are linearity versus nonlinearity, particularity versus holism, fundamental probability versus a deterministic theory with no fundamental probability in its basis.

6.6 Sufficiency of the Three Conditions for Proof of the Pauli Principle

What has been shown thus far is that the three conditions specified above are necessary for the proof of the Pauli principle. We will now show that these conditions are also sufficient for the proof.

The proof of sufficiency starts with the assumption that if two of the constituent spinor fields of the n fields of a closed system, $\psi^{(u)}(x)$ and $\psi^{(v)}(x)$, for all x , solve the same differential equation (6.3), then the three special conditions automatically follow. According to this field theory for a closed system, the field solutions are mapped in the same spacetime, thus there is no ‘entanglement’ involved. It then follows that they solve the same differential equation with the same boundary conditions, imposed by the remainder of the matter of the closed system. Thus,

$$\psi_{uv}(-) = [\psi^{(u)} - \psi^{(v)}](x) = 0, \quad \forall x.$$

It then follows that the derivatives of $\psi_{uv}(-)$ vanish everywhere. Thus, the continuity equation follows:

$$\partial_\mu [\psi_{uv}(-)^\dagger \gamma^0 \gamma^\mu \psi_{uv}(-)] = 0 . \quad (6.24)$$

Combining (6.24) with the separate continuity equations in $\psi^{(u)}$ and $\psi^{(v)}$, it follows that (6.11) must be true. With this result, the left-hand side of (6.12) automatically vanishes. Because of the independence of the terms on the right-hand side of this equation, we have

$$R_\nu = 0 , \quad \Phi_\alpha = 0 , \quad \lambda^{(u)} = \lambda^{(v)} .$$

Due to the independence of the separate terms in R_ν , it follows from its vanishing that the matter fields $\psi^{(u)}$ and $\psi^{(v)}$ correspond to the same state of motion, i.e., it follows that (6.11) must be true and that the mutual interaction must be repulsive, i.e., that $e^{(u)} = e^{(v)}$.

With the latter conclusion and the fact that $\Phi_\alpha = 0$, it follows that

$$e^{(u)} \phi_\alpha^{(v)} = e^{(v)} \phi_\alpha^{(u)} .$$

Thus, the u th and the v th matter fields are equivalent, since the electromagnetic spinor fields ϕ_α are the particular solutions of equations (6.16), which may then be expressed in the form of (6.18). That is, with the equivalence of the spinor matter fields, the right-hand side of (6.16) vanishes, and so the particular solutions of these differential equations vanish.

Summing up, starting with the equivalence of the spinor matter solutions $\psi^{(u)} = \psi^{(v)}$ for all x , if the contribution to the total interaction field amplitude Ψ from the uv coupled, nonlinear matter field components should vanish, it follows that the crucial three conditions are derived:

- the mutual uv interaction is repulsive,
- their inertial masses are equal,
- they are in the same state of motion.

Thus it has been demonstrated that the foregoing three conditions are necessary and sufficient for the physical implications of the Pauli exclusion principle, thereby completing the proof of this principle in the context of the field theory explored. It should be noted that this proof was based on features of relativity theory that are in principle excluded from the quantum theory – nonlinearity, holism (i.e., non-particularity) and determinism.

6.7 Fermi–Dirac Statistics from the Nonrelativistic Approximation for Ψ

Returning to the nonrelativistic approximation (6.10) for the interaction field amplitude Ψ , the time coordinate in this formulation becomes a common parameter in all factors. Any two factors, when corresponding to equivalent states of motion, contribute the product

$$\psi_{uv}^{(0)} = \psi^{(u)}(\mathbf{r}_u)\psi^{(v)}(\mathbf{r}_v = \mathbf{r}_u). \quad (6.25)$$

If (u) and (v) are interaction fields with the same mass and a mutually repulsive force, then the exact result derived in the preceding section implies that (6.25) is only an approximation for $\psi_{uv}(-) = 0$. That is to say, (6.25) is a non-zero amplitude that is an approximation for zero.

Finally, since the product (6.25) is a factor that multiplies the product of all other matter field solutions, it follows that the actual vanishing of the exact form of this product causes the interaction field amplitude to vanish identically under the conditions that cause the product $\psi_{uv}(-)$ to vanish.

To incorporate this conclusion into the asymptotic limit for the system that appears as ‘equivalent spin-1/2 particles’, we choose the phase in (6.10) as follows:

$$\alpha_P = \pm\pi P \longrightarrow \exp(i\alpha_P) = (-1)^P. \quad (6.26)$$

In this case, the equivalent way to express the nonrelativistic approximation for the interaction field amplitude Ψ is in terms of the totally antisymmetrized Slater determinant form for the many-fermion system:

$$\begin{aligned} \Psi &= \left(\frac{1}{n!}\right)^{1/2} \sum_P (-1)^P \prod_u \psi^{(u)}(\mathbf{r}_u) \\ &= \left(\frac{1}{n!}\right)^{1/2} \begin{vmatrix} \psi^{(1)}(\mathbf{r}_1) & \dots & \psi^{(1)}(\mathbf{r}_n) \\ \vdots & & \vdots \\ \psi^{(n)}(\mathbf{r}_1) & \dots & \psi^{(n)}(\mathbf{r}_n) \end{vmatrix}. \end{aligned} \quad (6.27)$$

This is the totally antisymmetrized wave function for the many-fermion system of particles in the quantum theory. It is the basis of Fermi–Dirac statistics for an ensemble of non-distinguishable spin-1/2 particles. It is the type of statistics that has been empirically successful in various predictions, such as Mendeleyev’s Periodic Table of the

atomic species, the properties of metals, the correct scattering amplitudes of charged spin-1/2 particle scattering, etc.

It should be noted at this juncture that, although the interaction field amplitude (6.27) is identical with the many-body wave function of quantum mechanics for a system of non-interacting fermions, the interpretation here is not in terms of an ensemble of discrete particles. The interaction field amplitude Ψ does not generally have the Slater determinant structure – it only appeared this way in this theory as a linear approximation. The derived Pauli exclusion principle, on the other hand, deals with the system of matter implied by a particular state of mutual interaction, independent of approximation.

Thus, Fermi–Dirac statistics appears in this holistic, continuum field theory of matter as a linear approximation, when the constituents of the matter of the actual closed system appear to be uncoupled. This approximation is valid only when the amount of energy–momentum transfer between interacting constituents of the closed system is non-relativistic in magnitude – it occurs when v/c is close to zero, where v is the relative speed between the interacting matter.

In Chap. 7, we will apply the formulas of the holistic field theory in relativity, derived thus far, to problems of atomic and elementary particle physics. In Chap. 8, we will apply this holistic field theory to problems in the astrophysical domain.

6.8 Bound Particle–Antiparticle Pairs. Ground State

The coupled equations for the particle–antiparticle pair in the Dirac bispinor notation are

$$[\gamma^\mu \partial_\mu - I(e^+) + \lambda] \psi(e^-) = 0, \quad (6.28a)$$

$$[\gamma^\mu \partial_\mu - I(e^-) + \lambda] \psi(e^+) = 0, \quad (6.28b)$$

where, as before, γ^μ are the Dirac matrices

$$\gamma^0 = \begin{pmatrix} \boldsymbol{\sigma}^0 & 0 \\ 0 & -\boldsymbol{\sigma}^0 \end{pmatrix}, \quad \gamma^k = -i \begin{pmatrix} 0 & \boldsymbol{\sigma}^k \\ -\boldsymbol{\sigma}^0 & 0 \end{pmatrix}. \quad (6.29)$$

The solution $\psi(e^+) = C\psi(e^-)$ is the charge conjugate of $\psi(e^-)$, reversing the sign of the electric charge, $e \rightarrow -e$, and $C = \gamma^2 K_0$ is the charge conjugate operator, where K_0 is the operation of taking the complex conjugate of the function to its right.

The total electromagnetic interaction functional is $I = I_1 + I_2$, where

$$I_1(e^\pm) = -e^2 \gamma^\nu \int \psi(e^\pm)^\dagger \gamma^0 \gamma_\nu \psi(e^\pm) G(x - x') d^4 x', \quad (6.30a)$$

$$I_2(e^\pm) = -ig_M(e^- \text{ or } e^+) \sum_\alpha (-1)^\alpha \left(\phi_\alpha^\dagger \cdot \Gamma_\alpha - \gamma_0 \Gamma_\alpha^\dagger \gamma_0 \cdot \phi_\alpha \right), \quad (6.30b)$$

where $G(x - x')$ is the symmetric Green's function defined in (5.53), the matrices Γ_α are defined in (6.17a-d) and the notation above is

$$\phi_\alpha^\dagger \cdot \Gamma_\alpha \equiv \phi_\alpha^*(1) \Gamma_\alpha(1) + \phi_\alpha^*(2) \Gamma_\alpha(2),$$

where the asterisk here denotes the complex conjugate.

It is rigorously demonstrated in a previous publication ([136], Chap. 7), that an exact bound state solution, when the particle and antiparticle states are in the same state of motion, is

$$\psi(e^+) = -\psi(e^-) = \begin{pmatrix} \exp(-i\lambda t) \\ 0 \\ 0 \\ \exp(i\lambda t) \end{pmatrix}, \quad (6.31)$$

in the reference frame of the pair (i.e., its proper Lorentz frame). In any other Lorentz frame, the argument of the exponential in (6.31) would generalize to the scalar invariant $ik^\mu x_\mu = i(\mathbf{k} \cdot \mathbf{r} - \lambda t)$, where the wave vector $\mathbf{k}$ takes account of the motion of the pair relative to the observer.

With this matter field solution for the particle and antiparticle, the electromagnetic spinor equations have the form

$$\begin{aligned} \sigma^\mu \partial_\mu \phi_1(e^+) &= e\psi^\dagger \gamma^0 \Gamma_1 \psi = -8\pi i e \begin{pmatrix} 1 \\ -\exp(2i\lambda t) \end{pmatrix}, \\ \sigma^\mu \partial_\mu \phi_2(e^+) &= e\psi^\dagger \gamma^0 \Gamma_2 \psi = 8\pi i e \begin{pmatrix} -\exp(-2i\lambda t) \\ 1 \end{pmatrix}. \end{aligned} \quad (6.32)$$

With this state of the matter fields of the particle and antiparticle, and the electromagnetic spinor equations with this matter field solution shown in (6.32), the two independent electromagnetic spinor solutions of (6.32) have the form [136]

$$\phi_1(x) = \frac{4\pi e}{\lambda} \begin{pmatrix} 1 \\ \exp(2i\lambda t) \end{pmatrix}, \quad (6.33a)$$

$$\phi_2(x) = \frac{4\pi e}{\lambda} \begin{pmatrix} \exp(-2i\lambda t) \\ 1 \end{pmatrix}. \quad (6.33b)$$

With these solutions, it can be shown [136] that

$$(I_1 + I_2)\psi(e) = I\psi(e) = 0.$$

Note at this stage that while the operation of the interaction operator I on the bound state of the pair $\psi(e)$ is zero, I itself is not zero. This is a consequence of the nonlinear features of the coupled field equations for the pair, under the special circumstances specified in this case.

In summary, the spinor function (6.31) is an exact solution of the coupled equations for the pair (6.28a) and (6.28b), when the particle and its bound antiparticle are in the same state of motion.

6.9 Energy and Momentum of the Bound Pair in its Ground State

It follows from Noether's theorem that the covariance of a field theory with respect to continuous transformations in time and space imply, respectively, the conservation of energy P_0 and the components of linear momentum P_k , where

$$P_0 = \int \sum_i \left[\frac{\partial L}{\partial(\partial_0 \Lambda_\zeta^{(i)})} \partial_0 \Lambda_\zeta^{(i)} - L \right] d\mathbf{r}, \quad (6.34a)$$

$$P_k = \int \sum_i \frac{\partial L}{\partial(\partial_0 \Lambda_\zeta^{(i)})} \partial_k \Lambda_\zeta^{(i)} d\mathbf{r} \quad (k = 1, 2, 3), \quad (6.34b)$$

and the index i above is summed over all fields (as the variational parameters). The index ζ denotes the components of the respective fields. The total Lagrangian density is $L = L_D + L_M$.

In the special case of the bound electron-positron pair (or the proton-antiproton pair), there are 12 such fields as variational parameters:

$$\left\{ \psi(e^-), \psi(e^-)^\dagger \gamma^0, \psi(e^+), \psi(e^+)^\dagger \gamma^0, \phi_\alpha(e^-), \phi_\alpha(e^-)^\dagger, \phi_\alpha(e^+), \phi_\alpha(e^+)^\dagger \right\}, \quad (6.35)$$

where $\alpha = 1, 2$ are the two separate spinor variables.

The first part of the Lagrangian density above is

$$L_D = \frac{hc}{2\pi} \sum_u \left\{ \psi^{(u)\dagger} \gamma^0 [\gamma^\mu \partial_\mu + I_1(u) + \lambda(u)] \psi^{(u)} + \text{h.c.} \right\}, \quad (6.36)$$

where $u = 1, 2$ denotes the particle and antiparticle variables. Since the particle and antiparticle have the same inertial mass, $\lambda(1) = \lambda(2)$. The

constant $hc/2\pi$ is inserted here to demonstrate that this is the place in the formalism where Planck's constant h appears, most fundamentally, rather than in commutation relations between the position and momentum operators, as in the quantum theory. However, we do not see these constants explicitly, since we use units with $\hbar/2\pi = 1 = c$ throughout.

The second part of the Lagrangian density has the form

$$L_M = i g_M \sum_{u \neq v} \sum_{\alpha} (-1)^{\alpha} \phi_{\alpha}^{(u)\dagger} [\sigma^{\mu} \partial_{\mu} \phi_{\alpha}^{(v)} - 2e^{(v)} \psi^{(v)\dagger} \gamma^0 \Gamma_{\alpha} \psi^{(v)}] + \text{h.c.} . \quad (6.37)$$

With the fields indicated above and the solution (6.31) for the matter field of the pair, it is readily shown that the expression for the energy and the linear momentum in (6.34a) and (6.34b) yield null values for this bound state:

$$P_0 = P_1 = P_2 = P_3 = 0 .$$

This is the true ground state of the pair. In units of energy, this bound state is $2mc^2$ below their energy when the particle and antiparticle would be (asymptotically) free of each other.

Because all four components of the energy-momentum four-vector P_{μ} are null in the Lorentz frame described above, it must be a null vector in all other Lorentz frames of reference. That is, the ground state of the pair so-derived is Lorentz invariant.

6.10 Dynamical Properties of the Pair in its Ground State

It also follows from the general form of the Lagrangian density above, due to its invariance with respect to rotations in the spacetime coordinate system, that Noether's theorem predicts the conservation of the components of the angular momentum of the pair. With the field solutions for the ground state of this pair, derived above, it turns out that the angular momentum of the pair in this ground level is a singlet 1S_0 state.

Noting that the source terms on the right-hand side of (6.32) are

$$\Upsilon_1 = 4\pi i \begin{pmatrix} -\rho + j_3 \\ j_1 + i j_2 \end{pmatrix} , \quad \Upsilon_2 = 4\pi i \begin{pmatrix} -(j_1 - i j_2) \\ \rho + j_3 \end{pmatrix} ,$$

it follows that the ground state of the pair may be expressed as follows with the usual variables:

$$\alpha = 1 : \quad \rho = 2e^{\pm} , \quad j_3 = 0 , \quad j_1 + ij_2 = 2e^{\pm} \exp(2i\lambda t) , \quad (6.38a)$$

$$\alpha = 2 : \quad \rho = 2e^{\pm} , \quad j_3 = 0 , \quad j_1 - ij_2 = 2e^{\pm} \exp(-2i\lambda t) . \quad (6.38b)$$

These are two oppositely polarized currents that are mutually transverse with respect to the x_3 direction. Thus, the coincident response of two equidistant counters, placed on opposite sides of the pair in its ground state, at the time t , would be to two spatially transverse currents that are 90° out of phase. These responses of the detectors would be in terms of direct coupling with two distinguishable currents. The latter assertion will now be demonstrated.

To derive the currents $j_{\pm} = j_1 \pm ij_2$ associated with the pair at the origin of our coordinate system, on the two counters equidistant at $\pm r$, it will be necessary to calculate the electric field intensities $E_{\pm}$ at the locations of the counters. This is because it is the electric field intensity that determined the motion of a test charge in the detecting apparatus.

In the Lorentz frame of the detecting device, the vector potential that corresponds to the oppositely polarized current densities $j_{\pm}$ follows from the particular solutions of d'Alembert's equation:

$$\square A_{\pm}(\mathbf{r}', t') = 4\pi j_{\pm}(\mathbf{r}', t') , \quad (6.39)$$

where $(\mathbf{r}', t')$ are the coordinates of the test charge of the apparatus, while the coordinates of the emitting source, the pair itself, are $(\mathbf{r} = 0, t)$. The solution of (6.39) is

$$A_{\pm}(\mathbf{r}', t') = \int j_{\pm}(t) G(x - x') d^4x , \quad (6.40)$$

where $G(x - x')$ is the symmetric Green's function (5.53). Although this corresponds to a symmetry between the retarded and the advanced potentials, it should be noted that because of the symmetry of the experimental setup in this particular problem, the result to be derived regarding the responses of two sets of apparatus, on opposite sides of the pair, along a common axis and equidistant from the plane of polarization of the pair currents $j_{\pm}$, is insensitive to the appearance or lack of appearance of the advanced term in the Green's function.

Since $j_3 = 0$, substitution of the Green's function (5.53) into (6.40) (with $r = 0$) yields the solution

$$A_3(r', t') = \int j_3(t) G(x - x') d^4x = 0 ,$$

$$\mathbf{A}_\pm(r', t') = \frac{2e^\pm}{2r'} \left\{ \exp [\pm 2i\lambda(t' + r')] + \exp [\pm 2i\lambda(t' - r')] \right\} \mathbf{e}_\pm , \quad (6.41)$$

where $\mathbf{e}_\pm = \mathbf{e}_1 \pm i\mathbf{e}_2$ and $\mathbf{e}_j$ is the unit vector in the j th direction.

It then follows from (6.41) that the electric field intensities $\mathbf{E}_\pm(r', t')$ at the counters, due to the effects of the polarized currents $j_\pm$ of the particle–antiparticle sources, at the origin, have the form

$$\mathbf{E}_\pm(r', t') = -\frac{\partial \mathbf{A}_\pm}{\partial t'} \quad (6.42)$$

$$= -\mathbf{e}_\pm(\pm) \frac{2i\lambda e^\pm}{r'} \left\{ \exp [\pm 2i\lambda(t' + r')] + \exp [\pm 2i\lambda(t' - r')] \right\} .$$

Thus we see that $\mathbf{E}_\pm$ describes a wave motion of an oscillating charge, whose angular frequency is $\omega = 2\lambda = 4\pi mc^2/h$ with a wave vector magnitude $k = \omega/c$. It then follows that when the phase of the current density has some fixed value at $t = 0$, say it is equal to zero then, the phase of the electric field oscillation at $r' \neq 0$ that is produced by this current would not become zero until the later time $t' = r'/c$ in the retarded solution, and at the time $t' = -r'/c$ in the advanced solution. Thus, the magnitude of the time taken for the propagation of the electromagnetic interaction between the pair and each of the detectors (each a distance r' from the pair) is r'/c seconds.

A salient point is that for the solution $\mathbf{E}_+$ due to the current source j_+ , the sign of the propagation vector is positive in the retarded term and the sign of this vector is negative in the advanced term. The oppositely polarized current density j_- gives rise to the electric field intensity $\mathbf{E}_-$ with the same functional form as $\mathbf{E}_+$, except that the propagation vector in this case is negative in the retarded term and positive in the advanced term.

We see, then, that the oppositely polarized currents $j_\pm$, at the common spatial location $\mathbf{r} = 0$, give rise to oppositely polarized electric field vectors that propagate in opposite directions – such that when each counter is an equal distance r' on each side of the pair, along a common axis, they will simultaneously detect oppositely polarized currents at the time $t' = r'/c$, each absorbing the energy mc^2 .

This derived result agrees with the experimental facts and with the law of energy conservation. [See C.S. Wu and I. Shaknov, *The Angular Correlation of Scattered Annihilation Radiation*, Phys. Rev. **77**, 136 (1950).] It also agrees with the prediction of the model proposed

in the quantum theory asserting that two photons, correlated with a 90° phase difference, are simultaneously created when the particle–antiparticle pair annihilates. However, the present theory does not require the introduction of photons nor that matter should actually annihilate into a vacuum. An important difference of the two theories, beside the fact that here there are no photons involved – they are superfluous – is that this is a deterministic field theory of matter, in all domains (micro- as well as macro-physics). This is in contrast with the intrinsically nondeterministic character of the quantum theory, where the laws of matter are laws of probabilities.

6.11 Pair Creation, the Physical Vacuum, and Blackbody Radiation

We have seen that the derived ground state of the bound particle–antiparticle pair is at zero energy, relative to the energy of this matter when the particle and antiparticle would be (asymptotically) free of each other, at $2\lambda \equiv 2mc^2$, where m is the inertial mass of each of the particles. If this quantity of energy (approximately 1 Mev for the electron–positron pair or 2 Gev for the proton–antiproton pair) were delivered to the pair in its ground state, it would dissociate. Each of the matter components would then be independently available to interact with its surroundings. Thus one should see tracks in a cloud chamber, interpreted as pair creation. But in this interpretation of these data, particles are not created out of a vacuum! The particle and its antiparticle twin were there all the time. It is only that they were invisible to the detecting apparatus when they were bound in their ground state. When this bound pair is dissociated (deterministically), the particle and antiparticle then make themselves visible by interacting with the matter of the apparatus to create tracks, which are then identified with the ‘particle’ and ‘antiparticle’.

A second interesting implication of the derived ground state of the pair is the assertion that there is no reason why any domain would not be populated with a very high density of such pairs, as a countable set making up a dense gas, that we call the physical vacuum. (The magnitude of this density will be determined in the next chapter.) But this is not the type of physical vacuum asserted by quantum field theory. The latter is made up of a non-countable set of particles, antiparticles and radiation that continually transmute into each other, non-deterministically. In the field theory of this approach in relativity theory, the physical vacuum does not contain any radiation:

it is made up of a countable set of particle–antiparticle pairs in their ground states of null energy–momentum.

It should be noted here that since all of the constituent pairs of this physical vacuum have null energy and momentum for their ground states, the ground state of the physical vacuum itself must be at zero energy. This result is in contrast with that of quantum field theory, where there are an infinite number of negative energy values, below that of positive value energy levels. The reason that the negative energy levels do not appear in the classical field theory is that the definition of energy here is in terms of covariance under continuous transformations from one energy value to the next, yielding the expression (6.34a) according to Noether’s theorem. One cannot continuously transform from a positive energy value to a negative one in the classical field theory, as one may do in the quantum theory, where one has only discrete jumps in energy values. Thus, the minimum energy value must be at zero in the covariant field theory in relativity, while in the quantum field theory the minimum energy is at minus infinity – unless, as it is done, one postulates (without derivation) a zero energy for the vacuum. The derivation of the zero energy of the constituent pairs of the physical vacuum in the non-quantum relativity theory then proves, from first principles, that the minimum energy of the physical vacuum in this theory must be zero.

With this model of the physical vacuum, we may consider the experiment on blackbody radiation in terms of the coupling of the charged matter of an apparatus to the ensemble of constituent pairs of the physical vacuum through the windows of a cavity that is maintained at a constant temperature. From this experimental setup, it is possible to derive the formula for the spectral distribution of blackbody radiation, assuming with Planck that the energy in each of the pair modes in the physical vacuum contained in the cavity is linearly proportional to its frequency. In this way, the correct Planck formula was derived from this field theory (see [136], Chap. 7, and [66]).

Thus we have explained blackbody radiation without the need for the photon model of electromagnetic radiation. It followed from a statistical analysis of the direct coupling of charged matter of a detecting apparatus to the countable set of particle–antiparticle pairs in their ground states of null energy–momentum, that make up the physical vacuum inside of a cavity maintained at a fixed temperature. This result reinforces the contention of this theory that one may dispose of the photon as an elementary particle. It only appears as a coupling that

transmits the electromagnetic interaction between charged matter, but not as a ‘thing in itself’.

There are several important roles of the physical vacuum of particle pairs in their ground states of null energy–momentum.

- It is responsible for what is interpreted as spontaneous decay of excited atomic or molecular systems. In this deterministic theory, there are only ‘induced’ interactions. The coupling of the excited atomic or molecular systems to the background of pairs in their ground state – the physical vacuum – is then the cause of these events.
- A second important consequence of this physical vacuum is the observation of blackbody radiation, as we have discussed above. It is not due to a gas of photons confined in a cavity, in this theory, in which the photon does not exist as a separate entity. It is rather due to the coupling of the measuring apparatuses to a dense gas of particle–antiparticle pairs – the physical vacuum – in a cavity maintained at a fixed temperature, in thermodynamic equilibrium with the walls of the cavity.
- A crucial role of the physical vacuum of pairs in their ground states of null energy and momentum is that it is the cause of the inertial mass of any observed matter, due to its coupling with this matter background. The latter is a consequence of the Mach principle, where the primary matter that gives rise to the inertial mass of observed matter is not the distant stars of the universe, but rather the nearby particle–antiparticle pairs of the physical vacuum. It will be seen (Chap. 7) that the inertial mass of elementary matter appears here in terms of mass doublets, and that the heavy sister of the electron of this doublet is the muon. The correct mass and lifetime of the muon will then be calculated.
- A further possible role of the physical vacuum of pairs is a contribution to the dark matter of the universe, which astrophysicists see today as essential to explain astrophysical phenomena, e.g., the rotations of galaxies.

7. Atomic and Elementary Particle Physics

The field theory of matter developed in this book reveals that there are only correlated spin-1/2 matter field components that comprise any complex system of matter. The ‘photon’ and the ‘neutrino’ are not elementary particle fields. Rather, they are virtual fields that couple matter to effect their mutual interaction. The bound electron–proton coupled with the long range electromagnetic force (with the virtual photon) is called hydrogen. The bound electron–proton, coupled with the short range electromagnetic force (with the virtual neutrino) is called neutron. These two types of electromagnetic force follow from the factorization of the electromagnetic field theory into a scalar and a pseudoscalar part, which may be identified, respectively, with the standard and the weak electromagnetic interactions. Such a factorization follows from the rejection of reflection symmetry in the electromagnetic (and all other) interactions in nature, in accordance with the underlying symmetry groups of the theory of relativity – the Poincaré group for special relativity and the Einstein group for general relativity.

The extra interaction in the case of hydrogen predicts a lifting of the accidental degeneracy in the states of hydrogen. Thus this extra contribution to the total interaction in the Hamiltonian for hydrogen predicts the Lamb shift. Its calculation yields results that agree (within experimental error) with the experimental results for the states with principal quantum numbers 2, 3 and 4.

A phenomenological discussion follows for properties of the neutron, pion and kaon, as composites of the electron, proton, their antiparticle fields and the virtual pair field corresponding to the ‘neutrino’. The composite neutron is shown to lead to the correct order of magnitude of its binding energy, lifetime and magnetic moment.

Next, it is shown that in general relativity there is a prediction of mass doublets. The heavy sister of the electron is the muon, created by the excitation of a background pair that is close to the observed electron, giving rise to a change in the geometrical field that affects the mass of the observed electron. The mass of the muon is then found to

be of the order of $(3/2\gamma)m_e \approx 206m_e$, where m_e is the electron mass and $\gamma = 1/137$ is the fine structure constant. The lifetime of the excited pair in the vicinity of the heavy electron is found, with perturbation theory, to be the order of 10^{-6} seconds. Both of these predictions for the heavy electron are compatible with the empirical parameters for the muon. Thus general relativity explains the existence of the muon in nature. There is a further prediction that the 'heavy proton' has a mass of the order of 193 Gev.

Generally, any number of pairs in the vicinity of an observed electron can be excited. Thus there is a potentially infinite spectrum of leptons: e , μ , τ , τ' , and so on.

Next, we analyze the charged and neutral pions. This is a particle that has been evoked to describe the coupling of neutrons and protons in the nuclear medium, effecting the strong (nuclear) force. Viewing the pion as a composite in the context of this theory, it is shown that the ratio of the masses of the neutral to charged pions and the ratio of their lifetimes is in agreement with the empirical facts. It is then seen how the model of the kaon as a composite leads to the correct CP violation in a decay channel of the long-lived neutral kaon decay, compared with the CP conserving decay channel.

Finally, it is argued how charge quantization of elementary matter is predicted in general relativity. It follows from the irreducible representations of the Einstein symmetry group of general relativity, expressed as a first order approximation beyond the irreducible representations of the Poincaré group in special relativity.

7.1 Introduction

In this chapter, we will apply the field theory based on general relativity theory and its special relativity limit to the atomic and elementary particle domains of matter. The first example will be the bound electron–proton system, coupled with the (long-range) electromagnetic interaction identified with the virtual photon. That is, the transfer of the virtual photon between them affects the binding force between the electron and the proton. This bound state of the atom, electron–proton, is called hydrogen.

The next phase of this analysis will treat the bound electron–proton system, coupled by the (short-range) electromagnetic interaction, identified with the virtual neutrino. This bound state of the electron–proton system, affected by the transfer of the virtual neutrino between

them, is called the neutron. The long-range electromagnetic interaction is governed by the scalar electromagnetic coupling discussed in the preceding chapter. We saw there that the generalized electromagnetic interaction has a scalar and a pseudoscalar part, because of the lack of reflection symmetry in the generalized theory. Thus the pseudoscalar part of the generalized electromagnetic interaction plays the role of the binding of the electron and proton to form the composite neutron, while the scalar part plays the role of binding the electron and proton to form hydrogen.

The successive addition of protons and neutrons (bound by the short-range nuclear force) to form the nuclei of atoms and the electrons whose number is equal to the number of protons to neutralize the charge of the created atoms then yields the atomic species beyond hydrogen, as they appear in the Periodic Table. The Pauli exclusion principle determines their classification in this table, as it has been derived in Chap. 6. At the end of this chapter, we will see how the group structure of the theory of general relativity yields the empirical fact of integral electric charge of the units of elementary matter.

Generally, the irreducible group of the theory of general relativity implies that the fundamental fields that represent elementary matter are the basis functions of its quaternion representations – they are the spinor variables. Thus, in fundamental terms, there are no elementary bosons – the photon is a virtual coupling between real matter fields of spin-1/2, and the pion, that binds the nuclear particles to affect the nuclear force field, as well as the kaon, are composites of spin-1/2 fields. Physical features of these composites will be derived in a phenomenological fashion.

In the next section, we will develop the theory of hydrogen per se, including a derivation of the Lamb shift in its spectrum. After this, we will develop features of the composites, neutron, pion and kaon, such as the ratios of the masses of charged to neutral pions and their lifetimes, and the reason for CP violation in neutral kaon decay. We will see how this theory correctly predicts quantitative features of these matter fields. It will then be demonstrated how, in general relativity, this theory predicts the occurrence of mass doublets, and associates the heavy sister of the electron with the muon, as well as predicting a heavy proton, whose mass is of the order of 193 GeV. Finally, there will be a derivation of the integral charge (referred to as charge quantization) of interacting electric matter, from the irreducible representations of the symmetry group of the theory of general relativity.

7.2 Hydrogen

7.2.1 Linearization of the Hydrogen Field Equation

Our determination (in Chap. 6) of the exact solution (6.31) of the ground state of the bound e^-e^+ was facilitated by the symmetry in the nonlinear field equations for this system – specifically that their inertial masses were the same and that they were in the same state of motion.

We will now consider the case where the masses of the bound particles are very different – the electron with mass m and the proton with a mass of the order of $m_p = 2000m$. In this case, the respective states of motion of the bound electron and proton are quite different – that of the proton may be assumed to be stationary relative to the electron, while the electron has orbital motion relative to the proton.

In this problem, there is no exact solution for the hydrogen atom to be displayed. The solution below follows from an approximation whereby the coupled electron–proton equations are linearized. It follows from considering the ratio m/m_p to be close to zero, so that we may consider that the transfer of energy and momentum to the proton field from the electron field is negligible. Thus we will assume that the proton matter field – the source of binding the electron field – is stationary relative to the electron.

Starting with the matter field equation (6.3), we now consider the coupled electron and proton equations:

$$[\gamma^\mu \partial_\mu - eI(p) + \lambda]\psi(e) = 0, \quad (7.1a)$$

$$[\gamma^\mu \partial_\mu + eI(e) + \Lambda]\psi(p) = 0, \quad (7.1b)$$

where the mass parameters are $\lambda = 2\pi mc/h$ ($= m$, the mass of the electron, with the units chosen $h/2\pi = 1 = c$) and $\Lambda = 2\pi Mc/h$ ($= M$, the mass of the proton, with the units chosen). λ and Λ are explicitly the inverses of the Compton wavelengths of the electron and the proton, respectively.

Consider now that the proton solution has the form of a stationary field, which in its proper frame is

$$\psi^{(p)} = \exp(-i\Lambda t) f(\mathbf{r})s, \quad (7.2)$$

where s is a constant four-component spinor whose only restriction is that it is normalized, i.e., $s^\dagger s = 1$.

Dividing (7.1b) by Λ , we have the equivalent proton equation

$$\left[\frac{1}{\Lambda} \gamma^0 \partial_0 - \frac{1}{\Lambda} \gamma^k \partial_k + \frac{e}{\Lambda} I(e) + 1 \right] \psi^{(p)} = 0 . \quad (7.3)$$

The operator $(1/\Lambda) \gamma^k \partial_k$ corresponds to the ratio of the proton's kinetic energy to its rest energy. Since the proton's motion relative to the electron is negligible compared with its rest energy (around 1 GeV), we may neglect this term in (7.3). Further, the operator $eI(e)/\Lambda$ corresponds to the ratio of the electron binding to the proton (around 1 eV) to the proton's rest energy. This may also be assumed to be negligible in the problem at hand. Neglecting the latter two terms in (7.3) and choosing $s^\dagger = (1000)$ in the proper Lorentz frame, (7.3) takes the form

$$\left[\frac{1}{\Lambda} \gamma^0 \partial_0 + 1 \right] \exp(-i\Lambda t) f(\mathbf{r}) s = 0 . \quad (7.4)$$

Next, it is observed that the space-dependent part $f(\mathbf{r})$ of the proton solution may be chosen arbitrarily without altering the validity of (7.4). Appealing once again to the physical argument that led to the proton field being at rest at the origin, relative to the electron field, we may take $f(\mathbf{r})$ to be

$$|f(\mathbf{r})|^2 = \delta(\mathbf{r}) , \quad (7.5)$$

where $\delta(\mathbf{r})$ is the three-dimensional Dirac delta function.

Using (7.5) in (7.2), we have

$$\psi^{(p)\dagger} \gamma^0 \gamma^k \psi^{(p)} = 0 , \quad \psi^{(p)\dagger} \gamma^0 \psi^{(p)} = \delta(\mathbf{r}) , \quad (7.6)$$

where $k = 1, 2, 3$ denote the three spatial coordinates.

According to (5.49) and (5.50), $I(p)$ is in two parts: $I(p)_1 + I(p)_2$. Inserting (7.6) into the first part of this interaction term,

$$I(p)_1 = \frac{\gamma^0 e}{r} . \quad (7.7)$$

Before the interaction $I(p)_2$ can be determined, we need to know the form of the spinor solutions of the electromagnetic field equations for the proton solution $\phi_\alpha^{(p)}$. These solutions have been derived for the case where the proton is stationary at the origin ([117], Sect. 5.5) where it is found that the spinor electromagnetic equations for the proton are

$$\sigma^\mu \partial_\mu \phi_1^{(p)} = -4\pi i e \delta(r) \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \quad \sigma^\mu \partial_\mu \phi_2^{(p)} = 4\pi i e \delta(r) \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \quad (7.8)$$

The particular solutions of these equations were found to be

$$\phi_1^{(p)} = -\frac{ie}{r^3} \begin{pmatrix} x_3 \\ x_1 + ix_2 \end{pmatrix}, \quad \phi_2^{(p)} = \frac{ie}{r^3} \begin{pmatrix} x_1 - ix_2 \\ -x_3 \end{pmatrix}. \quad (7.9)$$

Finally, with substitution of the solutions (7.9) into the form of $I(p)_2$ according to (5.49) and (5.50), it follows that

$$I(p)_2 = 16\pi ig_M \frac{e}{r^3} (\mathbf{r} \times \boldsymbol{\gamma})_3. \quad (7.10)$$

The noncovariant appearance of $I(p)_2$ in terms of the third component of a vector (the axis of quantization) is a consequence of the approximation in which the spatial components of the proton's vector potential were taken to be zero, leaving A_0 – the Coulomb potential of the proton – to act on the electron.

Adding (7.7) and (7.10) for the total interaction functional of the proton acting on the electron field, the linearized electron equation (the wave equation for hydrogen) takes the following form, in terms of the Dirac α and β matrices:

$$\left[-\alpha_r p_r - i \frac{\alpha_r}{r} \beta K^\wedge + \frac{e^2}{r} + 16\pi g_M \frac{e^2}{r^3} (\mathbf{r} \times \boldsymbol{\alpha})_3 - \beta \lambda + E \right] \psi^{(e)} = 0. \quad (7.11)$$

The following notation is used above:

$$\begin{aligned} \beta &= \gamma^0, \quad \boldsymbol{\alpha} = i\gamma^0 \boldsymbol{\gamma}, \quad K^\wedge = \beta(\boldsymbol{\sigma} \cdot \mathbf{L} + 1), \quad p_r = -i \left(\frac{\partial}{\partial r} + \frac{1}{r} \right), \\ \alpha_r &= \frac{\boldsymbol{\alpha} \cdot \mathbf{r}}{r}. \end{aligned} \quad (7.12)$$

The definition of the Dirac matrices γ^μ is given in (6.29) and $\mathbf{L} = -i\mathbf{r} \times \boldsymbol{\nabla}$ is the angular momentum operator.

It is convenient at this point to make the following change of variables: $\rho = \xi r$, with $\xi = \xi_1 \xi_2$, $\xi_1 = \lambda + E$, $\xi_2 = \lambda - E$. The electron equation then takes the Hamiltonian form:

$$H\psi = (H_0 + V)\psi = -i \frac{E}{\xi} \psi, \quad (7.13)$$

where

$$i\xi H_0 = i\xi \left[\alpha_\rho \left(\frac{\partial}{\partial \rho} + \frac{1}{\rho} \right) - \frac{\alpha_\rho \beta K^\wedge}{\rho} - i \left(\frac{\gamma}{\rho} + \frac{\beta \lambda}{\xi} \right) \right] \quad (7.14)$$

is the unperturbed Dirac Hamiltonian for the hydrogenic electron and $\gamma = 2\pi e^2/\hbar c$ is the fine structure constant ($\approx 1/137$). The unperturbed energy eigenvalues of iH_0 are then

$$\frac{E_{Jn}}{\xi} = \frac{\lambda}{\xi} \left[1 + \frac{\gamma^2}{(s+n)^2} \right]^{-1/2}. \quad (7.15)$$

The quantum numbers indicated above are

$$s^2 = K^2 - \gamma^2, \quad K^2 = \left(J + \frac{1}{2} \right)^2,$$

where J is the total angular momentum quantum number and n is a positive integer or zero. $N = n + 1$ is called the principle quantum number.

7.2.2 The Lamb Splitting

Treating V in (7.13) as a perturbation on the eigenfunctions of H_0 , the Dirac eigenvalues then follow for the energy levels of hydrogen. The extra interaction V in (7.13) comes from the interaction functional $I(p)_2$, which in turn follows from the spinor generalization of electromagnetic theory, as shown in Chap. 5.

The feature of the predicted energy level spectrum of hydrogen (7.15) that does not agree with the empirical spectrum is the accidental degeneracy that appears in the Dirac states of hydrogen, when the principle quantum numbers N are greater than unity (i.e., when $n > 0$). The Dirac states for the pair of eigenvalues $\pm K$ (of the operator $K^\wedge$) correspond to the same energy value E_{Jn} . For example, when the total angular momentum is $J = 1/2$, the states $NS_{1/2}$ and $NP_{1/2}$ are degenerate – they are at the same energy value. However, Lamb and his co-workers found that, empirically, the energies of the states $2S_{1/2}$ and $2P_{1/2}$ are not the same, with $E(2S_{1/2}) > E(2P_{1/2})$. It was also found in experimentation that $E(3S_{1/2}) > E(3P_{1/2})$ and $E(4S_{1/2}) > E(4P_{1/2})$. These Lamb splittings (called the Lamb shift in quantum electrodynamics) are not predicted by the Dirac theory for hydrogen.

In contemporary physics, these energy splittings in the fine structure of hydrogen were predicted by quantum electrodynamics (QED), in close agreement with the data, except for the case $N = 4$. In the latter case, the observation of the Lamb shift is not in agreement with the calculations from QED, to the accuracy of the measurements.

If the Lamb splittings are to be predicted by the theory proposed in this study, instead of QED, they must come from the perturbing potential in (7.13):

$$iV = \frac{\kappa}{\rho^2} \frac{(\boldsymbol{\rho} \times \boldsymbol{\alpha})_3}{\rho} . \quad (7.16)$$

The lifting of the accidental degeneracy in the Dirac states for hydrogen is due to the lower symmetry of this perturbing potential compared with the coulomb term in the unperturbed Hamiltonian H_0 . The strength of this potential is measured by the constant

$$\kappa = 16\pi g_M \gamma \xi = 16\pi [(s+n)^2 + \gamma^2]^{-1/2} \frac{g_M}{\lambda_C} \gamma^2 , \quad (7.17)$$

where $\lambda_C = h/2\pi mc$ is the (reduced) Compton wavelength of the electron.

To proceed with the perturbation treatment of the interaction (7.16) on the Dirac states, we note that the unperturbed states diverge as ρ^{-2} at the origin. To ensure that the solutions of the unperturbed problem lead to a rapidly convergent series solution at the origin, that depends successively on ρ^{-n} ($n \geq 2$), we will redefine the unperturbed and the perturbed energy by adding and subtracting the term $i\kappa\alpha_\rho/\rho^2$. For the unperturbed Hamiltonian we thus take

$$iH'_0 = i \left(H_0 - \frac{\kappa\alpha_\rho}{\rho^2} \right) \quad (7.18)$$

and for the perturbing part

$$iV' = \frac{\kappa}{\rho^2} \left[\frac{(\boldsymbol{\rho} \times \boldsymbol{\alpha})_3}{\rho} + i\alpha_\rho \right] . \quad (7.19)$$

The total Hamiltonian remains the same, $H = H_0 + V = H'_0 + V'$.

The new unperturbed electron equation with the new unperturbed Hamiltonian,

$$H'_0 \psi = -\frac{i}{\xi} E \psi , \quad (7.20)$$

can be solved exactly, giving the solution

$$\psi = \psi_D \exp \left(-\frac{\kappa}{\rho} \right) , \quad (7.21)$$

where ψ_D are the Dirac eigensolutions of the hydrogen equation (7.14). Note that the new exact solutions ψ converge at the origin more rapidly than any polynomial in $1/\rho$ would diverge there.

The accidentally degenerate states of the unperturbed Hamiltonian correspond to the quantum numbers $K = \pm 1$. It can be shown ([136], Sect. 8.2) that the Lamb splitting between these states is

$$\begin{aligned}\Delta_n &= E[(n+1)S_{1/2}] - E[(n+1)P_{1/2}] = \frac{3}{4}\tau_n E_n \\ &= \frac{64\pi}{3[(s+n)^2 + \gamma^2]} \frac{g_M}{\lambda_C} \gamma^4 E_n (mc^2),\end{aligned}\quad (7.22)$$

where $s^2 = K^2 - \gamma^2$, $E_1 = 1/3 + O(\gamma^2)$, $E_2 = 2/9 + O(\gamma^2)$ and $E_3 = 1/6 + O(\gamma^2)$, where $O(\gamma^2)$ represents the remainder of the respective series, starting with terms in γ^2 . Thus, to order γ^2 , the Lamb splitting in the first excited state of hydrogen, with principle quantum number $N = n + 1 = 2$ is

$$\Delta_1 = \frac{16\pi}{9} \gamma^4 \frac{g_M}{\lambda_C} mc^2. \quad (7.23)$$

We then find that the ratio, which is independent of the constant g_M to first order, is

$$\frac{\Delta_2}{\Delta_1} = 0.2965. \quad (7.24)$$

The experimental value of this ratio is

$$\left(\frac{\Delta_2}{\Delta_1}\right)_{\text{exp}} = 0.2976 \pm 0.0003. \quad (7.25)$$

We see, then, that the difference between the theoretical and the experimental values of this ratio for the Lamb splittings is of the order of only 0.2%. In view of the approximations that have been made in the theoretical analysis, it may be concluded that this theory is successful in its prediction of the hydrogen spectrum, including the Lamb splitting for the principle quantum number state $N = 2$.

Using the experimental value for $\Delta_1/h = 1057.8514 \pm 0.0019$ MHz in (7.23), we find that the new constant of this theory has the value

$$g_M = (2.087 \pm 0.001) \times 10^{-14} \text{ cm}. \quad (7.26)$$

With this value, the experimental Lamb splitting

$$\frac{(\Delta_2)_{\text{exp}}}{h} = 314.819 \pm 0.048 \text{ MHz}$$

is in agreement with the theoretical value from (7.22), within the experimental accuracy. The theoretical value for $N = 3$ is then

$$\frac{\Delta_3}{h} = \frac{E(4S_{1/2}) - E(4P_{1/2})}{h} = 132.22 \text{ MHz}.$$

This is to be compared with the experimental value

$$\frac{(\Delta_3)_{\text{exp}}}{h} = 132.53 \text{ (+0.58, -0.78) MHz}.$$

Thus, the theoretical and the experimental value of the Lamb splitting Δ_3 are in agreement within the experimental accuracy of this measurement.

The theoretical result calculated with QED for Δ_3 is 133.084 ± 0.001 MHz [R.A. Brown and F.M. Pipkin, *Annals of Physics* **80**, 479 (1973).] The theoretical determination in this book is then in better agreement with the data than is the calculation from QED.

Aside from these differences, it is interesting to note that, while the two theories are totally different from each other, both conceptually and mathematically, they give results that are so close on the matter of the Lamb shift. It is my speculation that this is not purely coincidental!

Summing up, we have seen that the special relativity limit of a generally covariant field theory of inertia that is incorporated in a general unified field theory yields a prediction of the entire hydrogen spectrum, including the Lamb splittings in the otherwise degenerate states of hydrogen. The theory, along with the earlier predicted Pauli exclusion principle, then extends to the spectra of all atomic species, beyond hydrogen. In the next section, we will examine features of elementary particles. Numerical predictions will be made in regard to features of the neutron, pion and kaon – all predicted as composites of more elementary matter physical fields in the context of this field theory. We will see that the only fundamental matter fields are: electron, positron, proton and antiproton. All other elementary particles fields are (unstable) composites of these, along with the virtual photon and virtual neutrino, to affect their coupling.

7.3 The Neutron

The neutron, in this theory, is a composite of an electron and a proton (e-p) coupled by a virtual neutrino field – a short-range pseudoscalar coupling. This is the second state of the e-p, the first being its state

bound by the long-range scalar coupling, represented by the virtual photon. These two electromagnetic interactions – one scalar and the other pseudoscalar – are predicted by the factorized spinor formulation of electromagnetic field theory, which does not recognize reflection symmetry in (any of) the laws of nature. This is in accordance with the symmetry group of the theory of relativity – the Einstein group – a 16-parameter Lie group which is a set of continuous, analytic transformations that leave the laws of nature covariant.

In this context, the neutrino field is a special solution of the spinor formulation of electromagnetism. In special relativity, it is the solution of the Weyl equation, for the proton pair in its ground state of null energy-momentum, as derived in the last chapter:

$$\sigma^\mu \partial_\mu [\phi_\alpha^{(p+)} + \phi_\alpha^{(p-)}] = 0 . \quad (7.27)$$

When the spinor field of the pair in the ground state, $\phi_\alpha^{(\text{pair})} = \phi_\alpha^{(p+)} + \phi_\alpha^{(p-)}$ couples to the electron and proton fields to bind them, there is a polarization induced due to the coupling of $\phi_\alpha^{(p+)}$ to the electron component and the coupling of $\phi_\alpha^{(p-)}$ to the proton component of e-p. The composite matter field

$$n \equiv e-\phi_\alpha^{(\text{pair})}-p \quad (7.28)$$

then forms the neutron. Because all three of these constituent fields are fermions (spin-1/2 fields), their composite n is also a fermion, as required empirically.

7.3.1 Binding Energy of the Neutron: A Phenomenological Determination

The binding of the electron to the proton due to their electromagnetic coupling to the proton pair field (7.28) of the physical vacuum, is in the domain of the nuclear force range (of the order of 10^{-13} cm). The binding energy of this composite is then of the order of $2e^2/\langle r \rangle$, where $\langle r \rangle$ is the ‘neutron radius’. The latter is of the order of the potential well width, which is the range of influence (i.e., the nuclear force) of a neutron on another neutron or proton. This is equivalent to replacing the Yukawa nuclear potential $[\exp(-\mu r)]/r$ with a square well of width μ^{-1} . This is the order of magnitude of the deuteron well width, around 4×10^{-13} cm. According to this phenomenology, then, the neutron binding energy should be of the order of

$$B(n) = \frac{2e^2}{4 \times 10^{-13} \text{ cm}} = 0.72 \text{ MeV} .$$

In terms of this model, the empirical value of the binding energy of the neutron is

$$\begin{aligned} B(n)_{\text{exp}} &= m_n - (m_p + m_e) \\ &= 939.5527 - (938.2592 + 0.5110) = 0.7825 \text{ MeV} . \end{aligned}$$

These two values of $B(n)$ are close enough, in view of the approximation used here, to give credence to the composite model of the neutron considered.

If $B(n)$ units of energy are then delivered to this composite neutron, it will decay as follows:

$$n \longrightarrow p + e + \phi_{\alpha}^{(\text{pair})} , \quad (7.29)$$

where $\phi_{\alpha}^{(\text{pair})}$ plays the role of the (anti)neutrino that is emitted in the neutron decay. The mechanism of the decay is the pseudoscalar part of the electromagnetic interaction that is identified as the weak interaction. The neutrino energy (the pair) is then absorbed by the sea of pairs of the physical vacuum in the vicinity of the decayed neutron.

7.3.2 The Neutron Lifetime

The empirical value for the lifetime of the neutron, decaying according to (7.29), is

$$\tau_n^{-1} = (1.114 \pm 0.016) \times 10^{-3} \text{ s}^{-1} . \quad (7.30)$$

Using time-dependent perturbation theory, the decay lifetime is

$$\tau_n^{-1} = \frac{4\pi^2}{h} |\langle f | I | i \rangle|^2 \rho_f \text{ s}^{-1} , \quad (7.31)$$

where we have inserted Planck's constant h . I is the interaction functional that projects the state of the composite neutron – the initial state $|i\rangle$ of n and its surroundings of pairs of the physical vacuum, that causes the breakup, into the final state $\langle f|$ of freed breakup products, e , p and the neutrino field $\phi_{\alpha}^{(\text{pair})}$. The density of final states is ρ_f for the three breakup products, electron, proton and (anti)neutrino. In earlier work [133] it was found to have the form

$$\rho_f = \frac{V^2 (2\pi)^6}{2\pi^3 h^6} \left(\frac{m_e m_p m_{\nu}}{m_e + m_p + m_{\nu}} \right)^{3/2} W^2 \text{ erg}^{-1} . \quad (7.32)$$

In this formula, V is the interaction volume defined by Fermi in his analysis of breakup reactions [E. Fermi, Prog. Theor. Phys. **5**, 570 (1950)]. W is the total kinetic energy of the three breakup products.

According to (6.30b), derived in the last chapter, the interaction functional I has the form

$$I = ie g_M \sum_{\alpha} (-1)^{\alpha} (\phi_{\alpha}^{\dagger} \cdot \Gamma_{\alpha} - \gamma^0 \Gamma_{\alpha} \gamma^0 \cdot \phi_{\alpha}) , \quad (7.33)$$

where g_M , determined from the Lamb shift, is shown in (7.26).

At this stage of the analysis, we see that the electromagnetic field intensity spinor ϕ_{α} is proportional to e/r^2 . Thus the magnitude of the matrix element in (7.31) is of the order of $g_M e^2 \langle f | 1/r^2 | i \rangle$. When evaluating this term, we should note that the matter field of the composite neutron is weighted with the screening factor $\exp(-r/\lambda)$, where λ is the proton Compton wavelength. This follows from the screening action of the proton pairs of the physical vacuum on the interacting constituents of the composite neutron. It then follows that

$$\langle f | 1/r^2 | i \rangle = \frac{4\pi \int_0^{\infty} \exp(-r/\lambda) (1/r^2) r^2 dr}{4\pi \int_0^{\infty} \exp(-r/\lambda) r^2 dr} .$$

Thus we find, after integration, that

$$\langle f | I | i \rangle \sim \frac{e^2}{\lambda} .$$

It then follows that the neutron decay rate is

$$\tau_n^{-1} = \frac{V^2 (e^2/\lambda)^2 (2\pi)^2}{\pi^2 h^7} \left(\frac{m_e m_p m_{\nu}}{m_e + m_p + m_{\nu}} \right)^{3/2} W^2 \text{ s}^{-1} . \quad (7.34)$$

At the present stage of experimental elementary particle physics, there seems to be general agreement that the neutrino has a small but non-zero inertial mass. In some of my early research, I found that a neutrino field must have non-zero inertial mass in a curved spacetime. But it was found to be very small in the curvature determined by the Schwarzschild field of a neutron. Still, it was non-zero [81]. If we take $m_{\nu} = 17.1 \text{ keV}$ in (7.34), we obtain a result for the neutron decay rate that is compatible with the observed neutron decay rate. Future experimentation will be needed to establish the unique theoretical prediction for the neutrino mass in this theory. One feature of this theory

is that indeed the neutrino has a non-zero inertial mass, which is a consequence of the curvature of spacetime, according to the theory of general relativity.

Taking the Fermi interaction volume to be a sphere of radius 10^{-15} cm (the range of the weak interaction) then $V = (4\pi/3)10^{-45}$ cm³. The Compton wavelength of the proton is the order of 2×10^{-14} cm and

$$W = m_n - m_p - m_e = 0.78 \text{ MeV} .$$

Assuming the mass of a neutrino to be $m_\nu = 17.1$ keV, we find that the decay rate of the neutron is

$$\tau_n^{-1} = 0.497 \times 10^{-3} \text{ s}^{-1} . \quad (7.35)$$

The empirical value of this decay rate, shown in (7.30), is then only of the order of two times greater. In view of the assumptions used in this phenomenological derivation for the neutron decay rate, it gives enforcement to the model of the neutron as the composite $e-\phi_\alpha^{(\text{pair})}-p$.

7.3.3 The Neutron Magnetic Moment

A feature of the neutron that must be addressed is that, while it is charge-neutral, it has a non-zero magnetic moment. It is negatively polarized with the same order of magnitude as the positively polarized magnetic moment of the proton. Empirically,

$$\mu_n = -1.9103 \frac{eh}{4\pi m_p c} , \quad \mu_p = 2.7896 \frac{eh}{4\pi m_p c} .$$

We saw in the preceding discussion that the binding energy of the composite neutron has the correct order of magnitude. The magnitude of the negative polarization of the neutron magnetic moment, μ_n , viewed as the composite $p-\phi_\alpha^{(\text{pair})}-e$ is sensitive to the mass of the particles of the mediating proton pair (that is, the mediating ‘neutrino’). If the predominant contribution to μ_n is due to the rotation about the ‘observed’ proton of the antiproton component of the proton pair, then the predicted sign of μ_n must be negative, and its magnitude must be proportional to e/m_p , as is the case, empirically.

We have seen from this phenomenological analysis that the proposed composite model of the neutron is compatible with the neutron’s binding energy (as measured in its beta decay) and the sign and order

of magnitude of the neutron magnetic moment. Furthermore, it identifies the breakup products with the electron and the proton, and identifies the (anti)neutrino as transferring electromagnetic energy from the neutron to the physical vacuum in terms of the factorized spinor electromagnetic field for the proton pair in its ground state of null energy–momentum.

7.4 Mass Doublets: The Electron–Muon

A mystery in elementary particle physics, ever since the cosmic ray observations in the 1930s, has to do with the existence of the muon. In the context of physical interactions with other elementary matter, the muon is a particle that is identical with the electron, except for its mass and instability. Why does the muon exist?

The mass of the muon is approximately 206 times that of the electron and its lifetime is the order of 2×10^{-6} s. It decays in this time to an electron and two (conjugated) neutrinos.

The field theory of inertia from general relativity, developed in Chap. 4, leads to a possible answer to this problem. Because of the spinor structure of the formalism, there is a prediction here of mass doublets – for every spinor matter field with mass m , there is a heavier matter field with mass $m^* > m$, wherein m and m^* are the mass eigenvalues of a mass operator defined in general relativity.

From the analysis in Chap. 4, combining (4.13) and (4.15) yields the following eigenvalue equation:

$$\Xi \eta \equiv (q^\mu \Omega_\mu)(-q^{*\nu} \Omega_\nu^\dagger) \eta = \lambda^2 \eta . \quad (7.36)$$

Here, Ξ is a two-dimensional mass operator, with two (squared) mass eigenvalues λ_1^2 and λ_2^2 . The two distinct mass eigenvalues are then $\lambda_1 = +\sqrt{\lambda_1^2}$ and $\lambda_2 = +\sqrt{\lambda_2^2}$. The plus sign only must be taken because we have seen earlier that the mass field λ , being the modulus of a complex variable, is positive-definite.

In the asymptotic limit that approaches a flat spacetime, the solutions η of (7.36) become the set of eigenfunctions $\{|\eta_s\rangle\}$ of a linear Hilbert space, where the normalization and integrability of this space of solutions of (7.36) yields an infinite spectrum of mass eigenvalues:

$$\lambda_s^2 = \langle \eta_s | (q^\mu \Omega_\mu)_a (-q^{*\nu} \Omega_\nu^\dagger)_a | \eta_s \rangle . \quad (7.37)$$

The subscript ‘a’ refers to the asymptotic values of the quaternion fields in the near flat spacetime limit.

Thus, in addition to the physical prediction of mass doublets, this theory predicts an infinite mass spectrum of elementary matter. In the actual limit, whereby all other matter surrounding the ‘observed matter’ is depleted, the spin affine connection Ω_μ vanishes and the mass of the observed matter vanishes. This is in accordance with the Mach principle.

It was found, to a first approximation in the fine structure constant $\gamma = 2\pi e^2/hc \approx 1/137$, that the ratio of mass eigenvalues for the electron pair and the proton pair, viz., $m(\mu)/m(e) = m(p_h)/m(p) = 3/2\gamma \approx 206$, where p_h is the predicted ‘heavy proton’. The empirical discovery in the 1930s that the mass of the muon $m(\mu) \approx 206m(e)$ is then interpreted as meaning that the muon is the heavy member of the electron mass doublet in general relativity [83]. Thus far, there is no experimental evidence for the existence of the heavy proton p_h , whose mass is $206m(p) \approx 193 \text{ GeV}$ [106].

The cause for the generation of the inertial mass of elementary matter is the coupling of the ‘observed matter’ to the particle–antiparticle pairs of the physical vacuum, in which it is embedded. The scenario for the generation of the heavier sister of the stable particle (electron or proton) is the following. The observed matter excites a pair in its vicinity, electromagnetically. This then changes the geometrical fields in the environment of the observed particle, which in turn alters its mass to the higher value of the mass doublet, $e \rightarrow \mu$ or $p \rightarrow p_h$. The coupling of the excited pair of the physical vacuum to its environment of pairs in their ground states of null energy–momentum then causes the de-excitement of this pair. The higher mass particle, the muon or p_h , then decays back to the ground level mass of the electron or the proton, with

$$p_h \text{ (or } \mu) \rightarrow p \text{ (or } e) + \nu + \nu^*, \quad (7.38)$$

where $\nu \equiv \phi_\alpha(p^+ - p^-)$ and $\nu^* \equiv \phi_\alpha^*(p^+ - p^-)$ are the time-reversed electromagnetic spinor fields, transferred back to the physical vacuum in this decay process. A phenomenological calculation was carried out on muon decay, where it was found to be the same order of magnitude as the empirical result, around $2 \times 10^{-6} \text{ s}$ [84].

7.5 Infinite Lepton Spectrum

In the nomenclature of elementary particle physics, a lepton is any particle that does not couple to nuclear particles with the strong interaction – such as the strength of coupling of a neutron to a proton

in the domain of an atomic nucleus. Examples of such matter fields are the electron, muon and neutrino.

The generation of the muon from an electron was seen to be the effect of the observed electron exciting a particle–antiparticle pair of the physical vacuum to an excited level, thereby altering the geometrical fields that define the inertial mass of matter. But if the observed electron excited two or more particle pairs of the physical vacuum, other heavier sisters – leptons – of the electron could be generated, which we may call τ , τ' , τ'' , etc., and their accompanying neutrinos. That is, an infinite spectrum of leptons could be generated. This result would be in contradiction with the prediction of the Standard Model of contemporary elementary particle physics, which predicts the existence of only three leptons, and their accompanying neutrinos. Experimental physics has thus far verified the existence of the τ particle – in this theory, the first of an infinite spectrum of such particles of matter.

In the analysis of the decay rate of the muon, it was found that if the breakup is into n particles, the density of final states in (7.31) depends on the factor $W^{(3n-5)/2}$ [101], where W is the kinetic energy of the final state and n is the number of discrete matter fields that the initial body breaks into. In (7.32), $n = 3$. Since the final fields and the interaction I are the same in channels

$$e^* \equiv \mu \longrightarrow e + \phi_\alpha + \phi_\alpha^* + \phi_\alpha + \phi_\alpha^* , \quad (7.39a)$$

$$e^{**} \equiv \tau \longrightarrow e + \phi_\alpha + \phi_\alpha^* + \phi_\alpha + \phi_\alpha^* , \quad (7.39b)$$

and since there are $n = 5$ breakup fields in both reactions, so that in both cases $W^{(3n-5)/2} = m^5$, it follows that the ratio of lifetimes depends inversely on the fifth power of the mass ratio of the tau and the mu particles:

$$\frac{T_\tau}{T_\mu} = \left(\frac{m_\mu}{m_\tau} \right)^5 . \quad (7.40)$$

According to the experimental data, $m_\tau = 1784$ MeV, $m_\mu = 106$ MeV, $T_\mu = 2 \times 10^{-6}$ s, $T_\tau = 3 \times 10^{-13}$ s. This gives

$$\left(\frac{T_\tau}{T_\mu} \right)_{\text{exp}} = \left[0.2 \left(\frac{m_\mu}{m_\tau} \right)^5 \right]_{\text{exp}} . \quad (7.41)$$

This indicates the same order of magnitude as the theoretical prediction (7.40), thereby verifying the theoretical structure proposed.

The lifetime T_τ corresponds to the excitation of two proton pairs of the physical vacuum rather than one, which generates the tau particle. The electron excitation of three, four or more pairs in its vicinity

would correspond to $n = 7, 9, \dots$, with the corresponding exponents $(3n - 5)/2 = 8$ for the τ' lepton, $(3n - 5)/2 = 11$ for the τ'' lepton, etc. Thus, the ratio of lifetimes of the higher mass tau particles would be

$$\frac{T_{\tau'}}{T_{\mu}} = \left(\frac{m_{\mu}}{m_{\tau'}}\right)^8, \quad \frac{T_{\tau''}}{T_{\mu}} = \left(\frac{m_{\mu}}{m_{\tau''}}\right)^{11}, \dots \quad (7.42)$$

The lifetimes of the higher mass tau particles and their masses are yet to be determined experimentally. But the ratios (7.42) are the predictions thus far for the infinite spectrum of higher mass tau leptons, compared with that of the muon.

7.6 The Pion

The pion plays a crucial role in the binding of the nucleus by propagating between the neutron and proton constituents (called nucleons), to effect the nuclear force field.

The pion is a pseudoscalar field that is a boson, since it has zero angular momentum. It solves the Klein–Gordon field equation

$$(\square + \Delta)\phi(\pi) = \lambda^2\phi(\pi),$$

where $\square = \partial_0^2 - \nabla^2$ is the d'Alembertian operator and Δ will be defined below.

The pion decays by means of the weak interaction. There are charge-neutral and charged pions. In this theory, the neutral pion decays as follows:

$$\pi^0 \equiv (p^+ - p^-)_{-} \longrightarrow (p^+ - p^-)_{+} \equiv 2\gamma. \quad (7.43)$$

The subscripts $-$ and $+$ respectively denote the antisymmetric (pseudoscalar) neutral pion and the symmetric (scalar) composites of proton and antiproton matter fields. The symmetric state is mathematically equivalent to two photons, as we discovered in the case of the dynamics of the null-energy ground state of the electron–positron (or proton–antiproton) pair.

The charged pions are composites that decay as follows:

$$\pi^{\pm} \equiv [p^{\pm} - \phi_{\alpha}(p^+ - p^-)] \longrightarrow (e^{\pm})^* + \phi_{\alpha}(p^+ - p^-) \equiv \mu^{\pm} + \nu. \quad (7.44)$$

In this case, the mass-excited electron – the muon – is created by a transfer of the proton of the composite pion to the physical vacuum,

absorbing it in the sea of proton–antiproton pairs that constitute the physical vacuum. This transfer, in turn, excites an electron–positron pair, giving rise to an elevation of an electron in this sea, thus yielding its heavy sister, the muon.

Empirically, we observe that the pion transforms as a pseudoscalar field in space. The pseudoscalar pion may be expressed as the following combination of first-rank spinor variables:

$$|\pi\rangle \propto \eta^\dagger \chi - \chi^\dagger \eta ,$$

where (η, χ) are the two-component matter field solutions and $\chi = \varepsilon \eta^*$ is the space (or time) reflection of η .

In the linear limit of a Hilbert space, the (squared) pion mass eigenvalues have the expectation values [86]

$$\lambda^2 = \langle \phi(\pi) | \square + \Delta | \phi(\pi) \rangle , \quad (7.45)$$

where

$$\Delta = II^* + (I^* q^\mu + q^{\mu*} I + q^\nu \partial_\nu q^{\mu*}) + q^{\mu*} \partial_\mu I ,$$

and $\square$ is the d'Alembertian operator $-\partial^\nu \partial_\nu$.

7.6.1 Ratio of Neutral and Charged Pion Masses

In the case of the neutral pion, the expectation value of I (representing the electromagnetic coupling to the proton and electron pairs of the physical vacuum) is much smaller than the kinetic energy portion of the expectation value of $\langle \square \rangle$. If we assume that the kinetic energy terms are the same for the neutral and charged pions, i.e.,

$$\langle \phi(\pi^0) | \square | \phi(\pi^0) \rangle \approx \langle \phi(\pi^\pm) | \square | \phi(\pi^\pm) \rangle ,$$

it then follows that the (squared) ratio of their respective masses is

$$\left[\frac{m(\pi^0)}{m(\pi^\pm)} \right]^2 = (1 + R)^{-1} , \quad (7.46)$$

where

$$R = \frac{\langle \phi(\pi^\pm) | I | \phi(\pi^\pm) \rangle}{\langle \phi(\pi^0) | \square | \phi(\pi^0) \rangle} . \quad (7.47)$$

In the paper cited above, it is shown that $R \approx 18/2\pi^3$. Thus, the ratio of the masses of the neutral to the charged pions is

$$\frac{m(\pi^0)}{m(\pi^\pm)} = \left[1 + \frac{18}{2\pi^3}\right]^{-1/2} = 0.966. \quad (7.48)$$

The empirical value for this ratio is

$$\left[\frac{m(\pi^0)}{m(\pi^\pm)}\right]_{\text{exp}} = \frac{134.9645}{139.5688} = 0.967. \quad (7.49)$$

The difference between the theoretical and experimental values above is of the order of only 0.1%, thus verifying the composite model of the pion that is shown in (7.43) and (7.44).

7.6.2 Ratio of Neutral and Charged Pion Lifetimes

While the masses of the charged and neutral pions are close, it is significant that the ratio of their lifetimes is the order of 10^{-8} :

$$\pi^0 \longrightarrow 2\gamma, \quad T(\pi^0) = 0.828 \times 10^{-16} \text{ s}, \quad (7.50a)$$

$$\pi^\pm \longrightarrow \mu^\pm + \nu, \quad T(\pi^\pm) = 2.603 \times 10^{-8} \text{ s}. \quad (7.50b)$$

To calculate the theoretical ratio of these lifetimes, we use the golden rule from time-dependent perturbation theory (7.31). With Fermi's phenomenological approach for breakup reactions, the matrix element $|\langle f|I|i\rangle|^2$ is replaced with $D^2(\Omega/V)^{n-1}$, where D is the effective potential well depth for the intrinsic interaction that binds the initial composite pion, Ω is the interaction volume of the pion's breakup into n objects, and V is an arbitrary volume in which the decay is observed. For the decay products indicated in (7.43) and (7.44), and the density of states for n decay products [101],

$$\rho_f = \frac{V^{n-1} \prod_i (2I_i + 1) (2\pi)^{3(n-1)/2} \prod_i m_i^{3/2} W^{(3n-5)/2}}{h^{3(n-1)} \Gamma(3(n-1)/2) \left(\sum_i m_i\right)^{3/2}}, \quad (7.51)$$

where $\Gamma(n) = (n-1)!$, $\Gamma(1/2) = \sqrt{\pi}$ is the gamma function, I_i is the spin of the i th product of the breakup and m_i is its mass. It then follows that

$$\frac{T(\pi^0)}{T(\pi^\pm)} = \frac{(m_p m_\pi)^{3/2} c^3 \Omega_\pi}{(48\sqrt{3}) 8\pi^4 h^3}. \quad (7.52)$$

We have assumed here that the interaction volume Ω is the same for the neutral and charged pions. The latter is the only unknown in (7.52). From the analysis of the electron-positron pair volume that determined the electron and muon masses, the interaction volume is a sphere with a radius of the order of magnitude 10^{-15} cm. Thus, we take $\Omega_\pi = 10^{-45}$ cm³/pair. The following result is then obtained from (7.52):

$$\frac{T(\pi^0)}{T(\pi^\pm)} = 2.3 \times 10^{-8} . \quad (7.53)$$

The experimental value for this ratio is 0.32×10^{-8} . These results are in agreement within a factor of 10. In view of the assumptions used to obtain the theoretical result (7.53), I believe that the composite model of the pion used here [equations (7.43) and (7.44)] is justified.

7.7 CP Violation in Neutral Kaon Decay

An elementary matter field that decays into pions and plays a role in the nuclear domain of matter is the kaon. As in the case of pions, there is a charge-neutral kaon and oppositely charged kaons. The neutral kaon decays into two channels: one is short-lived, K_S^0 , and the other is long-lived, K_L^0 .

In the context of this field theory, which does not recognize reflection symmetry, K_S^0 is the scalar part of the scalar-pseudoscalar pair of kaon matter fields. It is seen to decay as follows:

$$K_S^0 \longrightarrow \pi^+ + \pi^- .$$

The long-lived neutral kaon is more interesting. It is a pseudoscalar matter field that decays to three pions in one channel whereas, for a small fraction of the time, it decays to two pions:

$$K_L^0(1) \longrightarrow \pi^+ + \pi^- + \pi^0 \quad 0.998 \text{ of the time} , \quad (7.54a)$$

$$K_L^0(2) \longrightarrow \pi^+ + \pi^- \quad 0.002 \text{ of the time} . \quad (7.54b)$$

The decay of $K_L^0(1)$ establishes that it is a pseudoscalar matter field, since it transforms into three pions, each of which is pseudoscalar. The second decay, $K_L^0(2)$ violates spatial reflection symmetry (parity P), since the combination of two pions is scalar while the combination of three pions is pseudoscalar. The decay (7.54b) also violates charge conjugation symmetry C ($e \rightarrow -e$), since the charge conjugate of π^0

in the first channel (7.54a) should also be the neutral pion, which is not present in the decay (7.54b). The latter decay then automatically violates the combination CP.

In the proposed field theory, the long-lived neutral kaon is the composite

$$K_L^0 \equiv [p^+ \phi_\alpha(p^+ - p^-)] [p^- \phi_\alpha(p^+ - p^-)] (p^+ - p^-)_- . \quad (7.55)$$

Here, as earlier, $\phi_\alpha(p^+ - p^-)$ is the electromagnetic spinor solution for the proton pair in its (derived) ground state of null energy-momentum. The decay (7.54b) is to two pions and the (invisible) scalar component of the proton pair $(p^+ - p^-)_+$. The latter then blends with the background sea of similar such pairs of the physical vacuum, thus making this third component in the decay undetectable. That is, the transformation from (7.54a) to (7.54b) corresponds to the transformation $(p^+ - p^-)_- \rightarrow (p^+ - p^-)_+$.

With this view, and the equality of the density of final states in the transitions (7.54a) and (7.54b), the ratio of CP-violating K_L^0 decays to two pions to the (expected) decay of K_L^0 to three pions, in accordance with the golden rule of time-dependent perturbation theory, is the ratio of squared matrix elements

$$R = \frac{|\langle 2\pi | I_- | K_L^0 \rangle|^2}{|\langle 3\pi | I_+ | K_L^0 \rangle|^2} , \quad (7.56)$$

where I_+ and I_- are respectively the scalar and pseudoscalar parts of the generalized electromagnetic interaction predicted by this theory.

The pseudoscalar part of the interaction is of the order of

$$\langle I_- \rangle = 32\pi g_M e^2 \langle 1/r^2 \rangle , \quad (7.57)$$

where, as we learned from the Lamb shift determination, $g_M = 2.087 \times 10^{-14}$ cm. The scalar part of the interaction is the order of the usual coulomb term:

$$\langle I_+ \rangle = 2e^2 \langle 1/r \rangle . \quad (7.58)$$

The factor of 2 appears because of the coupling of two charged particle fields, p^+ and p^- to the proton or antiproton that constitute the kaon.

The theoretical ratio (7.56) is then

$$R = \left\langle \frac{16\pi g_M \langle 1/r^2 \rangle}{\langle 1/r \rangle} \right\rangle^2 .$$

The expectation values above are

$$\langle 1/r^n \rangle = \int_0^\infty \frac{\exp(-r/\lambda_e)}{r^n} r^2 dr .$$

The exponential above is the Yukawa short range nuclear force factor. The parameter $\lambda_e \approx 3.26 \times 10^{-11}$ cm is the Compton wavelength of the pion, as it was originally determined in Yukawa's analysis.

Thus we find that $\langle 1/r^2 \rangle / \langle 1/r \rangle = 1/\lambda_e$, giving

$$R = \left(16\pi \frac{g_M}{\lambda_e} \right)^2 = 0.7 \times 10^{-3} .$$

The experimental value for this ratio [J.W. Cronin, Rev. Mod. Phys. **53**, 373 (1981); V.L. Fitch, Rev. Mod. Phys. **53**, 367 (1981)] is

$$R_{\text{exp}} = 2 \times 10^{-3} .$$

The same analysis gave the ratio of lifetimes of the neutral short-lived to neutral long-lived kaons, in agreement with the data [152]. These results then reinforce the proposed composite model of the kaon.

7.8 Charge Quantization in General Relativity

A prominent feature of elementary matter is the quantization of electrical charge of interacting matter. This is the empirical fact that all electrically charged elementary matter interacts in integral units of the constant e^2 – the fine structure constant $\gamma = 2\pi e^2/hc = 1/137$, which measures the strength of the electromagnetic interaction. (The only counter-case in contemporary elementary particle physics in the Standard Model is based on the elementarity of the fractionally charged particle called the quark. The fractional charge of the quark, postulated to be a fundamental constituent of the proton and neutron, has never been detected directly in experimentation to this date.)

The quantization of electric charge of matter follows from the asymptotic features of the irreducible representations of the Einstein group, as they approach the representations of the Poincaré group of special relativity theory. This is the approach toward linearity in the flat spacetime limit from the curved spacetime of general relativity.

As we discussed in Chap. 4, the representations P of the Poincaré group of special relativity solve the set of relations

$$P^\dagger \sigma_\mu P = \frac{\partial x^{\mu'}}{\partial x^\nu} \sigma_\nu, \quad (7.59)$$

where $\partial x^{\mu'}/\partial x^\nu$ are the set of 10 continuous transformations in the flat spacetime of special relativity that relate the x' coordinate system to the x system in which the laws of nature are compared, depending on the constants of the transformation $\theta_0^{\mu\nu}$.

The double-valued solutions of this equation are the irreducible representations of the Poincaré group:

$$P(\mu\nu, 0) = \exp \frac{\sigma_\mu \sigma_\nu \theta_0^{\mu\nu}}{2}. \quad (7.60)$$

The irreducible representations E of the Einstein group are a global extension of the representations P . The extension of the constant parameters $\theta_0^{\mu\nu}$ are the spacetime-dependent parameters $\theta^{\mu\nu}(x)$. The extension of the Pauli matrices σ_μ are the quaternion fields $q_\mu(x)$. Thus, the global extension of the irreducible representations of the Poincaré group of special relativity are the irreducible representations of the Einstein group E of general relativity:

$$P(\mu\nu, 0) \longrightarrow E(\mu\nu, x) = \exp \frac{q_\mu(x) q_\nu(x) \theta^{\mu\nu}(x)}{2}. \quad (7.61)$$

In the first approximation for the quaternion metrical fields, we have the expansion

$$q_\mu(x) = \sigma_\mu + \varepsilon \Lambda_\mu(x) + O(\varepsilon^2), \quad (7.62)$$

and the transformation parameters have the expansion

$$\theta^{\mu\nu}(x) = \theta_0^{\mu\nu} + \varepsilon \vartheta^{\mu\nu} + O(\varepsilon^2). \quad (7.63)$$

The expansion parameter ε is a measure of the curvature of spacetime in the vicinity of the observed ‘particle’. This curvature is in turn a manifestation of the ‘other matter’ that couples to the observed matter.

With the approximations (7.62) and (7.63) in (7.61), the corresponding asymptotic form of the irreducible representations of the Einstein group are

$$E(\mu\nu, x)_a = P(\mu\nu, 0) P(\mu\nu, \varepsilon) \left[1 + \varepsilon (\sigma_\mu \Lambda_\nu(x) + \Lambda_\mu(x) \sigma_\nu) \theta_0^{\mu\nu} \right], \quad (7.64)$$

where $P(\mu\nu, 0)$ is given in (7.60) and

$$P(\mu\nu, \varepsilon) = \exp \frac{\varepsilon \sigma_\mu \sigma_\nu \vartheta^{\mu\nu}(x)}{2}$$

are the infinitesimally displaced representations of the Poincaré group.

The main focus of attention is the part of the representation E_a that has the product form

$$P(\mu\nu, 0)P(\mu\nu, \varepsilon) [\sigma_\mu \Lambda_\nu(x) + \Lambda_\mu(x) \sigma_\nu] . \quad (7.65)$$

It is this term that displays the symmetry of the interaction between the observed particle and its environment. When the interaction is purely electromagnetic, the term $\varepsilon \Lambda_\mu(x)$ is proportional to e^2 , since Λ_μ is the first order deviation of the metrical field q_μ from the flat spacetime form σ_μ – a deviation due to the electromagnetic coupling with strength e^2 .

We see, then, that while the eigenvalues in the first part of E_a are the spin quantum numbers, which may be generated to integral and half-integral values, in accordance with the Lie algebra of this group, so the eigenvalues of the representations (7.65) of the Einstein group are proportional to a discrete set of numbers that multiply the constant e^2 . The form $\Lambda_\nu \sigma_\mu$ is a second-rank quaternion variable, whose basis functions transform as those of the direct product representations $D^{(1/2)} \otimes D^{(1/2)} = D^{(0)} \oplus D^{(1)}$, implying that the quantum numbers multiplying e^2 may be generated only in integral steps. Thus, we have proved the charge quantization of elementary matter.

Note that it is not the charge e that is quantized in integral steps. It is rather the square of the charge e^2 that expresses the interaction with other matter, generally as e^2 , $2e^2$, $3e^2$, and so on. This is in accord with the empirical facts since one never measures e by itself. It is always measured in terms of its electromagnetic coupling with other charged matter, in terms of the constant e^2 – the fine structure constant.

What we have seen here, then, is that the origin of the empirically confirmed charge quantization is in the features of the symmetry group of the curved spacetime that prescribes the general covariance of the laws of nature, according to the theory of general relativity.

8. Astrophysics and Cosmology in General Relativity

In the context of the theory of general relativity, the universe as a whole – the subject of cosmology – is not a collection of individual, separable ‘things’, interacting with each other at distances over vast domains of space. Rather, the universe is a continuous whole, with an infinite number of distinguishable manifestations (modes).

From the logic of this theory of matter, we see that the equation of motion of a test body of the universe, say a constituent galaxy, is the geodesic equation in quaternion form. We find that the Hubble law for the ‘expanding universe’ is a first approximation of a covariant law for a spirally evolving universe, alternating between expansion and contraction, i.e., an oscillating universe without any singularities.

The principle of equivalence of the theory of general relativity is discussed in this context, asserting the equivalence between a geometrical path of a test body in the curved spacetime and the exertion of an external force on a body, causing it to move on the same curved path.

The Schwarzschild problem in general relativity is discussed and the effect of the perihelion precession of a planetary orbit is derived in the quaternion formulation of general relativity. It is compared with the standard Einstein result for this effect in his tensor formulation and with the classical formulation for planetary motion.

The spiral structure of the oscillating universe is demonstrated with the solution from the quaternion field theory in terms of Fresnel integrals.

There is a discussion of black holes and pulsars in the context of this theory.

Finally, a scenario is proposed to explain the separation of matter from antimatter in the universe, with a predominance of matter in our domain. It is proposed to be due to the competition between the unidirectional spiral motion of matter and antimatter in the early universe and the effect of the magnetic fields caused by rotating matter and oppositely charged antimatter in the spiral expansion, induced due

to the opposite polarizations of the created cosmic magnetic fields. It is then predicted that in far-away (spacelike) separated regions from our domain, there are predominantly antimatter galaxies and stars, planets, molecules and elementary particles of matter, duplicating the matter distributions in our own sector of the universe.

8.1 Introduction

If there is one set of laws of nature for phenomena in one domain – say the domain of micro-matter of elementary matter (electrons, protons, etc.) – and a different set of laws (not compatible with the first set) for other physical phenomena in another domain – say that of the universe as a whole – then precisely where, in between these two domains of the universe, does one turn one set of laws off and turn the other set on?

If the two sets of laws – such as the quantum theory and the theory of general relativity – are fundamentally incompatible with each other, both conceptually and mathematically, one cannot superpose them in any intermediate domain. The answer to this question, then, is that there cannot be a precise prescription as to where one law would be turned off and the other turned on. It then follows that only one of these laws of nature or the other (or neither!) may be accepted as a true law of nature.

A pertinent example is the change from Newton's theory of universal gravitation, which appears to work well in the terrestrial domain in the vicinity of the Earth, and its interactions with the Moon and the Sun, to Einstein's theory of gravity, represented in general relativity theory, and superceding Newton's theory in the interplanetary and interstellar domains. These two theories of gravity are incompatible, both conceptually and mathematically. Newton's theory is based on the ideas of atomism and action-at-a-distance, and its mathematical basis is linear and Euclidean. Einstein's theory of gravity is based on the continuous field concept and holism, wherein forces propagate at a finite speed between interacting bodies. Its mathematical basis is nonlinear and non-Euclidean. Still, in a particular limit, the formal expression of the laws of gravity, according to Einstein's field theory, asymptotically approaches (but does not exactly incorporate) the formal description of gravity according to Newton's theory.

But this does not mean that Einstein's theory of gravity and Newton's theory of gravity are both true in the domain where the expression of one of these laws mathematically approaches that of the other!

It is rather that, accepting the truth of Einstein's theory, it is useful to use the mathematical expression of Newton's theory (say, for an engineer who is designing a bridge), as an approximation for Einstein's theory, rather than starting from scratch with the formal expression of Einstein's field equations. That is to say, Newton's theory of gravity is descriptive, where it works, but not explanatory. This is an example of the well-known principle of correspondence.

It is in this same sense that the formal expression of 'low energy' quantum phenomena, using quantum mechanics as their description, can only be viewed as a mathematical approximation for a generally relativistic field theory in the micro-domain of matter. The single theory that we examine, then, with physical implications in all domains of nature, from the microscopic domain of 'particle physics' to the domain of cosmology – that of the universe as a whole – is the theory of general relativity. Its ontological stand is different from that of the quantum or classical (Newtonian) theories of matter. In this view, then, the universe is not a collection of separable 'things', interacting with each other or not. Instead, the universe that the scientist investigates is a single holistic continuum, with an infinite variety of (nonseparable, correlated) modes. It is the modes of the universe that appear to us as its constituent things – such as the galaxies, individual stars and planets, people, trees, electrons and protons.

In this final chapter, we will examine aspects of problems of astrophysics and cosmology, in the context of the generalized theory of matter that is based on exploiting the theory of general relativity, as we have developed it thus far.

We will start out with a derivation of the geodesic equation in quaternion form – a generalization of the geodesic equation of Einstein's tensor formalism. Viewing this equation as an equation of motion of a test body in the spacetime whose curvature is determined by all of the 'other' matter of the universe, we will examine specific problems of planetary motion in the astronomical test of general relativity in the quaternion formalism, and the implications of the Hubble law in cosmology, as a first approximation. We will then see how a spiral configuration of the evolving universe as a whole, could follow from the quaternion field equations in general relativity, and how the cosmology appears to be oscillating between expansion and contraction, without any singularities, in contrast to the contemporary model of a single big bang, wherein the entire universe evolves out of a point singularity.

A specific problem in astrophysics that will then be addressed is a possible pulsar model for a star oscillating in and out of the black hole

density state. Finally, there will be discussion of the fascinating problem of the separation of matter and antimatter in the early universe. We will see how this separation may have come about due to the competition of gravitational and electromagnetic forces at the early stages when the expansion (at the beginnings of each cycle) of an oscillating, spiral universe starts to evolve.

8.2 Principle of Equivalence

There has been a great deal of discussion in the literature on the theory of general relativity that pertains to its ‘principle of equivalence’. Before deriving the geodesic equation in quaternion form, it will be instructive to discuss the meaning of this principle and its relation to the geodesic equation as an equation of motion of a test body.

Going back to Newtonian dynamics, his second law of motion is the equation of motion of a test body with mass m , subject to an external force $\mathbf{F}$:

$$\mathbf{r}''(t) = \mathbf{F}/m. \quad (8.1)$$

With the boundary conditions that are the initial values of the position and velocity, $\mathbf{r}(0)$ and $\mathbf{v}(0) = \mathbf{r}'(0)$, a unique trajectory in three dimensions $\mathbf{r}(t)$ follows. (Henceforth we will use the prime symbol to denote the time derivative of a function unless this derivative is explicitly given.)

If there is no external force acting on the test body, $\mathbf{F} = 0$, then the solution of (8.1), $\mathbf{r}''(t) = 0$, is a straight line path, linear in the time parameter: $\mathbf{r}(t) = \mathbf{r}(0) + \mathbf{v}(0)t$. In a two-dimensional plane, eliminating the time parameter t in the equations for $x(t)$ and $y(t)$, the y -coordinate is linear in terms of the x -coordinate, i.e., $y(t) = A + Bx(t)$, where A and B are the integration constants.

The latter straight-line path is a geodesic in the Euclidean space. If there is no external force acting on the test body, it is then predicted (with this theory) to move naturally in a straight line at the constant speed v , forever. This is a statement of Galileo’s principle of inertia, also known as Newton’s first law of motion. Indeed, it would require external energy to do the work that is necessary to move the test body off of this geodesic path.

In the theory of general relativity, the spacetime is non-Euclidean, and its geodesics are curved paths. A continuous family of such geodesics characterizes a curved spacetime.

The extension of Galileo's principle of inertia is then the assertion that the natural path of a body, not acted upon by an external force, is along a curved geodesic. But in a Euclidean spacetime, such a curve would be a consequence of an external force acting on the body. Thus there is an equivalence between, on the one hand, a body moving freely on a curved geodesic of a non-Euclidean spacetime and, on the other hand, a body in a Euclidean spacetime acted upon by an external force. This is a general expression of the principle of equivalence.

In the theory of general relativity, the curvature of spacetime corresponds to the existence of 'other matter' in the universe, apart from the test body. This correspondence is spelled out in the field equations that relate the geometrical properties of the spacetime to the material content of the system, such as Einstein's tensor field equations, or the generalization in terms of the quaternion field equations derived in this text. The Sun's existence then corresponds to a particular curved spacetime. The test body, such as a planet of the solar system, would then move freely on one of the curved geodesics associated with the Sun, that is the orbital path of the planet relative to the Sun's position. This is equivalent to postulating that the planet in a Euclidean space is subject to an external force of the Sun which in turn causes the planet to move on the orbital path that we observe.

However, the equation of motion of the test body along a geodesic in a curved spacetime is only the Newtonian trajectory when Einstein's field equations for gravitation approximate Newton's theory of universal gravitation. As we depart from this approximation, deviations from the Newtonian trajectory are predicted. That is, the 'equivalence' is not exact. Three important deviations predicted and empirically verified, which substantiated Einstein's theory, thereby refuting Newton's theory of gravitation, were:

- the aperiodic motion of a planet about a center of force, i.e., the perihelion precession of a planetary orbit,
- the shift of an emitted radiation spectrum from matter that is in a correspondingly increasing gravitational potential, toward the red end of the spectrum, i.e., the gravitational redshift,
- the curving of a path of starlight as it propagates past the rim of the Sun.

All three of these predictions that are not predicted by Newton's classical theory were determined theoretically and verified empirically in the first half of the 20th century. Thus, Einstein's theory superceded

Newton's theory as an explanation of the phenomenon of gravity. In this chapter, we will analyze the first of these effects in detail.

The 'equivalence principle' is often referred to in saying that the gravitational mass m_G that appears in Newton's equation for gravitational force, and the inertial mass m_i that appears in Newton's second law of motion, are numerically equal to each other. I do not believe that this is an accurate definition, since the gravitational mass m_G is only defined in relation to the Newtonian approximation for the gravitational force while the inertial mass m_i is a general feature of a material system, without approximation.

8.3 The Quaternion Geodesic for a Stationary Orbit

According to Galileo's principle of inertia (Newton's first law of motion) the natural path of a body, that is free of external forces, is a geodesic of the Euclidean space – a straight-line motion. That is, the principle asserts that the natural motion of the body is along a straight line with constant velocity, so long as there are no external forces. The generalization of this principle in Einstein's general relativity is that the natural path of the unobstructed body is a curve in a Riemannian spacetime. The quantity and character of the curvature of this path is predicted by the nature of the 'other matter' of the closed system, of which the test body is a constituent. This relationship follows from Einstein's field equations. In the context of Einstein's tensor formalism, there is no rotation of the body predicted here.

The full generalization of the geodesic path in terms of the irreducible form of general relativity must incorporate a natural rotation of the test body, as well as its translation, as it propagates along the curved geodesic of the Riemannian spacetime. This is in terms of the quaternion metric derived in Chap. 3. For the general continuous change in spacetime, prescribed by the Einstein group which underlies the intrinsic symmetry of the theory of general relativity, includes rotation as well as translation.

As discussed in Chap. 3, the first step in deriving the equation of motion of a test body in the irreducible form is to start with the geodesic equation in its irreducible quaternion expression. This starts with the factorization of the Riemannian metric,

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \longrightarrow ds ds^*, \quad (8.2)$$

where ds and ds^* are the (quaternion) conjugated differential metrics,

$$ds = q_\mu dx^\mu, \quad (8.3)$$

and q_μ is the (16-component) generalized metrical field, derived in Chap. 3, with the correspondence

$$g_{\mu\nu} \Longleftrightarrow q_\mu q_\nu^* + q_\nu q_\mu^*.$$

It is important to emphasize here that the fundamental metrical field is the 16-component quaternion q_μ , rather than the 10-component symmetric tensor field $g_{\mu\nu}$ of Einstein's tensor expression of general relativity.

In the stationary orbit problem in Einstein's tensor formulation, the metric tensor $g_{\mu\nu}$ is time-independent. But the quaternion factorization ds in (8.2) need not be time-independent. It can depend on the time coordinate in a phase factor $q_\mu \rightarrow q_\mu \exp[i\theta(t)]$, so long as the time reversal is $q_\mu^* \exp[i\theta(-t)]$, where the asterisk denotes the quaternion conjugate – its time reversal. Note that the phase factor could be a complex number, so long as its imaginary part is odd under reflections, in space or time. Thus, if $\theta = \theta_r + i\theta_i$ and $\theta_i(-t) = -\theta_i(t)$, then under time reversal $q_\mu \exp(i\theta_r - \theta_i) \rightarrow q_\mu^* \exp(-i\theta_r + \theta_i)$, and the combination $q_\mu q_\nu^* + q_\nu q_\mu^* \Leftrightarrow g_{\mu\nu}$ still represents a time-independent metric tensor, while q_μ itself is time-dependent in the phase factor. (This is similar to the situation in quantum mechanics where the wave function ψ might be time-dependent in a phase factor while the probability density $\psi^* \psi$ remains time-independent.)

We saw in Chap. 3 that if the trajectory is parameterized in terms of the determinant $dS = |q_\mu| dx^\mu$, then the geodesic equation with this parameterization is given in (3.42):

$$\frac{d^2 x^\rho}{dS^2} + \Gamma_{\mu\nu}^\rho \frac{dx^\mu}{dS} \frac{dx^\nu}{dS} = 0. \quad (8.4)$$

This equation is the determinant of the quaternion form of the geodesic equation (3.36).

8.4 Planetary Motion

Let us now commence by defining the interval dS in the reference frame wherein the test body is located at the spatial origin. In this frame,

$$\frac{d}{dS} = |q_0|^{-1} \frac{d}{dX^0} , \quad (8.5a)$$

$$\frac{d^2}{dS^2} = |q_0|^{-2} \frac{d^2}{dX^{02}} + |q_0|^{-1} \frac{d}{dX^0} |q_0|^{-1} \frac{d}{dX^0} , \quad (8.5b)$$

where X^0 is the time measure in the reference frame that is in motion with respect to the global coordinate frame. It follows from (8.5a) and (8.5b) that the geodesic equation (8.4) may be expressed in the form

$$|q_0|^{-2} x^{\rho''} + |q_0|^{-1} (|q_0|')^{-1} x^{\rho'} + \Gamma_{\mu\nu}^{\rho} x^{\mu'} x^{\nu'} = 0 . \quad (8.6)$$

This is equivalent to the equation of motion

$$x^{\rho''} + \Gamma_{\mu\nu}^{\rho} x^{\mu'} x^{\nu'} = -\frac{1}{2} Q^{-1} Q' x^{\rho'} , \quad (8.7)$$

where $Q = |q_0|^{-2}$, and the affine connection has the usual form

$$\Gamma_{\mu\nu}^{\rho} = \frac{1}{2} g^{\rho\lambda} (\partial_{\mu} g_{\lambda\nu} + \partial_{\nu} g_{\mu\lambda} - \partial_{\lambda} g_{\nu\mu}) . \quad (8.8)$$

(For a derivation of this relation, see [117], Sect. 2.14.) The prime in (8.6) and (8.7) denotes differentiation with respect to the time coordinate X^0/c .

The equation of motion (8.7) differs from the usual geodesic equation because of the non-zero term on the right-hand side. But the two equations merge in the local domain, where the Riemannian spacetime approaches a Euclidean spacetime. In this limit, the quaternion metrical field components approach the constant Pauli matrices so that Q then vanishes, and hence the right-hand side of (8.7) also vanishes. The affine connection terms $\Gamma_{\mu\nu}^{\rho}$ vanish in this limit. Thus, the geodesic equation approaches the local equation of motion of classical physics, $x''_{\text{loc}} = 0$. This describes the free motion of a test body, according to Newton's first law of motion.

Since in the case of planetary motion $v/c \ll 1$, it follows that to order v/c the time derivatives of the global coordinates are $x^{0'} = c$, $x^{0''} = 0$. With $\rho = 0$, (8.7) then becomes

$$\Gamma_{\mu\nu}^0 x^{\mu'} x^{\nu'} = -\frac{1}{2} c Q^{-1} Q' . \quad (8.9)$$

The equation of motion in terms of the spatial coordinates ($k = 1, 2, 3$) is then

$$x^{k''} + \Gamma_{\mu\nu}^k x^{\mu'} x^{\nu'} = \Gamma_{\mu\nu}^0 x^{\mu'} x^{\nu'} (x^{k'}/c) . \quad (8.10)$$

Thus, the explicit generalization of the geodesic equation (to order v/c) that follows the quaternionic description of the stationary state in general relativity entails the introduction of the non-zero term on the right-hand side of (8.10) – a term that is non-invariant under time reversal. This result implies that there would be damping in the oscillatory harmonics of planetary motion.

8.5 The Schwarzschild Problem

As discussed in Chap. 2, in the tensor form of general relativity, we assume that, far from a matter distribution such as a single star, Einstein's equations may be approximated by the vacuum equation, $R_{\mu\nu} = 0$. Assuming then, with Schwarzschild, that the tensor solutions are not only stationary but also spherically symmetric, the following squared differential (2.23) was found:

$$ds^2 = \left(1 - \frac{2\alpha}{r}\right) (dx^0)^2 - \frac{dr^2}{1 - 2\alpha/r} - r^2 d\Omega^2, \quad (8.11)$$

where α is an integration constant called the Schwarzschild radius and $d\Omega^2 = \sin^2 \theta d\theta d\phi$ is the (squared) differential solid angle in terms of the polar angle coordinate θ ($0 < \theta < \pi$) and the azimuthal angle coordinate ϕ ($0 < \phi < 2\pi$).

It follows from the metric in (8.11) that the corresponding affine connection components (8.8) are determined. It then follows that the equation of motion (8.10) takes the form (in terms of the spherical polar coordinate system)

$$\begin{aligned} \frac{2\alpha}{r^2} \left(1 - \frac{2\alpha}{r}\right)^{-1} (r')^2 = r'' - \frac{\alpha}{r^2} \left(1 - \frac{2\alpha}{r}\right)^{-1} (r')^2 \\ + \left(1 - \frac{2\alpha}{r}\right) \left[\frac{\alpha c^2}{r^2} - r(\theta')^2 - r \sin^2 \theta (\phi')^2 \right], \end{aligned} \quad (8.12a)$$

$$\frac{2\alpha}{r^2} \left(1 - \frac{2\alpha}{r}\right)^{-1} r' \theta' = \theta'' + \frac{2}{r} \theta' r' - \sin \theta \cos \theta (\phi')^2, \quad (8.12b)$$

$$\frac{2\alpha}{r^2} \left(1 - \frac{2\alpha}{r}\right)^{-1} r' \phi' = \phi'' + 2 \cot \theta \theta' \phi' + \frac{2}{r} r' \phi'. \quad (8.12c)$$

In the problem of planetary motion, r is the distance from the center of mass of the Sun to that of the orbiting planet, ϕ is the azimuthal angle swept out by the radius vector in the orbital plane, and the polar

angle θ measures the angle of orientation of the orbital plane relative to the observer's reference frame.

Choosing the boundary condition so that, at the initial time $t = 0$, the orbital plane is at $\theta_0 = \pi/2$, it follows that the solution of (8.12b) is the orbital plane with this fixed orientation, $\theta(t) = \pi/2$, as in classical mechanics. With this solution for θ , the three coupled differential equations (8.12a–c) reduce to the following two coupled differential equations:

$$0 = r'' + \frac{\alpha c^2}{r^2} \left(1 - \frac{2\alpha}{r}\right) - \frac{3\alpha}{r^2} \left(1 - \frac{2\alpha}{r}\right)^{-1} (r')^2 - r \left(1 - \frac{2\alpha}{r}\right) (\phi')^2, \quad (8.13a)$$

$$\frac{d(r^2\phi')}{dt} = 2\alpha \left(1 - \frac{2\alpha}{r}\right)^{-1} r' \phi'. \quad (8.13b)$$

The corresponding equations in the Schwarzschild formulation are the same as above, except that there is a 1 rather than a 3 in the second term of (8.13a) and the term on the right-hand side of (8.13b) is zero.

To first order in α/r , (8.13b) becomes

$$\frac{d(r^2\phi')}{dt} = 2\alpha r' \phi'. \quad (8.14)$$

Using the change of variables, $y = r^2\phi'$, (8.14) takes the form

$$\frac{dy}{dt} = 2\alpha r' \frac{y}{r^2}. \quad (8.15)$$

The solution of this equation is

$$y = 2\alpha \int \frac{yr'}{r^2} dt = 2\alpha \int \frac{y}{r^2} dr.$$

Thus,

$$\frac{dy}{dr} = \frac{2\alpha y}{r^2}.$$

The solution of this equation is

$$y \equiv r^2\phi' = K \exp\left(-\frac{2\alpha}{r}\right), \quad (8.16)$$

where K is an integration constant. For $r \gg 2\alpha$, $\exp(-2\alpha/r) \rightarrow 1$ and $r^2\phi' \rightarrow K$. Thus,

$$K = r_C^2 \phi'_C = L/m \quad (8.17)$$

is Kepler's constant for the planet, where L is its angular momentum and m is its inertial mass. In the classical theory, the orbital speed of the planet is

$$v_{\text{orb}} = r_C \phi'_C = K/r. \quad (8.18)$$

In Einstein's tensor formalism for general relativity, $ds = cdt(1 - 2\alpha/r)^{1/2}$, so that the corresponding Einstein constant is

$$K_E = cr^2 \frac{d\phi}{ds} = cr^2 \phi' \left(\frac{ds}{dt} \right)^{-1} = \frac{L}{m} \left(1 - \frac{2\alpha}{r} \right)^{-1/2}. \quad (8.19)$$

Thus, in Einstein's tensor formulation of the problem of planetary motion, in the Schwarzschild approximation, the planet's angular momentum and orbital speed are

$$L = mK_E \left(1 - \frac{2\alpha}{r} \right)^{1/2}, \quad (8.20a)$$

$$v_{\text{orb}} = \frac{K_E}{r} \left(1 - \frac{2\alpha}{r} \right)^{1/2}. \quad (8.20b)$$

In the quaternionic expression of general relativity, with the approximations used, the angular momentum and orbital speed of the planet are

$$L = mK \exp \left(-\frac{2\alpha}{r} \right), \quad (8.21a)$$

$$v_{\text{orb}} = r\phi' = \frac{K}{r} \exp \left(-\frac{2\alpha}{r} \right), \quad (8.21b)$$

where K is Kepler's constant.

When the radial distance from the center of mass of the planet to that of the Sun is sufficiently large, then as r decreases, the orbital speed increases as $1/r$. But in the quaternionic theory, the increase in v_{orb} slows down as $r \rightarrow 2\alpha$. According to (8.21b), the orbital speed approaches zero at the origin, rather than diverging to infinity, as the classical theory shows according to (8.18).

In Einstein's formulation, v_{orb} , as a real number, must cut off at $r = 2\alpha$. This is a discontinuity that does not occur in the quaternion formulation, which instead predicts that $v_{\text{orb}} \rightarrow 0$ in this limit, since $\exp(-2\alpha/r)$ approaches zero faster than $1/r$ approaches infinity as r approaches zero.

8.6 The Radial Solution and Perihelion Precession

The perihelion precession of a planetary orbit follows from the radial solution as a function of the azimuthal angle swept out in the orbital plane. Substituting (8.16) into (8.13a) and expressing the derivatives with respect to the azimuthal angle ϕ rather than the time, the equation of motion in the radial variable takes the form

$$\frac{d^2 r}{d\phi^2} - 2 \frac{dr}{d\phi} \left(1 + \frac{\alpha}{2r}\right) - r \left(1 - \frac{2\alpha}{r}\right) + \frac{\alpha c^2 r^2}{K^2} \left(1 + \frac{2\alpha}{r}\right) = 0. \quad (8.22)$$

With the change of variables $u(\phi) = 1/r(\phi)$, (8.22) may be expressed in the form

$$\frac{d^2 u}{d\phi^2} + u = H + 2\alpha u^2 + 2\alpha H u - \alpha \left(\frac{du}{d\phi}\right)^2, \quad (8.23)$$

where the constant $H = \alpha c^2/K^2$.

Assuming the method of perturbations ([117], Sect. 7.6), the solution $u(\phi)$ of the radial equation (8.23) is

$$u(\phi) = \left(H + 4\alpha H^2 + \frac{1}{2}a_0^2\alpha\right) + \left(a_0 \cos \phi - \frac{1}{2}a_0^2\alpha \cos 2\phi\right) + 3a_0 H \alpha \phi \sin \phi, \quad (8.24)$$

where $a_0 = (1 - r_p H)/r_p$, and r_p is the distance from the center of mass of the planet at the perihelion point (the point of closest approach to the Sun on the planet's elliptical orbit) to that of the Sun.

The first two terms in (8.24) are periodic, but the third term is not. It is the latter that predicts that the planet will take a longer time in each of its orbits to reach the perihelion point. With the approximation that for small p ,

$$\cos(\phi - p\phi) \approx \cos \phi + p\phi \sin \phi,$$

equation (8.24) then becomes

$$u(\phi) = H + 4\alpha H^2 + \frac{1}{2}a_0^2\alpha - \frac{1}{2}a_0^2\alpha \cos 2\phi + a_0 \cos[(1 - 3\alpha H)\phi]. \quad (8.25)$$

The maximum value of $u(\phi)$ occurs at

$$(1 - 3\alpha H)\phi_n = 2\pi n,$$

where n is an integer that signifies the orbit of the planet. The perihelion shift for each orbit from the preceding orbit is

$$\Delta = \phi_{n+1} - \phi_n - 2\pi = 2\pi(1 - 3\alpha H)^{-1} - 2\pi \approx 6\pi H\alpha .$$

With $\alpha = GM/c^2$, $H = \alpha c^2/K^2$ (G is Newton's universal gravitational constant), we find the result for the perihelion precession:

$$\Delta = \frac{6\pi G^2 M^2}{c^2 K^2} . \quad (8.26)$$

This is the same result as Einstein's. When M is the mass of the Sun and $K = L/m$ is Kepler's constant for the planet Mercury, $\Delta \approx 42.86$ arcsec per century. The observation for this is $\Delta_{\text{exp}} \approx 42.6$ arcsec per century.

It should be noted that in higher approximations, the quaternion field prediction and Einstein's tensor prediction would start to deviate from each other. This is something that could be pursued in future astronomical studies, experimentally and theoretically.

8.7 The Hubble Law and Cosmology

Newton's theory of gravitation predicted that the force of gravitation between massive bodies depends on their mutual separation, but not on the time parameter. This led to the prediction that all massive bodies move in stationary orbits about other massive bodies, such as the planets of our solar system relative to their mother Sun. However, Einstein's theory of general relativity superceded Newton's theory as an explanation for gravitation. It predicts that the force of gravity is a field that depends on the time parameter as well as the spatial parameters. In principle, then, Einstein's theory of general relativity ruled out the dynamics of stationary orbits, as in Newton's theory of universal gravitation.

We saw in the preceding section that the planetary orbit is, in Einstein, non-periodic – the anomalous perihelion precession of a planetary orbit was calculated in the case of Mercury's orbit to be in close agreement with the empirical data. Applied to stellar systems of the night sky, the implication of general relativity theory is that individual stars, galaxies, clusters of galaxies, and so on, do not orbit about centers of force, in stationary orbits, as Newton's theory would have anticipated.

In the 1920s, E. Hubble discovered that Einstein's prediction was indeed correct. He saw that each galaxy of the night sky is moving away from all other galaxies – that the universe is 'expanding'. Specifically,

what Hubble discovered was that the speed v of a galaxy relative to another galaxy, that is r cm away from it, depends linearly on their separation, i.e.,

$$v = Hr . \quad (8.27)$$

This is called the Hubble law and H is Hubble's constant. The observation was based on the Doppler effect, whereby the observed wavelength of radiation from a moving source is displaced toward the red end of the spectrum (the longer wavelengths) when the emitter of this radiation moves away from the observer. Thus, Hubble saw a predominant redshift of the radiation emitted by the galaxies of the universe, thus leading to the conclusion that the universe is expanding, rather than being a set of stationary orbits of stellar matter relative to other stellar matter – the cosmic clock that was envisioned by Newtonian theory.

In the context of the theory of general relativity, the Hubble law (8.27) is not relativistically covariant – it changes its form when transformed to other spacetime reference frames. Thus, the theory of relativity implies that this empirically correct law (8.27) must be a mathematical approximation for a truly covariant law of stellar dynamics of the constituents of the universe as a whole. Theoretically, then, we must start with a cosmological model that is compatible with the theory of general relativity – a theory of relativistic cosmology.

The presently accepted single big bang cosmology assumes a universal time measure, called cosmological time, which starts at the initiation of the assumed single expansion of the universe. Thus, all time measures, in this view, may be with reference to the beginning of this 'absolute time'. But such a view is incompatible with the relativity of time measure, according to the basis of the theory of relativity – the principle of relativity. In the latter view, all time measures depend on the frame of reference in which they are defined. To counter this assertion by saying that the reference frame of the universe as a whole is indeed absolute does not refute Einstein's requirement – that the expression of the laws of nature cannot depend on any absolute reference frame, not even that of the universe as a whole! Thus, a feature of the single big bang cosmology that makes it incompatible with a theory of relativistic cosmology is its incorporation of an absolute cosmological time measure.

8.7.1 The Oscillating Universe Cosmology

A way out of this difficulty is to recognize that an oscillating universe cosmology is indeed compatible with the requirement of relativistic

covariance. The terms in the tensor form of the geodesic equation that play the role of force are the affine connection components $\Gamma_{\mu\nu}^{\rho}$. What is important here is that these are not positive-definite fields. Under conditions of sufficiently dense matter and relative velocities comparable with c , between interacting matter, the forces between matter can be repulsive, dominating the intrinsic attractive forces. This would then give rise to an expansion of the matter of the universe. But when the matter becomes sufficiently rarefied, an inflection point is reached when the attractive forces dominate the repulsive forces, thus changing the global expansion of the universe into a global contraction. The contraction continues until the matter density and relative speeds of its components become large enough, when the matter of the universe is sufficiently dense, to reach another inflection point where the predominant attractive force changes once again to a predominant repulsive force, starting the expansion once again. This scenario continues ad infinitum. It is the oscillating universe cosmology. With this cosmology, there is no scientific answer to the question: when did the universe begin?

According to this model, there is no absolute time measure or a unique singular ‘big bang’ to start off the expansion of the universe. The time when the last big bang happened depends on the reference frame from which it was determined. In our reference frame, here on Earth, our astronomers say that, according to the Doppler measurements and their fit to the Hubble law (8.27), the determination of the constant H implies that the ‘big bang’ happened between 10 and 15 billion years ago. But from a different reference frame, say on some other planet in some other distant galaxy that is moving at a high speed relative to us, the determination might be 1 billion years ago, or from another reference frame, it might be 300 billion years ago.

Further, as distinct from the single big bang model, the beginning of any particular cycle (big bang) is not discrete or mathematically singular in space and time. The lack of singularities in physics is indeed a requirement of the theory of general relativity, as a field theory based entirely on the continuous field concept.

8.7.2 Dynamics of the Expansion and Contraction of the Universe

With the assumption that the terms $\Gamma_{\mu\nu}^k x^{\mu'} x^{\nu'}$ in the equation of motion (8.10) are time independent (over the time scale of our observations), this equation may be integrated, yielding the solution

$$x^k(t) = K_1 + K_2 \exp(c^{-1} \Gamma_{\mu\nu}^0 x^{\mu'} x^{\nu'} t) + \frac{\Gamma_{\mu\nu}^k x^{\mu'} x^{\nu'}}{\Gamma_{\mu\nu}^0 x^{\mu'} x^{\nu'}} ct, \quad (8.28)$$

where $k = 1, 2, 3$ are the three spatial directions and K_1, K_2 are the two integration constants, to be determined by the boundary conditions.

We will now assume the validity of the oscillating universe cosmology, for the reasons of covariance discussed above. Since the velocity of the test matter in motion changes its direction from the imploding stage to the exploding stage of the universe as a whole, there is an inflection point in dx^k/dt at the times of alternation of these phases. Calling $t = 0$ the time when the presently observed explosion started in our cycle of the oscillating universe, and locating the test matter at the origin, the boundary conditions imposed by this cosmology are

$$\frac{dx^k}{dt}(0) = 0 = x^k(0). \quad (8.29)$$

With (8.29) in the solution (8.28), the integration constants are found to be

$$K_1 = -K_2 = c^2 \left[\frac{\Gamma_{\mu\nu}^k x^{\mu'} x^{\nu'}}{(\Gamma_{\mu\nu}^0 x^{\mu'} x^{\nu'})^2} \right]_0, \quad (8.30)$$

where 0 refers to the value of the ratio in the bracket at the beginning of a cycle of the oscillating universe.

With the assumption that $dx^k/dt \ll c$, (8.30) becomes

$$K_1 = -K_2 = \left[\frac{\Gamma_{00}^k}{(\Gamma_{00}^0)^2} \right]_0,$$

and the solution (8.28) takes the form

$$x^k(t) = \left[\frac{\Gamma_{00}^k}{(\Gamma_{00}^0)^2} \right]_0 [1 - \exp(c^{-1} \Gamma_{\mu\nu}^0 x^{\mu'} x^{\nu'} t)] + \frac{\Gamma_{\mu\nu}^k x^{\mu'} x^{\nu'}}{\Gamma_{\mu\nu}^0 x^{\mu'} x^{\nu'}} ct. \quad (8.31)$$

Consider the two terms on the right-hand side of (8.31). The coefficient that multiplies the first term has the dimension of length:

$$\left[\frac{\Gamma_{00}^k}{(\Gamma_{00}^0)^2} \right]_0 \equiv R^k. \quad (8.32)$$

This may be interpreted as the radius of the universe when applied to the furthestmost stellar objects of the night sky. When R^k is large and spacelike, compared with ct , the second term on the right-hand side of

(8.31) may be neglected compared with the first. This approximation corresponds to the assumption that a distance from the observer to the faraway galaxy is large compared with the distance traveled by light from one time that the astronomer observes this galaxy to the next.

With the latter assumption, the resulting solution predicts that the speed of a galaxy, in any of the three spatial directions x^k , is

$$\frac{dx^k}{dt} = c^{-1} \Gamma_{\mu\nu}^0 x^{\mu'} x^{\nu'} (x^k - R^k). \quad (8.33)$$

Since R^k cancels in the comparison of two speeds at the corresponding times, (8.33) then predicts that dx^k/dt is linearly proportional to x^k . (This assumes that the term in front of the right-hand side of (8.33) is a constant in time. The latter term is then the Hubble constant,

$$H = c^{-1} \Gamma_{\mu\nu}^0 x^{\mu'} x^{\nu'} = -\frac{1}{2} Q^{-1} Q'. \quad (8.34)$$

With this result, (8.33) is indeed a derivation of the (empirically correct) Hubble law. The quaternion factor $Q = |q_0|^{-2}$ is defined in (8.7), dependent on the determinant of the time component of the quaternion metrical field q_0 . With the definition (8.34) of Hubble's constant H , we see that it actually is dependent on the time parameter. But with the approximation used, it appears to be a constant over the time span that it is observed in our view of the expanding universe. In principle, however, and independent of approximation, H must be space and time dependent, and thus the Hubble law (8.27) must only be a non-relativistic approximation to describe the expanding universe, because of the covariance of the laws of nature, including the laws of cosmology.

8.7.3 Spiral Structure of the Universe

Let us consider further the dynamics developed thus far for the oscillatory cosmological model. With the assumption of the boundary conditions (8.29) applied to the constituent 'test matter' of the universe, subject to the forces exerted on it by the rest of the matter of the universe, we will assume that the 'radius' of the system (8.32) is the same order of magnitude as the distance ct . Thus we may not in this case neglect the second term on the right-hand side of (8.31), compared with the first term.

In this case, however, we may consider the motion of a test mass over time intervals that are short compared with the time it would take light to traverse the entire domain of the universe. Under these conditions, and assuming that the speed of the test matter relative to the ‘center of mass’ of the entire universe is small compared with c , the argument of the exponential factor in (8.31) is small compared with unity, so that

$$c^{-1} \Gamma_{\mu\nu}^0 x^{\mu'} x^{\nu'} t \approx \Gamma_{00}^0(ct) \approx \frac{ct}{R} \ll 1 ,$$

where R is the order of magnitude of the interaction domain radius of the entire universe, according to (8.32).

With this approximation, the expansion of the exponential in (8.32), keeping the first two terms, gives the location of the moving matter as follows:

$$x^k(t) = ct \left\{ \frac{\Gamma_{00}^k}{(\Gamma_{00}^0)^2} - \left[\frac{\Gamma_{00}^k}{(\Gamma_{00}^0)^2} \right]_0 \right\} \Gamma_{00}^0 . \quad (8.35)$$

With the general expression for the affine connection (8.8), and a Taylor expansion of $g_{\mu\nu}(t)$, assuming it to be an analytic function of the time parameter t , it follows that for small t the spatial coordinates of the test mass are given by the relation

$$x^k(t) = \frac{1}{2} a^k t^2 + b^k t^3 ,$$

where a^k and b^k are complicated functions of the first and second derivatives of the metric tensor $g_{\mu\nu}$ with respect to the time coordinate. (For details, see [117], Sect. 7.8.)

With this result, the acceleration of the test body in the k th spatial direction is

$$\frac{d^2 x^k}{dt^2} = a^k + 6b^k t . \quad (8.36)$$

Substituting the variables $d^2 \zeta^k / dt^2 = d^2 x^k / dt^2 - a^k$, (8.36) takes the form

$$\left(\frac{d^2 \zeta^1}{dt^2} + \frac{d^2 \zeta^2}{dt^2} + \frac{d^2 \zeta^3}{dt^2} \right)^2 = 36 [(b^1)^2 + (b^2)^2 + (b^3)^2] t^2 . \quad (8.37)$$

The quaternion metrical field q_μ is a spin-one variable, implying that in addition to translational motion, there is a rotational motion of the test

body that moves on the quaternionic geodesic. Thus, the entire closed system (in principle the ‘universe’) must be set into rotational motion in a plane that is perpendicular to the orientation of the imposed spin of the system as a whole. Thus if x^3 is (defined to be) the direction of the axis of rotation of the closed system, the quaternion field equations must predict that $(b^1, b^2) \gg b^3$. We then formulate a (non-isotropic) two-dimensional displacement of the test mass (a constituent of the early universe) in terms of the vector

$$\zeta^r = e_x \zeta^x + e_y \zeta^y .$$

This is in reference to the arbitrarily chosen origin $x = 0, y = 0$ that defines the initial motion of the test mass. The nonlinear differential equation in terms of this coordinate is then

$$\left(\frac{d^2 \zeta^r}{dt^2} \right)^2 = \left(\frac{d^2 \zeta^x}{dt^2} \right)^2 + \left(\frac{d^2 \zeta^y}{dt^2} \right)^2 = A^2 t^2 , \quad (8.38)$$

where

$$A^2 = 36 [(b^x)^2 + (b^y)^2] . \quad (8.39)$$

Clearly, the coordinate of the test mass must obey the boundary conditions of the oscillating cosmology, $\zeta^r(0) = (d\zeta^r/dt)(0) = 0$. With these boundary conditions, the solutions of (8.38) are the Fresnel integrals that define the Cornu spiral:

$$\zeta^x = c \left\{ \int_0^t \cos [(A/2c)\tau^2] d\tau - t \right\} , \quad \zeta^y = c \int_0^t \sin [(A/2c)\tau^2] d\tau . \quad (8.40)$$

The total solution for the ‘test body’ in the universe in this ‘early time’ is then

$$x(t) = \zeta^x(t) + \frac{1}{2} a^x t^2 , \quad y(t) = \zeta^y(t) + \frac{1}{2} a^y t^2 . \quad (8.41)$$

It is readily verified with this solution that $x(0) = (dx/dt)(0) = y(0) = (dy/dt)(0) = 0$, in accordance with the boundary conditions of the oscillating universe cosmology.

The solution (8.41) is a superposition of the spiral motion of a test body of the universe in a two dimensional plane, characterized by Fresnel integrals, and a constant acceleration relative to an observer’s frame of reference. In the rest frame of the test body itself, then, $\mathbf{a} = 0$.

One then has a purely spiral motion with two inflection points – one at the beginning of the expansion, when there is a maximum density of matter and the dominant gravitational force is repulsive, and the other at minimum matter density, when the dominant gravitational force is attractive, leading to the beginning of the contraction phase of the cycle. These oscillations of the matter of the universe then continue indefinitely, into the past and into the future.

This cosmological model is then in contrast with the ‘single big bang’ model of present day thinking, wherein there is a singular beginning of all matter in space and time, and then a single expansion (with the possibility of a subsequent single contraction). With the latter cosmological model, the matter of the universe at the ‘beginning’ was isotropic and homogeneous, and continues to be so. The observations of the galaxies and the clustering of galaxies of the night sky are convincing to me that the latter model is not true to nature, nor is its assumption of an absolute, cosmological time. These do not conform with the covariance requirement of the theory of general relativity, as a basis for cosmology.

One further remark is that the same boundary conditions and approximations applied to the universe as a whole, as discussed above, could apply to individual galaxies. The same predictions must then follow, that a constituent star of a galaxy evolves along a spiral path in the confines of its mother galaxy, with two inflection points to describe an oscillating galaxy of stars, that is continually expanding and contracting. The empirical observation of the spiral structures of most of the observed galaxies may attest to this theoretical astrophysical prediction.

8.8 Black Holes and Pulsars

As we have seen earlier in this text, general relativity theory entails the curvature of spacetime as a representation of the existence of matter. The denser the matter of a star, the greater is the curvature of spacetime in its vicinity.

Thus, starlight propagating past the rim of a star, in moving along a geodesic of spacetime, will trace a curved path in this vicinity. This effect has been seen as one of the tests of general relativity – the bending of a beam of starlight as it passes the rim of the Sun. (The qualitative and quantitative effect is discussed in detail in [117], Sect. 6.11.) The effect is also seen in contemporary measurements of the lens effect,

whereby the astronomers view a double image of a star, due to beams of starlight that pass the opposite edges of a scattering stellar object. The amount of curvature viewed there depends on the mass density of the scatterer.

Imagine now a star that is so dense that the geodesics associated with it are so highly curved that they close on themselves. The family of such a set of geodesics is similar to the irrotational field of a magnet. One sees all of the lines of force emanating from the magnetic north pole bending around so as to arrive at the magnetic south pole, no matter how closely the north and south poles are located.

If the density of a star is so high that all of the geodesics associated with it close on themselves, then all signals emitted by such a star must be re-absorbed by the same star. Thus, all electromagnetic radiation (and any other signals) leaving this star must be reabsorbed by the same star. These signals could not arrive at our detectors – the star would be ‘black’ to us. Such a star is called a black hole. This scenario is contingent on the stability of such a star, which, in turn, depends on the existence of stable solutions of the metrical field equations in general relativity.

According to the foregoing definition of a black hole, all signals emitted by such a star – all virtual fields that would couple the star to the outside world, such as ‘photons’, neutrinos, or the field of a propagating gravitational force (which in contemporary physics is called a graviton) – that make its presence known, would never emerge, because they propagate along the geodesics associated with the black hole, and these are a field of lines that close on themselves. All such emitted signals from a black hole must then be re-absorbed by this star.

It should be pointed out here that, in this day, there is no empirical, conclusive evidence for the existence of the black hole, nor is there any theoretical proof that there are stable solutions of the equations of general relativity that correspond to the black hole state. But, in principle, it is still possible that black holes could exist in reality. In this regard, there has been a great deal in the literature on observational discoveries of black holes in the night sky. But this cannot yet be a reality, for the reasons discussed above. There is indeed astronomical evidence for very dense matter at the centers of galaxies, but there is no proof that these are indeed black holes as they are defined above! There is also discussion of the possibility that some observations can be associated with binary black holes – two of them that rotate about each other under the force of gravity. But this must also be ruled

out as a reality since a single black hole cannot couple to any other matter, including another black hole – since the latter would require its emitted gravitational force to be absorbed by another body, and this is not possible because these forces are propagating along closed geodesics! Earlier claims have been based on observing X-rays emitted by a visible star and then disappearing from view. The claim is that they are absorbed by a black hole that is binary to the visible star. But this is also not possible because the black hole cannot bind to other matter.

This author also does not believe that it is possible to express the nature of the black hole with quantum mechanical concepts – the Heisenberg uncertainty principle, quantum mechanical tunneling, etc. The reason is that it has never been shown that general relativity (the theoretical root of the black hole, if it exists) can be unified with quantum mechanics – it has never been demonstrated that there is a rigorous quantum theory of gravity! Indeed, there is reason to believe that such a unification is impossible, for both mathematical and conceptual reasons.

8.8.1 Possible Model of a Pulsar

The dynamics of all material systems – the universe as a whole, individual galaxies, individual stars, etc. – is oscillatory, according to the theoretical analysis of this monograph. Thus, the normal life of a star is one of pulsation between minimum and maximum density.

If a black hole star is a possibility, then while it is in the state of maximum density, no light (or other signals) may be emitted from it to the outside world. But if this star pulsates, then when it is in a density state that is lower than the black hole density, it will emit radiation to the outside world. Thus, our observations of such a star would reveal (damped) periodic pulses of radiation, just as it is seen in the night sky, interpreted as ‘pulsars’. Taking the time-irreversible equation of motion of the test body, subject to a host star, to be (8.10), the expression for the affine connection in terms of the metric tensor (8.8), and spherical symmetry, so that $g^{\mu\nu} = g^{\mu\nu}(r, x^0)$, it follows that the equation of motion satisfied by the radial coordinate of the test matter is as follows [115]:

$$\frac{d^2 r}{dt^2} + c \left(g^{rr} \partial_0 g_{rr} - \frac{1}{2} g^{00} \partial_0 g_{00} \right) \frac{dr}{dt} - g^{00} \partial_r g_{00} \left(\frac{dr}{dt} \right)^2 - \frac{c^2}{2} g^{rr} \partial_r g_{00} = 0. \quad (8.42)$$

Assuming now that the speed of the test body in the star is small compared with the speed of light, i.e., $(dr/dt)/c \ll 1$, the third term in (8.42) may be neglected compared with the second term. The resulting equation describes a damped oscillator. We then define terms as follows:

$$\partial_r g_{00} = -2 \left(\frac{\omega}{c} \right)^2 r g_{rr} , \quad (8.43)$$

$$c \left(g^{rr} \partial_0 g_{rr} - \frac{1}{2} g^{00} \partial_0 g_{00} \right) = b , \quad (8.44)$$

where ω is the angular frequency of the pulsation of the star and b is a constant parameter characteristic of the star. Inserting (8.43) and (8.44) into (8.42), the latter takes the form of the equation of motion of a damped oscillator:

$$\frac{d^2 r}{dt^2} + b \frac{dr}{dt} + \omega^2 r = 0 . \quad (8.45)$$

The general form of the solution of this equation is

$$r(t) = A \exp(-bt/2) \cos \{ [\omega^2 - (b/2)^2]^{1/2} t + \alpha \} , \quad (8.46)$$

where A, α are the two integration constants. They are determined from the boundary conditions

$$r(0) = A \cos \alpha = r_0 , \quad (8.47)$$

$$\frac{dr}{dt}(0) = - \left[(A^2 - r_0^2) [\omega^2 - (b/2)^2] \right]^{1/2} - (b/2) r_0 = 0 , \quad (8.48)$$

where r_0 is the gravitational radius of the star at the beginning of its pulsations. The boundary condition (8.48) follows because of the oscillatory nature of the pulsations of the star – at the inflection points, the rate of change of the star's radius with respect to time must vanish.

Applying (8.47) and (8.48) to the solution (8.46), we have

$$r(t) = \frac{r_0 \omega \exp(-bt/2)}{[\omega^2 - (b/2)^2]^{1/2}} \cos [\omega^2 - (b/2)^2]^{1/2} t + \alpha , \quad (8.49)$$

where

$$\alpha = \cos^{-1} \frac{[\omega^2 - (b/2)^2]^{1/2}}{\omega} . \quad (8.50)$$

8.8.2 Damped Oscillatory Motion and Pulsars

A solution of (8.43) is

$$g_{00} = -2 \left(\frac{\omega}{c} \right)^2 \int g_{rr} r dr, \quad (8.51)$$

yielding

$$\partial_0 g_{00} = 2 \left(\frac{\omega}{c} \right)^2 \int \partial_0 g_{rr} r dr. \quad (8.52)$$

Inserting this into (8.44) we have

$$b = c \left[g^{rr} \partial_0 g_{rr} + \left(\frac{\omega}{c} \right)^2 g^{00} \int \partial_0 g_{rr} r dr \right]. \quad (8.53)$$

Calling $\partial_0 g_{rr} = F(r, x^0)$, (8.53) may be expressed as follows:

$$\frac{F(r, x^0)}{\int F(r, x^0) dx^0} - \frac{1}{2} \frac{\int F(r, x^0) dr}{\int F(r, x^0) r dr dx^0} = \frac{b}{c}. \quad (8.54)$$

Assuming that the metrical field of the pulsating star describes a stationary state with angular frequency ω , F is separable as follows: $F(r, x^0) = f(r) \cos \omega t$. Equation (8.54) then reduces to

$$\omega = 2b \tan \omega t. \quad (8.55)$$

The solutions of (8.55) occur at the times when $\omega/2$ crosses the transcendental function $\tan \omega t$. The first crossing is at

$$\frac{\omega}{2b} = \omega t_1 = 0.1344.$$

This solution is unacceptable because it yields an imaginary frequency in the cosine function in (8.47), implying that $r(t)$ is not oscillatory, in contradiction with our starting assumption.

The next crossing in (8.55) occurs at

$$\frac{\omega}{2b} = \omega t_2 = 4.4932.$$

This gives $b = \omega/8.9864 \approx \omega/9$. In this case, the solution (8.47) has the form

$$r(t) = r_0 \exp \left(-\frac{0.349t}{T} \right) \cos \frac{2\pi t}{T}, \quad (8.56)$$

where $T = 2\pi/\omega$ is the period of the star's oscillation. According to this solution, the gravitational radius of the star would decrease to e^{-1} of its original size after the order of three pulsations.

The next crossing in (8.55) occurs at a time much greater than $3T$. Thus, within this analysis, the solution (8.56) determines the damping action in the star's pulsations.

Summing up, the prediction here is that, if there can indeed be a stable black hole solution of the field equations of general relativity, its pulsations in and out of the black hole density state would reveal a 'pulsar' effect, whereby one would observe pulses of radiation, whenever the star is out of the black hole density state. The intensity of the pulses would then dampen in accordance with the ever-decreasing size of the star to e^{-1} of its original radius in the order of three pulsations.

This mechanism is not the only possible one to reveal pulsar action. But in my view it is a possible candidate to explain (at least a portion of) the pulsar data. The commonly held model of a pulsar in contemporary astrophysics is a rotating neutron star.

In the limit that corresponds to damping the oscillations of the black hole toward the end of its life, the portion of its life corresponding to closed geodesics (the black hole portion) would change to only open geodesics. This would lead to a total disintegration of the star.

The evolution of this (very specialized type of) star would be as follows: supernova (or other violent stellar processes) that could lead to a maximal density star $\rightarrow$ black hole $\rightarrow$ pulsar ($\equiv n$ cycles of pulsation in and out of the black hole state) $\rightarrow$ total disintegration of the star. The time scale for this set of steps would depend on the physical parameters of the star and its rate of damping (attenuation). Generally, such a star would be a very rare object in the normal constituency of the night sky.

8.9 Separation of Matter and Antimatter in the Early Universe

An interesting question in cosmology and particle physics, that has persisted since the discovery of antimatter in the early decades of the 20th century, is: why is there a predominance of matter over antimatter in our domain of the universe? A second question is: is there another domain of the universe where there is a predominance of antimatter over matter, where the complex structures such as atoms, molecules, people, planets, galaxies, etc., are composed of antimatter rather than

matter? In the latter case, the dominant atoms would be anti-helium, rather than helium, whose nuclei are two antiprotons bound to two antineutrons, with two orbiting antielectrons (positrons).

I wish to propose the following scenario that could answer these questions. We start with the beginning of a spiral expansion, at the beginning of any particular cycle of the oscillating universe, as discussed earlier in this chapter. The spiral nature of the matter of the expanding and contracting universe followed mathematically from solutions in the equations of motion of a test body that is a constituent of the universe, subject to the forces of the rest of the universe. These are the Fresnel integrals (8.40) and (8.41). Because of the underlying quaternion structure of the metric of spacetime, the cosmological dynamics is found to be asymmetric, involving rotation, say in the clockwise direction relative to a universal cosmic plane.

Note that while we have specified ‘clockwise rotation’, this is tied to an arbitrarily chosen observer’s view. There is no absolute cosmic reference frame in the generally covariant theory of relativity. That is, from a different observer’s frame of reference, the cosmic rotation could be seen as counterclockwise.

The theory developed thus far reveals that in the background of all observed matter, there is a sea of electron and proton particle–antiparticle pairs, e^-e^+ and p^-p^+ , each in their ground states of null energy, null momentum and null angular momentum. In regard to energy, this means that the pair is at $2mc^2$ below the state where they would be free of each other, where m is the inertial mass of each of the members of the pair. The early universe, at the time of an explosion of its matter into an expansion phase (the phase that we witness at the present time) is then composed only of such a ‘physical vacuum’ of pairs.

The scenario for the separation of the particles from their bound antiparticles, at the time of the gravitational explosion, is as follows. The internal energy of the matter of the universe, at the inflection point wherein the explosion (‘big bang’) ensues, delivers 2 Gev units of energy to each of many of the proton pairs, so as to dissociate them, out of a much greater number of proton pairs of the universe, and about 1 Mev units of energy to each of the electron pairs to dissociate them.

The positively charged protons and positrons then move in the spiral gravitational field, say clockwise, giving rise to a magnetic field B^+ , perpendicular and below the cosmic plane of rotation. The dissociated, negatively charged electrons and antiprotons then rotate in the

opposite direction, counterclockwise, in the magnetic field B^+ , thus separating them from the positively charged positrons and protons.

The antiprotons and electrons, rotating in the clockwise direction in the spiral gravitational field, then give rise to a magnetic field B^- , perpendicularly oriented above the plane of cosmic rotation. This magnetic field, in turn, induces the positively charged protons and positrons to rotate counterclockwise in the cosmic plane.

However, the gravitational spiral rotation is unidirectional (say clockwise) and the magnetically induced rotation of the positively charged protons is also clockwise, while the magnetic effect on the antiprotons is to make them rotate counterclockwise. Thus, there are more protons than antiprotons rotating clockwise. But for every proton that rotates clockwise, there is an antiproton that rotates counterclockwise. Thus, there is a separation of the matter from the antimatter due to the competition between the asymmetrical gravitationally induced rotations and the symmetrical magnetically induced rotational forces of the early universe.

In this way, then, the gravitational motion of the ‘matter’ of the universe superposes a spiral motion (in the same direction of rotation) of the separated antimatter universe. Of course, there are not two separated distinct universes. They are merely the spacelike separations of the matter and antimatter components of a single universe, components that are galaxies, planets, people, etc., one set composed of antimatter and the other of matter.

A prediction that follows, from this scenario, is that galaxies predominantly composed of antimatter, such as an anti-Milky Way, must entail the decay of antineutrons (rather than neutrons) to antiprotons, positrons and virtual neutrinos with positive helicity (rather than to protons, electrons and virtual neutrinos with negative helicity).

It is interesting in this regard that in the 1987 supernova explosion, neutrinos with positive helicity were detected, correlated with this event. This observation implies the possibility that the supernova explosion may have entailed the decay of antineutrons rather than neutrons. These stellar components of the supernova may then have been antimatter rather than matter components of the universe – an antimatter galaxy!

Bibliography

A contemporary book on quantum mechanics and gravity would normally have voluminous references. This would be useful for the reader when the topics covered are closely connected with ongoing research programs on the subject. However, this book is based on a research program starting with the approach of Albert Einstein that has had very few followers in the literature, since the inception of the theory of general relativity. Thus, to be useful for the reader, I have chosen to refer, in the first section below, only to those papers and books that have influenced my own program, in addition to references given at appropriate places throughout the text to papers on my research program. In the second section, I give references to present-day works on quantum gravity from the standpoint of the quantum theory, rather than the classical field view of general relativity of the present book. Finally, in the third and fourth sections, I will list my (complete) bibliography. It contains titles of research that are closely connected with the topics in this book. These are presented in two parts: the first part lists my bibliography by subject from 1953–1980, and the second part, the continuation of my bibliography since 1980.

General Relativity from the Standpoint of Classical Field Theory

First and foremost among the physicists who have influenced my research is Albert Einstein. Out of a very large number of his publications, I wish to list the following:

1. A. Einstein, *The Meaning of Relativity: Relativity Theory of the Non-Symmetric Field* (Princeton, 1956), 5th edn. This book contains the most recent of Einstein's writing on ideas of the theory of general relativity and the unified field theory
2. A. Einstein, H. Lorentz, H. Weyl, H. Minkowski, *The Principle of Relativity* (Dover, 1923). This book contains many of the original papers on relativity (in English translation)
3. P.A. Schilpp (Ed.), *Albert Einstein: Philosopher–Scientist* (Open Court, 1949), A. Einstein, *Autobiographical Notes*. In addition, all of Einstein's publications are listed at the end of this book

4. K. Przibram (Ed.), *Letters on Wave Mechanics* (Philosophical Library, 1967). Contains interesting correspondence between Einstein and his colleagues, Schrödinger, Planck and Lorentz
5. A. Einstein, *Generalization of the Relativistic Theory of Gravitation*, *Annals of Mathematics* **46**, 578 (1945). A seminal paper on Einstein's attempt at a unified field theory. Its sequel is with E. Straus, *Annals of Mathematics* **47**, 731 (1947)
6. E. Schrödinger, *Space-Time Structure* (Cambridge, 1954). Seminal work by a very profound physicist who, in my view, comes closest to Einstein's approach

Other interesting papers by Schrödinger on the unified field theory are:

7. E. Schrödinger, *Electric Charge and Current Engendered in Combined Maxwell-Einstein Fields*, *Proceedings Royal Irish Academy* **56 A**, 13 (1954)
8. E. Schrödinger, *The General Theory of Relativity and Wave Mechanics*, *Scientific Papers* (1953), p. 65
9. E. Schrödinger, *Are there Quantum Jumps?*, *British Journal for the Philosophy of Science* **3**, 3 (1952). Explicates Schrödinger's critical view of the commonly accepted interpretation of the quantum theory

A deep thinker in theoretical physics and mathematics, who was Schrödinger's colleague at the Dublin Institute for Advanced Studies, was Cornelius Lanczos. His voluminous and pertinent writings on the quantum and general relativity theories are collected in the set:

10. C. Lanczos: *Collected and Published Papers and Commentary*, ed. by W.R. Davis, M. Chu, P. Dolan, J.R. McConnell, L.K. Norris, E. Ortiz, R.J. Plemmons, D. Ridgeway, B.K. Scaite, W.J. Stewart, J.W. York Jr., W.B. Boggett, B.M. Gellai, A.A. Gsponer, and C.A. Prioli (North Carolina State University, Raleigh, 1998)
11. C. Lanczos, *Variational Principles of Mechanics* (Toronto, 1962). An excellent book on the calculus of variations

A great 19th century astronomer-physicist-mathematician was Sir William Rowan Hamilton. In addition to his pioneering work in mechanics and astronomy, Hamilton discovered the algebra of quaternions. His papers on this subject are collected in:

12. H. Halberstam and R.E. Ingram (Eds.), *The Mathematical Papers of Sir William Rowan Hamilton*, Vol. III Algebra (Cambridge, 1967)

One of the chief architects of the contemporary theory of quantum mechanics and quantum field theory was Paul A.M. Dirac.

13. P.A.M. Dirac, *The Principles of Quantum Mechanics* (Oxford, 1958), 4th edn. This classic book is, in my view, the clearest and most incisive explication of this subject. On the theory of general relativity, his book:

P.A.M. Dirac, *General Theory of Relativity* (Wiley, 1975) is highly recommended

Wolfgang Pauli was a very clear 20th century writer on the quantum theory as well as the theory of relativity.

14. W. Pauli, *The Theory of Relativity* (Pergamon, 1958), including updated notes, is a very important contribution to the literature on this subject

On the quantum theory, I have benefited from a series of lectures given by Pauli at the University of California in 1958. They are reprinted in the University of California Radiation Laboratory paper:

15. W. Pauli, *Lectures on Continuous Groups and Reflections in Quantum Mechanics* (UCRL 8213), ed. by R.J. Riddell (October, 1958)

Other references to quaternion and spinor analyses that I have studied are the following:

16. H.S. Ruse, Proc. Roy. Soc. (Edinburgh) **57**, 97 (1937); C.W. Kilmister, Proc. Roy. Soc. (London) A **199**, 517 (1949); C.W. Kilmister, Proc. Roy. Soc. (London) A **207**, 402 (1951); A.W. Conway, Proc. Roy. Soc. (London) A **162**, 145 (1937)

Texts I have studied on group theory are:

17. L.P. Eisenhart, *Continuous Groups of Transformations* (Princeton, 1933); L. Pontjagin, *Topological Groups* (Princeton, 1958); E.P. Wigner, *Group Theory and its Application to Atomic Spectra* (Academic, 1959)

Another of the great architects of 20th century physics of the microdomain was Louis de Broglie. He sums up his life work in:

18. L. de Broglie, *Recherches d'un Demi Siècle* (Albin Michel, 1976)

The following are some classic works by Willis E. Lamb, Jr. on his discovery of the Lamb shift:

19. *Fine Structure of the Hydrogen Atom I* (with R.C. Retherford), Phys. Rev. **79**, 549 (1950); *Fine Structure of the Hydrogen Atom II* (with R.C. Retherford), Phys. Rev. **81**, 222 (1951); *Fine Structure of the Hydrogen Atom III*, Phys. Rev. **85**, 259 (1952); *Fine Structure of the Hydrogen Atom IV* (with R.C. Retherford), Phys. Rev. **86**, 1014 (1952); *Fine Structure of the Hydrogen Atom V* (with S. Triebwasser and E.S. Dayhoff), Phys. Rev. **89**, 98 (1953); *Fine Structure of the Hydrogen Atom* (with E.S. Dayhoff and S. Triebwasser), Phys. Rev. **89**, 106 (1953)

For Lamb's interpretation of the quantum theory, see:

20. W.E. Lamb, Jr., *Remarks on the Interpretation of Quantum Mechanics*, Science of Matter, ed. by S. Fujita (Gordon and Breach, 1979), p. 1

Quantum Gravity from the Standpoint of Quantum Mechanics

There is an extensive bibliography on quantum gravity, from the standpoint of the quantum theory. This is in contrast with this book, which is viewed from the the basis of the theory of general relativity, as a classical field theory, wherein quantum features appear only as approximations. Needless to say, it is still not proven that there is indeed a bona fide quantum theory of gravity, from the view of quantum mechanics. It has been my contention that it is indeed impossible to unify the quantum and general relativity theories because of fundamental logical and mathematical incompatibilities. Nevertheless, a great deal of work has been done in the latter part of the 20th century and into the 21st century, aimed at fixing the problem with a possible unification. Recent works that treat the subject of quantum gravity extensively, from the standpoint of quantum mechanics, are the following [in the *Fundamental Theories of Physics Series* (Reidel and Kluwer) except for [23]] (in alphabetical order):

21. B. Bertotti, F. de Felice, and A. Pascolini (Eds.), *General Relativity and Gravitation* (Reidel, 1984)
22. V.P. Frolov, I.D. Novikov, *Black Hole Physics* (Kluwer, 1998)
23. S.W. Hawking and W. Israel (Eds.), *300 Years of Gravitation* (Cambridge, 1987). A comprehensive outline of contemporary research in astrophysics and black holes
24. B.R. Iyer and B. Bhawal (Eds.), *Black Holes, Gravitational Radiation and the Universe* (Kluwer, 1999)
25. B.R. Iyer, N. Mukunda, C.V. Vishveshwara (Eds.), *Gravitation, Gauge Theories and the Early Universe* (Kluwer, 1989)
26. R. Iyer and C.V. Vishveshwara (Eds.), *Geometry, Fields and Cosmology* (Kluwer, 1997)
27. J.V. Narlikar and T. Padmanabhan, *Gravity, Gauge Theories and Quantum Cosmology* (Reidel, 1986)
28. E. Prugovecki, *Quantum Geometry: A Framework for Quantum General Relativity* (Kluwer, 1992)

Author's Publications 1953–1980

Philosophy of Physics

29. *A New Approach to the Theory of Fundamental Processes*, British Journal for the Philosophy of Science **15**, 213 (1964)
30. *On the Elementarity of Measurement in General Relativity*, Synthese **17** (1967), Boston Studies in the Philosophy of Science (Humanities, 1968), Vol. III, p. 56, ed. by R. Cohen and M. Wartofsky
31. *Positivism, Realism and Existentialism in Mach's Influence on Contemporary Physics*, Philosophy and Phenomenological Research **30**, 403 (1970)
32. *Philosophical Implications of Unity in the Contemporary Arts and Sciences*, Philosophy and Phenomenological Research **34**, 389 (1974)
33. *On Pair Annihilation and the Einstein-Podolsky-Rosen Paradox*, International Journal for Theoretical Physics **1**, 387 (1968)
34. *Is Quantization Really Necessary?*, British Journal for the Philosophy of Science **21**, 359 (1970)
35. *Comment on 'Alternative to the Orthodox Interpretation of the Quantum Theory'*, American Journal of Physics **36**, 463 (1968); **37**, 228 (1969)
36. *On the Nature of Light and the Problem of Matter*, in C.A. Hooker (Ed.), *Contemporary Research in the Foundations and Philosophy of Quantum Theory* (Reidel, 1973), p. 346.
37. *Space, Time and Elementary Interactions in Relativity*, Physics Today **22**, 51 (1969)
38. *A Resolution of the Clock Paradox*, Physics Today **24**, 23 (1971)
39. Replies in: *The Clock Paradox – Majority View*, Physics Today **25**, 9 (1972)
40. *Comments on the Clock Paradox*, International Journal of Theoretical Physics **7**, 281 (1973)
41. *Reply to Greenberger on 'The Reality of the Twin Paradox Effect'*, American Journal of Physics **41**, 748 (1973)
42. *On the Mach Principle and Relative Space-Time*, British Journal for the Philosophy of Science **23**, 117 (1972)
43. *Parmenidian Particulars and Vanishing Elements*, Studies in the History and Philosophy of Science **3**, 151 (1972), with E.H. Madden
44. *The Field Concept in Contemporary Science* (Thomas, 1973)
45. *The Search for a Theory of Matter* (McGraw-Hill, 1971)
46. *On the Meaning of $E = mc^2$* , International Journal of Theoretical Physics **8**, 377 (1973)
47. In: *Quantum Mechanics Debate*, Physics Today **24**, 36 (1971)
48. *Ideas of the Theory of Relativity: General Implications from Physics to Problems of Society* (Israel Universities, 1974)

49. *On the Philosophy of General Relativity and Ideas of Eastern and Western Cultures*, in S. Fujita (Ed.), *Festschrift for Ta-You Wu* (Gordon and Breach, 1978), p. 9
50. In: *More About Tachyons*, *Physics Today* **22**, 47 (1969)
51. *On Dingle's Controversy about the Clock Paradox and the Evolution of Ideas in Science*, *International Journal of Theoretical Physics* **10**, 321 (1974)
52. Book review article on: K.R. Popper, *Objective Knowledge*, *Philosophy of the Social Sciences* **4**, 399 (1974)
53. *Maimonides, Spinoza and the Field Concept in Physics*, *Journal of the History of Ideas* **37**, 125 (1976)
54. *On the Mach Principle and General Relativity*, *British Journal for the Philosophy of Science* **26**, 49 (1975)
55. *On the Logical Status of Equivalence Principles in General Relativity Theory*, *British Journal for the Philosophy of Science* **27**, 225 (1976)
56. *On Wave-Particle Dualism*, *Annales de la Fondation Louis de Broglie* **1**, 129 (1976)
57. *Popper and Reality*, *Synthese* **33**, 355 (1976)
58. *Le Concept de Temps en Physique et en Cosmologie*, *La Recherche* **9**, 104 (1978) (in French)

Continuum Field Approach to Microscopic Physics

59. *Une Théorie Continue de la Matière Élémentaire*, *La Recherche* **6**, 1008 (1975) (in French)
60. *A New Theory of Elementary Matter. Part I: Philosophical Approach and General Implications*, *International Journal of Theoretical Physics* **4**, 433 (1971)
61. *A New Theory of Elementary Matter. Part II: Electromagnetic and Inertial Manifestations*, *International Journal of Theoretical Physics* **4**, 453 (1971)
62. *A New Theory of Elementary Matter. Part III: A Self-Consistent Field Theory of Electrodynamics and Correspondence with Quantum Mechanics*, *International Journal of Theoretical Physics* **5**, 35 (1972)
63. *A New Theory of Elementary Matter. Part IV: The Particle-Antiparticle Pair and Hydrogen*, *International Journal of Theoretical Physics* **5**, 161 (1972)
64. *A Self-Consistent Field Theory of Quantum Electrodynamics*, *Nuovo Cimento Supplement* **21**, 197 (1961) (with S.L. Schwebel)
65. *The Pauli Exclusion Principle from a Self-Consistent Field Theory of Quantum Electrodynamics*, *Nuovo Cimento* **27**, 1138 (1963)
66. *Black Body Radiation from a Self-Consistent Field Theory of Quantum Electrodynamics*, *Nuovo Cimento* **37**, 977 (1965)

67. *Implications of a New Approach to Quantum Electrodynamics in Electron-Proton Scattering*, Nuclear Physics **43**, 204 (1963) (with S.L. Schwebel), and in *Nuclear Structure*, ed. by L.I. Schiff and R. Hofstadter (Stanford, 1964), p. 366
68. *Implications of a New Approach to Quantum Electrodynamics in e -He⁴ Scattering*, Nuovo Cimento **53** A, 561 (1968)
69. *A Theory of Light Without Photons: Implications from Fermis to Light-Years*, in *Coherence and Quantum Optics*, ed. by L. Mandel and E. Wolf (Plenum, 1973), p. 605
70. *Completion of the Proof of the Pauli Exclusion Principle from a Self-Consistent Field Theory of Electrodynamics*, Nuovo Cimento **44** B, 289 (1978)
71. *The Lamb Shift in Helium from a Self-Consistent Field Theory of Electrodynamics*, Lettere Nuovo Cimento **13**, 73 (1975) (with H. Yu)
72. *The Anomalous Magnetic Moment of the Electron from a Self-Consistent Theory of Electrodynamics*, Lettere Nuovo Cimento **13**, 169 (1975)
73. *The Lamb Shift in Atomic Helium and the Anomalous Magnetic Moment of the Electron from a New Theory of Elementary Matter*, in *Coherence and Quantum Optics*, ed. by L. Mandel and E. Wolf (Plenum, 1978), p. 487

Generalizations of Electromagnetic Theory

74. *Implications of Parity Nonconservation and Time Reversal Noninvariance in Electromagnetic Interactions*, Annals of Physics (New York) **6**, 244 (1959)
75. *Implications of Parity Nonconservation and Time Reversal Noninvariance in Electromagnetic Interactions. Part II: Atomic Energy Levels*, Annals of Physics (New York) **8**, 479 (1959) (with S.L. Schwebel)
76. *On Covariant Formulations of the Maxwell-Lorentz Theory of Electromagnetism*, Journal of Mathematical Physics **3**, 845 (1962) (with S.L. Schwebel)
77. *A Spinor Formulation of Electromagnetic Theory in General Relativity*, Nuovo Cimento **31**, 98 (1964)
78. *Classical Effects from a Factorized Spinor Representation of Maxwell's Equations*, International Journal of Theoretical Physics **4**, 145 (1971)
79. *Significance of the Measurement of the Electric Dipole Moment of the Cesium Atom*, Physics Letters **14**, 302 (1965)

Inertia from General Relativity

80. *On Spinor Connection in a Riemannian Space and the Masses of Elementary Particles*, Nuovo Cimento **34**, 81 (1964)

81. *The Mass of a Neutrino from Spinor Connection in a Riemannian Space*, Nuovo Cimento **37**, 888 (1965)
82. *On Positive-Definiteness of Inertial Mass, Non-Positive Definiteness of Electromagnetic Coupling and the Mach Principle from General Relativity*, Nuovo Cimento **53 B**, 398 (1968)
83. *On the Electron–Muon Mass Doublet from General Relativity*, Nuovo Cimento **7 B**, 247 (1972)
84. *On the Lifetime of the Muon State of the Electron–Muon Mass Doublet*, Nuovo Cimento **10 B**, 339 (1972)
85. *Matter and Antimatter from General Relativity*, Lettere Nuovo Cimento **5**, 947 (1972)
86. *The Inertial Mass of a Pion from a Quaternion Field Theory of Matter*, Nuovo Cimento **43 A**, 74 (1978)
87. *On the Ratio of Neutral to Charged Pion Lifetimes in General Relativity*, Lettere Nuovo Cimento **24**, 577 (1979)

Unified Field Theory

88. *On Factorization of Einstein's Formalism into a Pair of Quaternion Field Equations*, Nuovo Cimento **47**, 759 (1967)
89. *A Symmetric Tensor–Antisymmetric Tensor Theory of Gravitation and Electromagnetism from a Quaternion Representation of General Relativity*, Nuovo Cimento **55 B**, 199 (1968)
90. *On the Most General Form of Field Theory from Symmetry Principles*, Nature **226**, 138 (1970)
91. *On the Quantization of Electrical Charge from General Relativity*, Lettere Nuovo Cimento **21**, 123 (1978)

Astronomy and Cosmology

92. *The Geodesic Equation and Planetary Motion from a Quaternion Representation of General Relativity*, Nuovo Cimento **66 B**, 137 (1970)
93. *A General Relativity Test Using Two or More Solar Satellites*, Astronautica Acta **15**, 65 (1969) (with H.G. Lafromboise)
94. *The Hubble Law and the Spiral Structures of Galaxies from Equations of Motion in General Relativity*, International Journal of Theoretical Physics **14**, 115 (1975)
95. *On Stellar Collapse and the Black Hole Limit from a Dynamical View*, Annales Institut Henri Poincaré **28**, 399 (1978)

Atomic, Nuclear and Solid State Physics

96. *Solid State Theory* (McGraw-Hill, 1963) (republished by Dover, 1974)
97. *Nuclear Hyperfine Structure of Mn^{++}* , Phys. Rev. **90**, 1058 (1953)
98. *Paramagnetic Resonance Absorption of Mn^{++} in Single Crystals of $CaCO_3$* , Phys. Rev. **93**, 373 (1954) (with W. Hershberger and K. Hurd)
99. *Selection Rules for the Absorption of Polarized Electromagnetic Radiation by Mobile Electrons in Crystals*, Phys. Rev. **107**, 437 (1957)
100. *Rotational Properties of Paramagnetic Resonance Spectra in Non-Cubic Crystals*, Journal of Physics and Chemistry of Solids **15**, 291 (1960)
101. *Semiempirical Model for Direct Nuclear Breakup Reactions*, Phys. Rev. **103**, 671 (1956)

Author's Publications Since 1980

102. *On Einstein's Views of the Relativity and Quantum Theories and their Future Progress I*, Annales de la Fondation Louis de Broglie **4**, 85 (1979)
103. *On Einstein's Views of the Relativity and Quantum Theories and their Future Progress II: Toward a Unified Field Theory*, Annales de la Fondation Louis de Broglie **4**, 175 (1979)
104. *Implications of a New Approach to Electrodynamics in Proton-Nucleus Scattering at Small Distances*, Lettere Nuovo Cimento **27**, 586 (1980)
105. *Bell's Inequalities from the Field Concept in General Relativity*, Nuovo Cimento **58 A**, 1 (1980)
106. *A Proton Mass Doublet from General Relativity*, Nuovo Cimento **59 A**, 113 (1980)
107. *A New Look at Electromagnetic Field Theory*, Foundations of Physics **10**, 921 (1980)
108. Book review article on *Philosophical Problems of Modern Physics*, by P. Mittlestaedt, Philosophia **8**, 815 (1979)
109. Articles published in *American Academic Encyclopedia* (Arete Publishers, 1980), *Energy Level*, p. 173, Vol. 7; *Spectrum*, p. 172, Vol. 18
110. *Is the World Really Made out of Parts?*, Proc. Int. Conf. Cybernetics and Society IEEE, p. 635 (Oct. 8–10, Denver, Colorado, 1979)
111. *Elementary Particles from General Relativity*, Foundations of Physics **11**, 329 (1981)
112. *On the Question of Neutrino Emission from the Sun*, Lettere Nuovo Cimento **32**, 307 (1981)
113. *A Composite Neutron Model*, Nuovo Cimento **66 A**, 94 (1981)
114. *Ideas of Matter: From Ancient Times to Bohr and Einstein* (University Press of America, 1981)

115. *A Pulsar Model from an Oscillating Black Hole*, Foundations of Physics **12**, 689 (1982)
116. *On Time Reversal Noninvariance in Nuclear Forces*, Lettere Nuovo Cimento **34**, 441 (1982)
117. *General Relativity and Matter: A Spinor Field Theory from Fermis to Light-Years* (Reidel, 1982)
118. *A Generalized e - p Interaction at Small Distances from a New Approach to Electrodynamics*, Nuovo Cimento **70** A, 229 (1982)
119. *On the Incompatibility of the Quantum and Relativity Theories and a Possible Resolution*, Hadronic Journal **5**, 1781 (1982)
120. *On the Possible Origin of CP Violation in Neutral Kaon Decay*, Nuovo Cimento **72** A, 361 (1982)
121. *Antiscience Within Science*, B'Or Ha Tora, No. 2 (1982), 52
122. *On a Scientist's Comprehension of the Universe*, B'Or Ha Tora, No. 3 (1983), p. 23
123. *Comparison of the Field Concept of Matter in Relativity Physics and the Buddhist Idea of Nonself*, Philosophy East and West **33**, 395 (1983)
124. *Proton-Antiproton Collisions and the $W^\pm$ -Particle from General Relativity*, Lettere Nuovo Cimento **39**, 17 (1984)
125. *God and the Physicists* (book review article on: *God and the New Physics* by P. Davies, with M. Bunge), Free Inquiry **4**, 52 (1984)
126. *Reply to 'God and the Physicists'*, Free Inquiry **4**, 62 (1984)
127. *Spazio, Tempo e Movimento nella Teoria della Relativita di Einstein*, Laboratorio Scienze Lettere Arti **4**, 35 (1984)
128. Book review article on: *The Open Universe: An Argument for Indeterminism*, by K.R. Popper, Philosophy of Social Science **15**, 205 (1985)
129. *On Einstein's Later View of the Twin Paradox*, Foundations of Physics **15**, 977 (1985)
130. Book review article on: *Quantum Theory and the Schism in Physics*, by K.R. Popper, Philosophy of Social Science **15**, 321 (1985)
131. *On $\mu \rightarrow e^\pm + 2\gamma$ in General Relativity*, Nuovo Cimento **91** A, 241 (1986)
132. *Irreversibility in Deterministic Physics*, Annales de la Fondation Louis de Broglie **11**, 7 (1986)
133. *The Lifetime of a Neutron from a Composite Model*, Nuovo Cimento **93** A, 325 (1986)
134. *Investigation of the Weak Interaction from General Relativity*, Nuovo Cimento **94** A, 390 (1986)
135. Book review article on: *Whitehead's Metaphysics of Extension and Solidarity* by J.L. Nobo, Transactions of the Peirce Society **23**, 153 (1987)
136. *Quantum Mechanics from General Relativity: An Approximation for a Theory of Inertia* (Reidel, 1986)
137. *On the Tau Lepton in General Relativity*, Nuovo Cimento **96** A, 166 (1986)

138. *On the Mass of the $W^\pm$ -Particle from Generalized Electrodynamics*, Nuovo Cimento **98 A**, 197 (1987)
139. *The Precessional Frequency of a Gyroscope in the Quaternionic Formulation of General Relativity*, Foundations of Physics **19**, 105 (1989)
140. *Bohr Versus Einstein: A Possible Shaping of 21st Century Physics*, in: Problems in Quantum Physics: Gdansk '87, ed. by Kostro, Posiewnik, Pykacz and Zukowski (World Scientific, 1988), p. 745
141. *On the Inertial Mass Concept in Special and General Relativity*, Foundations of Physics Letters **1**, 163 (1988)
142. *Einstein versus Bohr: The Continuing Controversies in Physics* (Open Court, 1988)
143. *On Ideas of Physics and Kabbala*, B'Or Ha Tora, No. 6 (1987), p. 103
144. *On the Origin of Spin in Relativity*, British Journal for the Philosophy of Science **40**, 409 (1989)
145. *Considerations of an Oscillating Spiral Universe Cosmology*, Annales de la Fondation Louis de Broglie **14**, 361 (1989)
146. *Response to Rodrigues and Rosa on the Twin Paradox*, Foundations of Physics **19**, 1525 (1989)
147. *Space, Time and Motion in Einstein's Theory of Relativity*, Poznan Studies on the Philosophy of the Sciences and the Humanities **17**, 151 (1989)
148. *Einstein and the Evolution of Twentieth Century Physics*, Physics Essays **3**, 80 (1990)
149. Book review article on: *The World View of Contemporary Physics: Does it Need a New Metaphysics*, ed. by J.F. Kitchener, Philosophy of Social Science **20**, 233 (1990)
150. *Biographies of Albert Einstein*, in St James Guide to Biography (St. James Press, 1991)
151. *Einstein and the Evolution of Twentieth Century Physics*, Annales de la Fondation Louis de Broglie **16**, 241 (1991)
152. *On the Ratio of Lifetimes of the Neutral Short-Lived and Long-Lived Kaons* Nuovo Cimento **105 A**, 395 (1992)
153. *On Czachor's Error in his Analysis of Sachs' Unified Field Theory*, Physics Essays **5**, 548 (1992)
154. *Relativity in our Time: From Physics to Human Relations* (Taylor and Francis, 1993)
155. *On Hawking's 'A Brief History of Time' and the Present State of Physics*, British Journal for the Philosophy of Science **44**, 543 (1993)
156. *On the Problem of Cosmology*, Physics Essays **6**, 32 (1993)
157. *Relativistic Implications of Electromagnetic Field Theory*, in: Advanced Electromagnetism, ed. by T.W. Barrett and D.M. Grimes (World Scientific, 1995), p. 541
158. *On Quantum Mechanics from General Relativity: A Dialogue with J.S. Bell*, Annales de la Fondation Louis de Broglie **20**, 33 (1995)

159. *On Quantum Mechanics from General Relativity: A Paradigm Shift*, in: Courants, Amers, Ecueils en Microphysique (Directions in Microphysics), ed. by C. Cormier-Delanoue, G. Lochak, and P. Lochak (Fondation Louis de Broglie, 1993), p. 393
160. *On the Rotation of Galaxies from General Relativity*, Physics Essays **7**, 490 (1994)
161. *On the Spectrum of Neutrino Mass in General Relativity*, Nuovo Cimento **107 A**, 1709 (1994)
162. *On Comprehension of the Universe – from Spinoza to Einstein*, Science Philosophy Interface **1**, 15 (1996)
163. *Could the Photon be Superfluous?*, in: The Present Status of the Quantum Theory of Light, ed. by S. Jeffers et al. (Kluwer, 1997), p. 79
164. *On the Future Course of Quantum Physics*, Apeiron **2**, 104 (1995)
165. *An Interpretation of the Top Quark Mass in Terms of a Proton Mass Doublet in General Relativity*, Nuovo Cimento **108 A**, 1445 (1995)
166. Book review article on: *The Collected Papers of Albert Einstein Vol. 5: The Swiss Years Writings, 1902–1914*, English translation, Anna Beck, Physics Today **48**, 65 (1995)
167. *On the Separation of Matter and Antimatter in the Early Universe*, Annales de la Fondation Louis de Broglie **20**, 323 (1995)
168. *Changes in Concepts of Time from Aristotle to Einstein*, Proc. Conf. on Modern Mathematical Models of Time and Applications to Physics and Cosmology, University of Arizona (Kluwer, 1997)
169. *Changes in Concepts of Time from Aristotle to Einstein*, Astrophysics and Space Science **244**, 269 (1996)
170. *The Neutron-Proton Mass Difference from a Composite Neutron Model*, Nuovo Cimento **109 A**, 1641 (1996)
171. *On the Source of Anisotropy in Cosmic Radiation from General Relativity*, Nuovo Cimento **110 A**, 611 (1997)
172. *The Future of Physics?*, Annales de la Fondation Louis de Broglie **22**, 271 (1997)
173. *On the Role of Space and Time in Relativity Theory*, Apeiron **5**, 121 (1998)
174. *A Speculation on Einstein's Reply to 'Quantum Teleportation'*, Physics Essays **11**, 1 (1998)
175. *Quantum Mechanics from General Relativity: Particle Probability from Interaction Weighting*, Annales de la Fondation Louis de Broglie **24**, 25 (1999)
176. *Reinterpretation of the Fitzgerald-Lorentz Contraction*, with M. Harada, Physics Essays **11**, 521 (1998)
177. *Fundamental Conflicts in Modern Physics and Cosmology*, Frontier Perspectives **8**, 13 (1999)
178. *On Unification of Gravity and Electromagnetism and the Absence of Magnetic Monopoles*, Nuovo Cimento **114 B**, 123 (1999)

179. *Elementary Particles from General Relativity*, in: Photon: Old Problems in the Light of New Ideas, ed. by V.V. Dvoeglazov (Nova Science, 2000), p. 135
180. *Is there a Beginning of Time? Answers from Science and Theology*, Science Philosophy Interface **5**, 1 (2000)
181. *The Field Concept and Holism in Relativity Physics and Comparisons with Asian Philosophies*, Science Philosophy Interface **5**, 14 (2000)
182. *Will the 21st Century See a Paradigm Shift in Physics from the Quantum Theory to General Relativity?*, Revue Internationale de Philosophie **54**, 351 (2000)
183. *Dialogues on Modern Physics* (World Scientific, 1998)
184. *On the Meaning of Space and Time*, in: Recent Advances in Relativity Theory, Vol. 1, Formal Interpretations, ed. by M.C. Duffy and M. Wegener (Hadronic Press, 2000), p. 170
185. *From Atomism to Holism in 21st Century Physics*, Annales de la Fondation Louis de Broglie **26**, 389 (2001)
186. *Symmetry in Electrodynamics: From Special to General Relativity, From Macro to Quantum Domains*, in: Modern Nonlinear Optics, Part 1, 2nd edn., Advances in Chemical Physics **119**, ed. by M.W. Evans, Series editors I. Prigogine and S.A. Rice (Wiley, 2001), pp. 677–706
187. *The Influence of the Physics and Philosophy of Einstein's Relativity on my Attitudes in Science: An Autobiography*, in: Fragments of Science: Festschrift for Mendel Sachs, ed. by M. Ram (World Scientific, 1999) p. 201
188. *Cornelius Lanczos – Discoveries in the Quantum and General Relativity Theories*, Annales de la Fondation Louis de Broglie **27**, No. 1, pp. 85–91 (2002)
189. *Science and Religion*, Science Philosophy Interface (in press)
190. Website: www.compukol.com/mendel

Index

- absolute time 160
- accidentally degenerate states 129
- advanced potential 96
- anti-Lorentz force density 93
- anti-Ricci tensor 56
- antimatter universe 173

- big bang 149
- binding energy of neutron 132
- black hole 167
 - binary 167
- blackbody radiation 2
- Bohr model 5
- bound electron-positron 113

- charge conjugate 111
- charge-neutral kaon 141
- conservation
 - of electric charge 57
 - of interaction 101
- conservation equation 88, 92
- continuous field concept 20
- correlated matter fields 98
- covariance 25
- curved spacetime 36, 150

- d'Alembert's equation 115
- damped oscillator 169
- dark matter 119
- de Broglie wavelength 6
- delayed action-at-a-distance 83
- density of states 140
- deterministic theory 23, 119
- Dirac matter field equation 100
- Dirac spinor wave equation 68

- early universe 173
- Einstein constant 157
- Einstein group 45, 61, 144
- Einstein's field equations 29
- Einstein's theory of general relativity 19
- electric dipole moment 93
- electromagnetism 86
- electron 84
 - diffraction 6
- energy levels 127
- entanglement 99, 108

- factorized form of field equations 53
- Fermi-Dirac statistics 110
- force of gravity 16
- Fresnel integrals 165

- Galileo's law of gravity 12
- Galileo's principle of inertia 152
- gauge
 - covariance 73
 - transformation 74
- generalized Mach principle 80, 95, 99
- geodesic equation 36, 37, 154
- global extension 90
- gravitation 76
 - theories 148
- gravitational
 - mass 152
 - redshift 19
- graviton 23
- Green function 95

- symmetric 105
- ground state 114
- Hamilton, William 62
- Heisenberg's matrix mechanics 8
- hidden variables 10
- holism VIII, 99
- holistic
 - continuum 149
 - field theory 103
- Hubble law 160, 163
- hydrogen 122
 - spectrum 130
- inertial mass 15, 64, 72, 152
- infinite mass spectrum 136
- interaction Lagrangian 83
- interaction volume 141
- interaction weighting amplitude 104
- irreducible representations 91
- kaon, charge-neutral 141
- Kepler's constant 157
- Lagrange–Euler equation 51
- Lagrangian density 50
- Lamb shift 89, 127
- least action principle 88
- lens effect 166
- lepton 137
- Lie group 34
- lifetime 137
 - of neutron 132
- linear approximation VIII
- linear formalism 22
- linearity vs. nonlinearity 108
- linearization 124
- linearized electron equation 126
- logical positivism 7
- Lorentz force 95
 - density 81
- Mach principle 2, 64, 78, 136
 - generalized 80, 95, 99
- magnetic monopoles 57
- Majorana equation 69
- many-fermion system 110
- mass field 70
- matter–antimatter problem 171
- Maxwell's equations 81
- muon 135
- neutrino 133
- neutron 123
 - binding energy 132
 - lifetime 132
- neutron star 171
- Newton's theory of universal gravitation 17
- Newton's third law 15
- Newton, Isaac 13
- Newtonian approximation 77
- Noether's theorem 74
- nonlinear formalism 22
- nonlinearity 35
- opposing paradigms VII
- oscillating universe cosmology 77, 160
- oscillatory harmonics 155
- pair
 - annihilation 102
 - creation 117
- Palatini method 52
- parallel transfer 32
- Pauli exclusion principle 103
- perihelion precession 18, 159
- photoelectric effect 3
- photon 118
- physical vacuum 65, 117
 - of pairs 172
- pion 138
 - mass eigenvalues 139
- Poincaré group 45
- polarized currents 116
- principle of covariance 25
- principle of equivalence 150
- probability calculus 9
- proton equation 125

- pulsar 168, 171
- quantization of electrical charge 143
- quantum mechanics 66
- quantum theory of gravity 1, 168
- quark 143
- quaternion 46
 - algebra 59
 - factorization 153
 - line element 60
 - metric 46
- radiation solutions 83
- radius of the universe 162
- regular functions 44
- relativistic cosmology 160
- relativistic wave equation 67
- Ricci tensor 55
- Riemann curvature tensor 31, 55
- Riemannian geometry 28
- rotational motion 164
- Schrödinger wave equation 8
- Schwarzschild
 - approximation 157
 - problem 155
 - radius 41
 - solution 35, 38
- Science of Pure Time 59
- self-energy 107
- singularity 42
- Slater determinant 104, 110
- speed of light 27
- spin affine connection 47
- spin curvature tensor 49
- spinor 70
 - field equations 86
 - representation 85
 - variable 48
 - wave equation 69, 71
- spiral motion 165
- stationary orbit 17
- stationary state problem 61
- superposition principle 69
- unified field theory 51, 58, 80
- variational calculation 50
- virtual neutrino 130
- wave monism 11
- weak interaction 138
- weighting function 102
- zero energy 118

Mendel Sachs

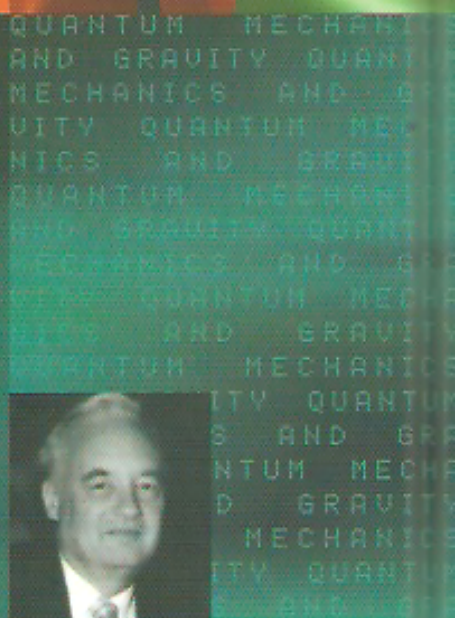
QUANTUM MECHANICS
AND GRAVITY

This book describes a paradigm change in modern physics from the philosophy and mathematical expression of the quantum theory to those of general relativity. The approach applies to all domains – from elementary particles to cosmology. The change is from the positivistic view in which atomism, nondeterminism and measurement are fundamental, to a holistic view in realism, wherein matter – electrons, galaxies, – are correlated modes of a single continuum, the universe. A field that unifies electromagnetism, gravity and inertia is demonstrated explicitly, with new predictions, in terms of quaternion and spinor field equations in a curved spacetime. Quantum Mechanics emerges as a linear, flat-space approximation for the equations of inertia in general relativity.

ISBN 3-540-00800-4



9 783540 008002

 springeronline.com


Professor Mendel Sachs (State University of New York/ Buffalo) is the author of ten other books; among them are: *Dialogues on Modern Physics* (1998), *Einstein Versus Bohr* (1988), *Quantum Mechanics from General Relativity* (1986) and *General Relativity and Matter* (1982). Mendel Sachs was born in Portland, Oregon and grew up in Toronto, Canada. He received his Ph. D. in Physics at UCLA. He is married and has four children.