CHAPTER

4

Tensors in Physics

We have now reached the point where our mathematical machinery is sufficiently well developed to allow us to begin applying it to physics. Our purpose in this chapter is twofold: First, by applying the techniques of tensor analysis to familiar areas of physics, we shall gain facility in the use of our mathematical tools and gain insight as to how the relativistic gravitational theory might be formulated. Second, and much more important, we shall be led to answer several fundamental questions concerning the background concepts of the gravitational theory. We shall, for instance, demonstrate the suitability of the Lorentz metric for describing electromagnetic phenomena in mass-free space; we shall discuss the nature of the relation between the line element ds and the proper-time interval $d\tau$, and we shall choose and justify a tensor equation of motion. In the process of the investigation, it is also hoped that some equations of classical physics will be considerably simplified.

4.1 Maxwell's Equations in Tensor Form

We assume the reader is familiar with Maxwell's equations and classical electrodynamics as well as a reasonable amount of special relativity. The task of this section will be to formulate Maxwell's equations in covariant tensor form—one which is valid in all Riemannian coordinate systems. To do this we begin by considering the equations in just *one coordinate system at rest* and make a purely *formal* change to *tensor notation.* Afterwards, we can consider the transformation to other systems, such as those in motion relative to the original one; but in what follows it should always be kept in mind that we deal only with a single rest system until we explicitly state otherwise.

In résumé of classical electrodynamics we list the Maxwell equations in two conveniently separated pairs. The first pair,

$$\nabla \times \mathbf{H} - \frac{1}{c}\dot{\mathbf{E}} = \frac{1}{c}\mathbf{j} \qquad \dot{\mathbf{E}} \equiv \frac{\partial \mathbf{E}}{\partial t} \tag{4.1}$$

$$\nabla \cdot \mathbf{E} = \rho \qquad \nabla \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \tag{4.2}$$

(in Heaviside-Lorentz units) involves the current density $\mathbf{j}$ and the charge density ρ, which are the sources of the electric field $\mathbf{E}$ and the magnetic field $\mathbf{H}$. We shall accordingly refer to this pair as the *source equations*. The second pair,

$$\nabla \times \mathbf{E} + \frac{1}{c}\dot{\mathbf{H}} = 0 \tag{4.3}$$

$$\nabla \cdot \mathbf{H} = 0 \tag{4.4}$$

is homogeneous and deals only with the relation between the electric and the magnetic fields. Since the components of the electric and magnetic vector fields will be seen to be also the components of an electromagnetic field tensor in what follows, we can say that (4.3) and (4.4) are internal relations between the components of a single tensor. Accordingly, these will be referred to as the *internal equations*.

The sources of the field, the ρ and $\mathbf{j}$ of the source equations, cannot be independently specified. In general, (4.1) and (4.2) are not compatible unless a consistency relation between the sources is satisfied. We can obtain this consistency relation by taking the divergence of both sides of (4.1):

$$\nabla \cdot \nabla \times \mathbf{H} - \frac{1}{c}\nabla \cdot \dot{\mathbf{E}} = \frac{1}{c}\nabla \cdot \mathbf{j} \tag{4.5}$$

and inserting

$$\dot{\rho} = \nabla \cdot \dot{\mathbf{E}} \tag{4.6}$$

which is obtained by differentiating (4.2). Noting that the divergence of a curl is identically zero for any vector, we then find that

$$\dot{\rho} + \nabla \cdot \mathbf{j} = 0 \tag{4.7}$$

We have here a differential relation between charge density and current density which follows directly from the Maxwell equations. Unless this relation between the sources is satisfied, the Maxwell equations can have no solution, so we shall always assume that, for all physical sources, (4.7) is satisfied.

If we interpret $\mathbf{j}$ as the convection current

$$\mathbf{j} = \rho \mathbf{v} \tag{4.8}$$

where $\mathbf{v}$ is the velocity field of the material with charge density ρ, then (4.7) becomes identical with the continuity equation of fluid mechanics.

$$\dot{\rho} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{4.9}$$

This equation states that the quantity having density ρ is neither created nor destroyed; i.e., it is conserved. Thus we can interpret (4.7) as the *physical law of conservation of charge* as well as the *mathematical requirement for consistency of the source equations.*

Our first step in writing the Maxwell equations (4.1) to (4.4) and the conservation equation (4.7) with tensor notation is to choose a convenient set of four-dimensional coordinates and an appropriate metric tensor in this coordinate system. Following the lead of the special theory of relativity we shall utilize the space and time coordinates

$$x^0 = ct \qquad x^1 = x \qquad x^2 = y \qquad x^3 = z \tag{4.10}$$

where x, y, z are Cartesian space coordinates, and t is a time coordinate.

It will prove most convenient to adopt the convention, in the following chapters, that Greek indices run from 0 to 3 and Latin indices run from 1 to 3:

$$\begin{pmatrix} \nu = 0, 1, 2, 3 & \text{(Greek)} \\ j = 1, 2, 3 & \text{(Latin)} \end{pmatrix}$$

The question of what is an "appropriate" metric tensor to use with these coordinates can be approached in several different ways. By an appropriate metric tensor we here mean one which allows us—using the coordinates of special relativity—to write Maxwell's equations in a simple and elegant form which is easily generalized to a truly covariant form. The authors take the viewpoint that the motivation for the choice of an appropriate metric tensor should come from the elegant theory of classical electrodynamics and the structure of the Maxwell equations alone; if at all possible, it should not be necessary to depend on other theories such as special relativistic mechanics. Thus, in the following paragraphs, we shall attempt to obtain a quadratic form which plays a distinguished role in the Maxwell equations and which may therefore be adopted as the quadratic form of the metric tensor. [If the reader is not interested in this particular motivational approach to the choice of a metric tensor, he may skip directly to Eq. (4.43).]

To begin our search for a distinguished quadratic form we shall show that, associated with the Maxwell equations in vacuum, there is a class of particularly interesting surfaces. To be precise, let us ask the following question: Do there exist unique three-dimensional hypersurfaces, imbedded in the four-dimensional space we have chosen, on which the first derivatives of the **E** and **H** fields can be *discontinuous?* (In the region of the hypersurface we still demand that Maxwell's equations be satisfied.) If so, what is the nature of these surfaces? Such surfaces are quite important in the study of partial differential equations; they are generally referred to as characteristic hypersurfaces, or simply as characteristics. In the case of the first-order Maxwell equations, the characteristic has been defined as a hypersurface over which a discontinuity in the first derivatives can occur, but in the general case of an nth-order equation, the characteristic is defined as a hypersurface over which a discontinuity in the nth derivatives can occur.

In addition to being mathematically interesting, the notion of a characteristic can be seen to be physically meaningful. Suppose one passes a sharp pulselike electromagnetic disturbance through empty space. The front of the pulse can be very sharp if discontinuous first derivatives in the **E** and **H** fields are allowable on the hypersurface corresponding to the wave front. Thus we expect the characteristic hypersurfaces to the wave front. Thus we expect the characteristic hypersurfaces to correspond to allowed physical wave fronts. We know that the ideas of special relativity rest heavily on the assumption that such wave fronts propagate with the constant velocity c, so the motivation for investigating the characteristic hypersurfaces is apparent.

In vacuum the densities ρ and $\mathbf{j}$ are zero and Maxwell's equations accordingly become

$$\nabla \times \mathbf{H} - \frac{1}{c}\dot{\mathbf{E}} = 0 \tag{4.11}$$

$$\nabla \cdot \mathbf{E} = 0 \tag{4.12}$$

$$\nabla \times \mathbf{E} + \frac{1}{c}\dot{\mathbf{H}} = 0 \tag{4.13}$$

$$\nabla \cdot \mathbf{H} = 0 \tag{4.14}$$

The equation of any three-dimensional hypersurface imbedded in a four-dimensional space can be written in the form

$$w(x^0,x^1,x^2,x^3) = 0 \tag{4.15}$$

We shall suppose that w has continuous first derivatives in all the variables, which means that it is a *smooth hypersurface.* If we further assume that $\partial w/\partial x^0$ is nonzero, we can solve (4.15) for x^0 as a function of x^1, x^2, and x^3 and write the equation of the three-dimensional hypersurface as

$$x^0 = h(x^1,x^2,x^3) \tag{4.16}$$

Then, on the hypersurface defined by (4.15), which we shall refer to as S, we have

$$w(x^0,x^1,x^2,x^3) = h(x^1,x^2,x^3) - x^0 = 0 \tag{4.17}$$

The electric and magnetic fields on S are functions of only x^1, x^2, x^3, since x^0 is specified by (4.17) when these coordinates are given. We shall therefore denote the electric and magnetic fields on S as the following vector functions of these three variables:

$$\hat{\mathbf{E}}(x^1,x^2,x^3) = \mathbf{E}(h,x^1,x^2,x^3) \tag{4.18}$$

$$\hat{\mathbf{H}}(x^1,x^2,x^3) = \mathbf{H}(h,x^1,x^2,x^3) \tag{4.19}$$

where h is the value of x^0 given by (4.17). The vector functions $\hat{\mathbf{E}}$ and $\hat{\mathbf{H}}$ are assumed to have continuous first derivatives. Using the above definitions and a bit of vector algebra, we shall be able to obtain a pair of very useful relations (V. Fock, 1959) which the functions $\dot{\mathbf{E}}$, $\dot{\mathbf{H}}$, $\hat{\mathbf{E}}$, $\hat{\mathbf{H}}$, and h must obey on S. These relations will be the key to obtaining the characteristic surfaces of Maxwell's equations.

From (4.18), (4.19), and (4.17) we have

$$\frac{\partial \hat{E}_k}{\partial x^i} = \frac{\partial E_k}{\partial x^i} + \frac{\partial E_k}{\partial x^0}\frac{\partial h}{\partial x^i} \tag{4.20}$$

Setting $k = i$ and summing from 1 to 3, we obtain

$$\nabla \cdot \hat{\mathbf{E}} = \nabla \cdot \mathbf{E} + \frac{1}{c}\dot{\mathbf{E}} \cdot \nabla h \tag{4.21}$$

From (4.12) this gives

(4.22) $$\nabla \cdot \hat{\mathbf{E}} = \frac{1}{c}\dot{\mathbf{E}} \cdot \nabla h$$

For the magnetic field we may obtain the analogous result

(4.23) $$\nabla \cdot \hat{\mathbf{H}} = \frac{1}{c}\dot{\mathbf{H}} \cdot \nabla h$$

From (4.20) we also obtain

$$\frac{\partial E_k}{\partial x^i} - \frac{\partial E_i}{\partial x^k} + \frac{\partial E_k}{\partial x^0}\frac{\partial h}{\partial x^i} - \frac{\partial E_i}{\partial x^0}\frac{\partial h}{\partial x^k} = \frac{\partial \hat{E}_k}{\partial x^i} - \frac{\partial \hat{E}_i}{\partial x^k}$$

that is, in vector notation,

(4.24) $$\nabla \times \mathbf{E} + \frac{1}{c}\nabla h \times \dot{\mathbf{E}} = \nabla \times \hat{\mathbf{E}}$$

and as the analogous result for the magnetic field, we have

(4.25) $$\nabla \times \mathbf{H} + \frac{1}{c}\nabla h \times \dot{\mathbf{H}} = \nabla \times \hat{\mathbf{H}}$$

Substituting into these last two equations from Maxwell's equations in vacuum, (4.11) and (4.13), we obtain

(4.26) $$-\frac{1}{c}\dot{\mathbf{H}} + \frac{1}{c}\nabla h \times \dot{\mathbf{E}} = \nabla \times \hat{\mathbf{E}}$$

(4.27) $$\frac{1}{c}\dot{\mathbf{E}} + \frac{1}{c}\nabla h \times \dot{\mathbf{H}} = \nabla \times \hat{\mathbf{H}}$$

The scalar product of (4.26) and (4.27) with ∇h gives

(4.28) $$-\frac{1}{c}\nabla h \cdot \dot{\mathbf{H}} + \frac{1}{c}\nabla h \cdot \nabla h \times \dot{\mathbf{E}} = -\frac{1}{c}\nabla h \cdot \dot{\mathbf{H}} = \nabla h \cdot \nabla \times \hat{\mathbf{E}}$$

(4.29) $$\frac{1}{c}\nabla h \cdot \dot{\mathbf{E}} + \frac{1}{c}\nabla h \cdot \nabla h \times \dot{\mathbf{H}} = \frac{1}{c}\nabla h \cdot \dot{\mathbf{E}} = \nabla h \cdot \nabla \times \hat{\mathbf{H}}$$

Also the vector product of (4.26) and (4.27) with ∇h gives

(4.30) $$-\frac{1}{c}(\nabla h \times \dot{\mathbf{H}}) + \frac{1}{c}\nabla h \times (\nabla h \times \dot{\mathbf{E}}) = \nabla h \times (\nabla \times \hat{\mathbf{E}})$$

(4.31) $$\frac{1}{c}(\nabla h \times \dot{\mathbf{E}}) + \frac{1}{c}\nabla h \times (\nabla h \times \dot{\mathbf{H}}) = \nabla h \times (\nabla \times \hat{\mathbf{H}})$$

Expanding the double cross product and substituting from (4.26) and (4.27), we have

(4.32) $$\frac{1}{c}\dot{\mathbf{E}} - \nabla \times \hat{\mathbf{H}} + \frac{1}{c}\nabla h(\nabla h \cdot \dot{\mathbf{E}}) - \frac{1}{c}\dot{\mathbf{E}}(\nabla h)^2 = \nabla h \times (\nabla \times \hat{\mathbf{E}})$$

(4.33) $$\frac{1}{c}\dot{\mathbf{H}} + \nabla \times \hat{\mathbf{E}} + \frac{1}{c}\nabla h(\nabla h \cdot \dot{\mathbf{H}}) - \frac{1}{c}\dot{\mathbf{H}}(\nabla h)^2 = \nabla h \times (\nabla \times \hat{\mathbf{H}})$$

Finally, substituting from (4.28) and (4.29), we get

(4.34) $$\frac{1}{c}\dot{\mathbf{E}} - \nabla \times \hat{\mathbf{H}} + \nabla h(\nabla h \cdot \nabla \times \hat{\mathbf{H}}) - \frac{1}{c}\dot{\mathbf{E}}(\nabla h)^2 = \nabla h \times (\nabla \times \hat{\mathbf{E}})$$

(4.35) $$\frac{1}{c}\dot{\mathbf{H}} + \nabla \times \hat{\mathbf{E}} - \nabla h(\nabla h \cdot \nabla \times \hat{\mathbf{E}}) - \frac{1}{c}\dot{\mathbf{H}}(\nabla h)^2 = \nabla h \times (\nabla \times \hat{\mathbf{H}})$$

Rearrangement now gives the two key relations that we have been working toward:

(4.36) $$\frac{1}{c}(1 - [\nabla h]^2)\dot{\mathbf{E}} = \nabla \times \hat{\mathbf{H}} - \nabla h(\nabla h \cdot \nabla \times \hat{\mathbf{H}}) + \nabla h \times (\nabla \times \hat{\mathbf{E}})$$

(4.37) $$\frac{1}{c}(1 - [\nabla h]^2)\dot{\mathbf{H}} = -\nabla \times \hat{\mathbf{E}} + \nabla h(\nabla h \cdot \nabla \times \hat{\mathbf{E}}) + \nabla h \times (\nabla \times \hat{\mathbf{H}})$$

Two situations are now possible: Either the factor $1 - (\nabla h)^2$ which appears on the left side of (4.36) and (4.37) is zero or it is nonzero. Consider first the case where it is nonzero. We can then solve for $\dot{\mathbf{E}}$ and $\dot{\mathbf{H}}$ in terms of the continuous first derivatives of h, $\hat{\mathbf{E}}$, and $\hat{\mathbf{H}}$, so $\dot{\mathbf{E}}$ and $\dot{\mathbf{H}}$ are themselves continuous on S. We have thus determined the values $\dot{\mathbf{E}}$ and $\dot{\mathbf{H}}$ on S which represent, in view of (4.17), the normal derivatives of the vector fields $\mathbf{E}$ and $\mathbf{H}$ on the hypersurface S. By the well-known theory of the initial-value problem for first-order differential systems, the solution fields $\mathbf{E}$ and $\mathbf{H}$ are therefore uniquely determined as continuously differentiable functions of all four variables in a neighborhood of the hypersurface S. Thus, if $1 - (\nabla h)^2$ is not zero, the first derivatives of

E and **H** must be continuous across S, and S cannot be a characteristic hypersurface.

In the second case, where $1 - (\nabla h)^2$ is zero, we cannot solve for $\dot{\mathbf{H}}$ and $\dot{\mathbf{E}}$, which therefore remain undefined on S. This is the only situation for which the first derivatives of **E** and **H** can be discontinuous across S, so we obtain the condition that, for S to be a characteristic, $1 - (\nabla h)^2$ must vanish. Note, furthermore, that if $1 - (\nabla h)^2$ vanishes, the equations (4.36) and (4.37) then provide a relation between $\hat{\mathbf{E}}$ and $\hat{\mathbf{H}}$; these functions cannot be arbitrarily prescribed on S. This is indeed a familiar property of characteristic surfaces: On a characteristic the fields must obey restrictive relations. (We shall see a further example of this property in Chap. 7, when we investigate the characteristic surfaces of Einstein's gravitational field equations.)

We now have the equation of the characteristics in the form

$$(\nabla h)^2 = 1 \tag{4.38}$$

which, by virtue of (4.17), is equivalent to

$$\left(\frac{\partial w}{\partial x^0}\right)^2 - (\nabla w)^2 = 0 \tag{4.39}$$

The reader may check that this is satisfied by the following class of hypersurfaces:

$$w = (x^0 - a^0)^2 - (x^1 - a^1)^2 - (x^2 - a^2)^2 - (x^3 - a^3)^2 = 0 \tag{4.40}$$

where the a's are arbitrary parameters. Indeed, one can also consider the a's to be functions of a single parameter λ. A continuous or discrete superposition of solutions of the form (4.40) is then easily made, and one can show that the envelopes of such superpositions constitute all the continuous and differentiable solutions of (4.39). In the theory of partial differential equations, a particular solution with the above outlined property is termed a *complete integral*. In the case of Maxwell's equations the complete integral (4.40) is the most important particular solution to (4.39).

By translating the origin of the coordinate system we can write the complete integral (4.40) as

$$(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = c^2t^2 - (x^2 + y^2 + z^2) = 0 \tag{4.41}$$

which clearly represents a three-dimensional sphere expanding at velocity c, or the four-dimensional *light cone*, which is familiar from special relativity. (In Chap. 8 characteristic surfaces will be discussed further.)

In (4.41) we have the equation of the most important characteristic written as a null quadratic form. In matrix notation it takes the form

$$(x^0x^1x^2x^3)\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}\begin{pmatrix}x^0\\ x^1\\ x^2\\ x^3\end{pmatrix} = 0 \tag{4.42}$$

We have here a quadratic form and its associated matrix which plays a very distinguished role in the propagation of electromagnetic wave fronts. On the other hand, the metric tensor is the associated matrix of another quadratic form, the line element, which also plays a fundamental role in the structure of a Riemann space. Thus, unless the matrix appearing in (4.42) and the metric tensor are the same, we are faced with the rather unsatisfactory situation of possessing two unrelated fundamental matrices. To avoid this situation and maintain the maximum amount of elegance and economy, we tentatively adopt the matrix appearing in (4.42) as the metric tensor in the ct, x, y, z coordinate system. The line element and the metric tensor are thus taken to be

(4.43)

$$ds^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 \qquad g_{\mu\nu} = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}$$

These are the familiar line element and metric tensor of special relativity which one usually obtains by quite different arguments. But let us note once again that the choice of this metric was here motivated directly by arguments based on the mathematical structure of the Maxwell equations, which is completely independent of all mechanical concepts.

It is now our task to show that the above metric is indeed a wise choice. To do this we shall rephrase Maxwell's equations in formal tensor notation. The elegance of the resultant equations will serve as justification for the choice. Further justification will also appear in Secs. 4.2 and 4.3, when we consider the consequences of also using this metric for the description of mechanical systems.

Let us begin the formal rephrasing of Maxwell's equations by defining a source four-vector.

$$s^\alpha = \left(\rho, \frac{1}{c}j_x, \frac{1}{c}j_y, \frac{1}{c}j_z\right) = \left(\rho, \frac{1}{c}\mathbf{j}\right) \tag{4.44}$$

The conservation of charge, Eq. (4.7), can be written using tensor notation as

(4.45) $$c\rho_{|0} + j^i{}_{|i} = 0$$

and in terms of the source vector s^α as

(4.46) $$s^\alpha{}_{|\alpha} = 0$$

Since the Christoffel symbols are zero in the coordinate system we are now using, we can just as well use the notation for *covariant* differentiation as ordinary differentiation. This switch to covariant differentiation is also advantageous if we wish to work in polar coordinates or in various other spatial curvilinear coordinate systems instead of rectangular coordinates, for then the equation needs no modification in the curvilinear system. Thus we write

(4.47) $$s^\alpha{}_{\|\alpha} = 0$$

When we later consider the transformation of the equations to a completely general coordinate system (in arbitrary motion), the use of covariant instead of ordinary differentiation will be of the utmost utility, as the reader has no doubt anticipated.

To write the source equations (4.1) and (4.2) in tensor notation, we define a 4×4 antisymmetric matrix composed of the components of the electric and magnetic fields in the following manner:

(4.48) $$F^{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & H_z & -H_y \\ -E_y & -H_z & 0 & H_x \\ -E_z & H_y & -H_x & 0 \end{pmatrix}$$

This is the Minkowski *electromagnetic field tensor*. (At this point, of course, we are not yet justified in calling it a tensor, but only a matrix.) Using this matrix we can write the source equations as

(4.49) $$F^{\mu\nu}{}_{|\nu} = s^\mu \qquad F^{\mu\nu}{}_{\|\nu} = s^\mu$$

where we have again used covariant differentiation instead of ordinary differentiation, for the same reasons as before. The easiest method of verifying that (4.49) is equivalent to the two Maxwell equations (4.1) and (4.2) is simply to substitute appropriate values of μ and ν and check that

(4.1) and (4.2) result. For $\mu = 0$, we have

(4.50) $$F^{00}{}_{|0} + F^{01}{}_{|1} + F^{02}{}_{|2} + F^{03}{}_{|3} = s^0$$

or equivalently,

(4.51) $$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = \rho$$

which is precisely the Maxwell equation (4.2). For $\mu = 1$, we have

(4.52) $$F^{10}{}_{|0} + F^{11}{}_{|1} + F^{12}{}_{|2} + F^{13}{}_{|3} = s^1$$

or equivalently,

(4.53) $$-\frac{1}{c}\dot{E}_x + \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} = \frac{1}{c} j_x$$

which is the x component of the Maxwell equation (4.1). The other components of (4.1) may be similarly checked.

In like manner the reader may satisfy himself that in tensor notation the internal Maxwell equations (4.3) and (4.4) are equivalent to

(4.54) $$F_{\mu\nu|\lambda} + F_{\lambda\mu|\nu} + F_{\nu\lambda|\mu} = 0$$

or, using the antisymmetrization notation introduced in Chap. 3,

(4.55*a*) $$\{F_{\mu\nu|\lambda}\} = 0$$

Here we need *not* replace ordinary by covariant differentiation since (4.55*a*) is already in covariant form, as we showed in Sec. 3.4. As we noted in Sec. 3.5, this is also equivalent to the dual tensor equation

(4.55*b*) $$*F^{\mu\nu}{}_{\|\nu} = 0$$

These formulas are best understood if we write out explicitly the dual matrix

(4.56*a*) $$*F_{\mu\nu} = \begin{pmatrix} 0 & H_x & H_y & H_z \\ -H_x & 0 & E_z & -E_y \\ -H_y & -E_z & 0 & E_x \\ -H_z & E_y & -E_x & 0 \end{pmatrix}$$

which arises from $F_{\mu\nu}$ by interchange of **E** and **H**. If we construct the contravariant dual tensor, using $g^{\mu\nu} = g_{\mu\nu}$ in (4.42), we obtain

$$*F^{\mu\nu} = \begin{pmatrix} 0 & -H_x & -H_y & -H_z \\ H_x & 0 & E_z & -E_y \\ H_y & -E_z & 0 & E_x \\ H_z & E_y & -E_x & 0 \end{pmatrix} \tag{4.56b}$$

That is, $*F^{\mu\nu}$ differs from $F^{\mu\nu}$ by the change of **E** into $-$**H** and **H** into **E**. The differential expressions $(F^{i\nu}{}_{|\nu}) = \nabla \times \mathbf{H} - (1/c)\dot{\mathbf{E}}$ therefore go over to $(*F^{i\nu}{}_{|\nu}) = \nabla \times \mathbf{E} + (1/c)\dot{\mathbf{H}}$.

Equation (4.55*a*) is, as we found in Sec. 3.5, simply the necessary and sufficient condition that $F_{\mu\nu}$ is closed and has a tensor potential. That is, there exists a four-vector ϕ_μ such that

$$F_{\mu\nu} = \phi_{\mu|\nu} - \phi_{\nu|\mu} \tag{4.57a}$$

Now, in classical electrodynamics, the only physically significant quantities are the electric and magnetic fields **E** and **H** which appear in the field tensor $F_{\mu\nu}$. Therefore the four-vector function ϕ_μ has no *direct* physical meaning; only its four-dimensional curl has physical meaning. It is thus clear that we may make a so-called *gauge transformation* on ϕ_μ; that is, we can add an arbitrary four-dimensional gradient $\Psi_{|\mu}$ to ϕ_μ, without altering $F_{\mu\nu}$ and therefore without altering the physical situation. That is,

$$(\phi_\mu + \Psi_{|\mu})_{|\nu} - (\phi_\nu + \Psi_{|\nu})_{|\mu} = \phi_{\mu|\nu} - \phi_{\nu|\mu} = F_{\mu\nu} \tag{4.57b}$$

We say, then, that the physically meaningful $F_{\mu\nu}$ tensor is *gauge-invariant;* i.e., it is not altered by a gauge transformation.

We have now obtained the Maxwell equations and their associated conservation law, all written formally in tensor notation but taken to be valid only in some specified rest system.

$$\begin{aligned} F^{\mu\nu}{}_{\|\nu} &= s^\mu \\ *F^{\mu\nu}{}_{\|\nu} &= 0 \qquad \text{or} \qquad \{F_{\mu\nu|\lambda}\} = 0 \\ s^\mu{}_{\|\mu} &= 0 \end{aligned} \tag{4.58}$$

In classical electrodynamics there is *no a priori rule for transforming the fields and the field equations* to a system in motion relative to the original rest frame. The rules proposed for the transformation previous to the application of relativity to electrodynamics were rather unconvincing. However, with the equations written in *tensor notation*, we are presented

with a more convincing solution to the problem. The temptation to postulate that the Maxwell equations in tensor notation are indeed *tensor equations* is very great, and we accordingly make this assumption. Then, under a general coordinate transformation, the *electric and magnetic fields transform as the indicated components of* $F_{\mu\nu}$, which is now a tensor, and the equations for the fields remain the same in all systems. The transformation properties of the fields which we postulate by this assumption can be tested in experiments involving moving material media and, indeed, agree quite well with experiment.

When Minkowski first introduced the $F_{\mu\nu}$ tensor into electrodynamics, he had in mind that it should transform as a tensor only under Lorentz transformations. However, as we see here, no such restriction appears necessary, for the Maxwell equations go over very easily indeed into a completely covariant form.

The reader should not be blinded by our mathematical transformations into assuming that the statement "$F^{\mu\nu}$ is a tensor" is a purely mathematical one. It is a very important and far-reaching physical principle which can be *motivated* by mathematical elegance, but must also be *tested* by physical experiment. The fact that $F^{\mu\nu}$ is a tensor under Lorentz transformation embodies a large part of special relativity theory. Our assumption that $F^{\mu\nu}$ is a tensor in a general Riemannian space-time leads to important consequences for the electrodynamics in accelerated systems of reference. The methodological principle that laws of nature which appear in tensor form in a particular coordinate system should be interpreted as valid in every system is called the *principle of covariance.* Its philosophical motivation is the postulate that no coordinate system should be distinguished in the formation of physical laws. It is, however, mathematically somewhat ambiguous and comes in practice to the old principle that we should try to explain facts with the simplest and most aesthetically satisfactory theory.

Now that we assume the Maxwell equations are tensor equations, we can investigate their form further. We note that the internal equations (4.55*a*) contain no terms involving the metric tensor. If the role of the metric tensor in the source equations (4.49) could be reduced to a minimum, these equations might be somewhat simplified. To see this let us write them out explicitly, working now in an arbitrary coordinate system:

$$F^{\mu\nu}{}_{\|\nu} = F^{\mu\nu}{}_{|\nu} + \begin{Bmatrix} \nu \\ \nu\ \alpha \end{Bmatrix} F^{\mu\alpha} + \begin{Bmatrix} \mu \\ \alpha\ \nu \end{Bmatrix} F^{\alpha\nu} = s^\mu \tag{4.59}$$

We found in (3.11) that the "contracted" Christoffel symbol can be written as

$$\begin{Bmatrix} \nu \\ \nu\ \alpha \end{Bmatrix} = \frac{1}{2g} g_{|\alpha} \tag{4.60}$$

where g is the determinant of the metric tensor. Next note that $F^{\alpha\nu}$ is antisymmetric while $\left\{ {\mu \atop \alpha\ \nu} \right\}$ is symmetric in α and ν. Thus the sum over α and ν of the product will cancel in pairs to give a zero result:

$$\left\{ {\mu \atop \alpha\ \nu} \right\} F^{\alpha\nu} = 0 \tag{4.61}$$

Therefore, (4.59) can be written

$$F^{\mu\nu}{}_{\|\nu} = F^{\mu\nu}{}_{|\nu} + \frac{1}{2g}(g_{|\alpha})F^{\mu\alpha} = s^\mu \tag{4.62}$$

or equivalently,

$$F^{\mu\nu}{}_{\|\nu} = F^{\mu\nu}{}_{|\nu} + \frac{1}{\sqrt{-g}}(\sqrt{-g})_{|\alpha}F^{\mu\alpha} = \frac{1}{\sqrt{-g}}(\sqrt{-g}\,F^{\mu\nu})_{|\nu} = s^\mu \tag{4.63}$$

In this form only the single quantity g enters the equation instead of the 16 components of $g_{\mu\nu}$ as in (4.49).

If we denote the antisymmetric tensor density (Sec. 3.5) associated with the $F^{\mu\nu}$ field tensor as

$$\mathfrak{F}^{\mu\nu} = \sqrt{-g}\,F^{\mu\nu} \tag{4.64}$$

and the density associated with s^μ as

$$\mathfrak{S}^\mu = \sqrt{-g}\,s^\mu \tag{4.65}$$

we can write (4.63) in another convenient form:

$$\mathfrak{F}^{\mu\nu}{}_{|\nu} = \mathfrak{S}^\mu \tag{4.66}$$

As an exercise to illustrate the great advantage of working with the Maxwell equations in the above tensor form, let us derive the charge conservation equation directly from (4.66). Differentiation with respect to x^μ gives

$$\mathfrak{F}^{\mu\nu}{}_{|\nu|\mu} = \mathfrak{S}^\mu{}_{|\mu} \tag{4.67}$$

Since $\mathfrak{F}^{\mu\nu}$ is antisymmetric in μ and ν while the differential operator $\partial^2/(\partial x^\nu \partial x^\mu)$ is symmetric, the sum on the left vanishes, and we obtain

$$\mathfrak{S}^\mu{}_{|\mu} = (\sqrt{-g}\,s^\mu)_{|\mu} = 0 \tag{4.68}$$

From (3.12) we have

$$s^\mu{}_{\|\mu} = \frac{1}{\sqrt{-g}}(\sqrt{-g}\,s^\mu)_{|\mu} = \frac{1}{\sqrt{-g}}\mathfrak{S}^\mu{}_{|\mu} \tag{4.69}$$

so since g is never zero, we obtain, from Eqs. (4.68) and (4.69), the conservation equation (4.47).

As a second illustration of the ease of working with the tensor form of Maxwell's equations, let us consider a "scale change" of the metric tensor: $g_{\mu\nu} \to A(x^\alpha)g_{\mu\nu}$, where $A(x^\alpha)$ is an arbitrary function of position. (It should be noted that such a scale change is completely unrelated to an ordinary transformation of coordinates.) We shall show that Maxwell's equations in free space are actually invariant under such a scale change. In order to show this, we write Maxwell's equations in the form displayed in (4.57*a*) and (4.66):

$$F_{\mu\nu} = \phi(x^\alpha)_{\mu|\nu} - \phi(x^\alpha)_{\nu|\mu} \qquad (\text{equivalently, } \{F_{\mu\nu|\lambda}\} = 0) \tag{4.70a}$$

$$\mathfrak{F}^{\mu\nu}{}_{|\nu} = (\sqrt{-g}\,F^{\mu\nu})_{|\nu} = 0 \tag{4.70b}$$

A scale change of the metric then replaces $g_{\mu\nu}$ by

$$\tilde{g}_{\mu\nu} = Ag_{\mu\nu} \tag{4.71}$$

The tensor $\tilde{g}^{\mu\nu}$ is defined as the inverse of $\tilde{g}_{\mu\nu}$ and is therefore clearly given by

$$\tilde{g}^{\mu\nu} = \frac{1}{A}g^{\mu\nu} \tag{4.72}$$

Under the scale change defined by (4.71) and (4.72), the four-vector $\phi(x^\alpha)_\mu$ can be consistently considered (by definition) to remain unchanged since it is simply a function of position:

$$\tilde{\phi}(x^\alpha)_\mu = \phi(x^\alpha)_\mu \tag{4.73}$$

This implies that $F_{\mu\nu}$ is unchanged also:

$$\tilde{F}_{\mu\nu} = \tilde{\phi}_{\mu|\nu} - \tilde{\phi}_{\nu|\mu} = \phi_{\mu|\nu} - \phi_{\nu|\mu} = F_{\mu\nu} \tag{4.74}$$

and that equation (4.70*a*) is therefore invariant under a change of the metric scale.

To obtain the doubly contravariant tensor $\tilde{F}^{\mu\nu}$ which appears in (4.70*b*), we must raise both indices of $\tilde{F}_{\mu\nu}$ (4.74), using $\tilde{g}^{\mu\nu}$ in (4.72). This gives

$$\tilde{F}^{\mu\nu} = \tilde{g}^{\mu\alpha}\tilde{g}^{\nu\beta}\tilde{F}_{\alpha\beta} = \frac{1}{A^2} g^{\mu\alpha}g^{\nu\beta}F_{\alpha\beta} = \frac{1}{A^2}F^{\mu\nu} \tag{4.75}$$

Combining this with the following expression for $\sqrt{-\tilde{g}}$, which follows from the definition (4.71),

$$\sqrt{-\tilde{g}} = \sqrt{-A^4 g} = A^2\sqrt{-g} \tag{4.76}$$

we see that

$$\sqrt{-\tilde{g}}\,\tilde{F}^{\mu\nu} = \sqrt{-g}\,F^{\mu\nu} \tag{4.77}$$

That is, the tensor density $\mathfrak{F}^{\mu\nu} = \sqrt{-g}\,F^{\mu\nu}$ is invariant; it then follows immediately that Eq. (4.70*b*) is also invariant under the scale change (4.71).

4.2 Proper-Time and the Equations of Motion via an Example in Relativistic Mechanics

In the previous section we dealt with the formulation of classical electrodynamics in a four-dimensional Riemann space; it did not prove necessary to consider questions of measurement concerning meter sticks and clocks and their relation to the four-dimensional Riemann space. Nor were we concerned with the paths of particles in the Riemann space, since we dealt only with tensor *fields*. The above are all specifically mechanical concepts, and, as we have noted, the results of the preceding section were purposely obtained in a way which was independent of mechanical notions. In this section we shall consider these mechanical questions, which we have so far carefully avoided.

In the hope of keeping the present development well grounded in familiar physical concepts, we shall consider the specific example of a coordinate system rotating at constant angular velocity and investigate the "fictitious" Coriolis and centrifugal forces associated with the rotation. Reasons for considering this example are twofold: First, we know that the equivalence principle asserts that the "fictitious" force due to acceleration of the coordinate system and the "real" force of gravity are in essence the same sort of phenomena. By considering these simple examples of fictitious forces we can hope to learn something of how forces

in general occur in the context of relativity theory. Second, the investigation of this example will lead naturally to a reasonable solution of the general mechanical questions we cited above—the nature of space-time measurements and the equations of motion for particles in a Riemann space.

We begin as in the previous section with an inertial system and the coordinates of special relativity, ct, x, y, z. Because of the success of using the Lorentz metric to formulate the free-space Maxwell equations in a four-dimensional Riemann space, we shall tentatively carry it over (in the spirit of elegance and economy) to use also in describing mechanical phenomena in a mass-free Riemann space. The line element and metric are then assumed to be those in (4.43):

$$\begin{aligned} ds^2 &= (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 \\ &= c^2\,dt^2 - (dx^2 + dy^2 + dz^2) \end{aligned}$$

$$g_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \tag{4.78}$$

This particular metric tensor will from now on be called the Lorentz metric tensor, and the corresponding metric, the Lorentz metric.

It will be convenient in this section to work with cylindrical coordinates instead of the Cartesian coordinates used in (4.78). In terms of cylindrical coordinates r, $\bar{\varphi}$, and z, we have

$$x = r\cos\bar{\varphi} \qquad y = r\sin\bar{\varphi} \qquad z = z \tag{4.79}$$

and

$$dx^2 + dy^2 + dz^2 = dr^2 + r^2\,d\bar{\varphi}^2 + dz^2 \tag{4.80}$$

Thus the Lorentz metric in cylindrical coordinates is

$$ds^2 = c^2\,dt^2 - (dr^2 + r^2\,d\bar{\varphi}^2 + dz^2) \tag{4.81}$$

We now define a transformation to a new t, r, φ, z system rotating about the z axis of the above inertial system with angular velocity w. To visualize the situation, one may think of the new system as being attached to a material disk which rotates with respect to the original coordinate system. This involves a change in $\bar{\varphi}$ only, which is clearly

$$\varphi = \bar{\varphi} - wt \tag{4.82}$$

The line element in this rotating system is easily obtained from (4.81) by the use of (4.82):

$$\begin{aligned} ds^2 &= c^2\,dt^2 - [dr^2 + r^2\,d\varphi^2 + 2wr^2\,d\varphi\,dt + w^2r^2\,dt^2 + dz^2] \\ &= (c^2 - w^2r^2)\,dt^2 - (dr^2 + r^2\,d\varphi^2 + 2wr^2\,d\varphi\,dt + dz^2) \end{aligned} \tag{4.83}$$

At this point we have obtained "complete" knowledge of the abstract geometry of space in the rotating coordinate system in the sense that the metric and the line element are known functions of the coordinates. We can at present go no further toward our goal of a description of physical measurements and mechanical processes in the rotating system, since we as yet have no way of relating the geometry of the space to clocks, measuring rods, and the motion of particles. To continue we must *interpret* the abstract geometry (embodied in the functional form of the line element) in mechanistic physical terms.

To interpret the abstract four-dimensional geometry and link it to reality by identifying a physical measurement with the evaluation of a geometrical object, we first need to define a few geometrical terms which one uses in four-space. An event is a point in four-space: a world-point. A single infinity of events forms a curve which we shall call a world-line or history if we are dealing with the representative point of a physical particle. The arc length between two events along a world-line is a geometrical invariant.

Let us now investigate what we mean by a measured proper-time interval in the special theory of relativity and obtain its relation to the line element ds^2. Suppose the line element between two events is $ds^2 = c^2\,dt^2 - (dx^2 + dy^2 + dz^2) \geq 0$, that is, a timelike interval. If we choose the Lorentz system in which the three-dimensional separation between the events is zero (the *proper* system), then

$$ds^2 = c^2\,dt^2 \tag{4.84}$$

Choosing the positive root, we then have

$$dt = \frac{ds}{c} \tag{4.85}$$

The interval ds/c is termed the proper-time interval between the events and corresponds, by the above comments, to the time interval that would be measured by a physicist to whom both events occurred at the same point in a three-dimensional frame to which he is attached. In practice

this corresponds, for instance, to a physicist measuring the proper lifetime of a μ-meson by riding with it from the event corresponding to the creation of the meson to the event corresponding to its decay. Now, in general relativity, we wish to define the proper-time interval between two infinitesimally close events in some *invariant* manner which agrees with the above expression for the proper-time interval in the special case of the Lorentz metric. Thus it is very reasonable to make the following definition: In general relativity, an infinitesimal proper-time interval between two neighboring events is defined as the invariant generalization of (4.85) to an arbitrary Riemann space:

$$d\tau \equiv \frac{ds}{c} = \frac{1}{c}\sqrt{g_{00}}\,dx^0 \qquad \text{where } dx^\mu = 0 \text{ for } \mu \neq 0 \tag{4.86}$$

Notice that this definition of proper-time interval has been given only for infinitesimals and that it is physically meaningful only for $ds^2 \geq 0$, since otherwise $d\tau$ would be imaginary.

To illustrate our definition of proper-time intervals, let us investigate the relation of the proper-time interval $d\tau$ to the coordinate-time interval dt for two events which occur with the same spatial coordinates r, φ, z in the rotating coordinate system in our example above. We set the coordinate intervals dr, $d\varphi$, and dz equal to zero in (4.83) and obtain

$$ds^2 = (c^2 - w^2r^2)\,dt^2 \tag{4.87}$$

Thus, by the proper-time definition (4.86),

$$d\tau = \frac{ds}{c} = \left(1 - \frac{w^2r^2}{c^2}\right)^{1/2} dt \tag{4.88}$$

This can be interpreted as follows: $d\tau$ is the time interval between the events as measured by an observer attached to the rotating disk. On the other hand, dt is the time interval as measured by an observer attached to the nonrotating coordinate system and who uses the standard coordinate time t. How the physical measurements of time are carried out in both cases will be elucidated when we talk of finite time intervals a few paragraphs later. Thus (4.88) tells us that the time interval between the two events is different for these two observers, and indeed the ratio is $\sqrt{1 - w^2r^2/c^2}$. We can easily verify that this relation, which we obtained by an interpretation of the abstract geometry of the rotating system, is in agreement with the notions of special relativity. The linear velocity of a fixed point in the rotating coordinate system is given by

$v = rw$, so we can write the proper-time relation as

$$d\tau = \left(1 - \frac{v^2}{c^2}\right)^{1/2} dt \tag{4.89}$$

which we recognize as the familiar Lorentz time-dilation equation. It should be noted that this interpretation places an upper limit on the useful range of values of r; unless r is less than c/w, a fixed point in the rotating coordinate system will exceed the velocity of light, making that system inaccessible to a physical observer. [Of course as a formal mathematical operation the transformation (4.82) may still be made.]

Let us next apply (4.83) to measure three-dimensional space intervals. The cross term $-2wr^2\,d\varphi\,dt$ presents here a mathematical difficulty. We are used to line elements ds^2, which are quadratic in all differentials of our space geometry, and the cross term violates this rule. We must thus introduce a new time differential in the rotating system which removes this term. We are forced to a redefinition of simultaneity:

$$dt^* = dt - \frac{wr^2}{c^2 - w^2r^2}\,d\varphi \tag{4.90}$$

The new choice of the time differential brings (4.83) into the form

$$ds^2 = (c^2 - w^2r^2)\,dt^{*2} - \left(dr^2 + \frac{c^2}{c^2 - w^2r^2}\,r^2\,d\varphi^2 + dz^2\right) \tag{4.91}$$

The intervals dr, $d\varphi$, and dz corresponding to simultaneity according to $dt^* = 0$ are the intervals which a physicist would measure on the rotating frame. We read off from (4.91) that, with this understanding, a correct longitudinal Lorentz dilatation takes place, while scales transverse to the local velocity remain unchanged. In general, a moving observer will try to define his local space-time intervals in such a way that the line element is *locally* Lorentzian, as in (4.91). Observe, however, that the differential (4.90) is not exact; that is, we cannot introduce markers t^*, r, φ, z *in the large* which would lead to the line element (4.91).

Let us investigate further the general properties of proper-time intervals and coordinate-time intervals and the relative usefulness of the two concepts. We now extend the definition of proper-time to finite intervals; a proper-time interval is defined to be the invariant proportional (with coefficient $1/c$) to the arc length along a world-line and therefore has the very desirable property that it is independent of any reference frame. It will be useful in formulating basic physical laws and in describing fundamental processes if we identify it, considered as a geometrical object, with a physically measurable time interval. This is always implicitly done by relativists when considering that physical time is measured by the ticking of an atomic oscillator and that the period of such an oscillator is an invariant that remains the same during the history of the atom considered (is independent of the atom's age and position in three-space). This identification amounts to a basic postulate in the theory of relativity. It has been called by Synge (Synge, 1956) the *chronometric hypothesis* and is stated most clearly using a four-dimensional picture: On the world-line of a material particle there exists or can be thought to exist a discrete set of events separated by equal proper-time intervals. These events can be created by conceiving a standard atomic clock carried by the material particle. Furthermore, this postulate can have meaning only if it does not depend on the type of atomic clock used; therefore one needs to make the consistency hypothesis that, for a fixed arc length of an arbitrary common world-line, the ratio of the number of ticks of two atomic clocks is a natural constant.

From the above paragraph it is clear that an infinitesimal proper-time interval is a very useful concept. However, let us now show that the notion of proper-time *in the large* as opposed to an infinitesimal proper-time interval meets with difficulty and is not such a useful notion. The total elapsed proper-time between widely separated events which we define as the integral

$$\int_{\text{event 1}}^{\text{event 2}} \frac{ds}{c} = \frac{s}{c} = \tau \tag{4.92}$$

is clearly dependent on the *path* of integration, i.e., on the world-line followed by the standard clock between events 1 and 2. Therefore, if we specify an initial zero point of proper-time at some space-time point x_0^μ, we cannot extend a set of proper-time values throughout space in a unique way and we cannot label each event with one proper-time value. We say, then, that proper-time is not integrable; we cannot uniquely extend a proper-time over the space-time manifold. This fact gives rise to the well-known twin paradox in the special theory of relativity. The paradoxical nature of an age difference between the twins disappears once one stops thinking in terms of Newtonian pictures; following our definition of proper-time (age), one has to compare the arc lengths of the two different world-lines (histories) of the twins in a four-dimensional diagram. These two lengths have no reason to be equal.

The situation with the coordinate-time is in a sense opposite to that with proper-time. The coordinate-time interval between events is not an invariant, but by the definition of a Riemann space, each space-time point possesses a unique coordinate-time label. Thus the coordinate-time has unambiguous meaning at each point in space-time, and is hence an integrable quantity. It follows that, when we speak of *widely separated events*, the concept of coordinate-time is useful, for the coordinate-time separation between any two events—however widely separated—is unique and well defined in any given coordinate system. For example, if the period of an atomic oscillator is expressed in coordinate-time as dt by an observer at some point in space-time, it is the same coordinate-time interval dt to all observers throughout space-time. We may sum up: Proper-time intervals $d\tau$ are coordinate-invariant, but not integrable, whereas coordinate-time intervals dt are integrable, but not invariant.

The concepts of proper-time and coordinate-time which we have discussed in the foregoing paragraphs often prove to complement each other in practice in a very satisfying way. As an example, suppose we wish to compare "corresponding" proper-time intervals at different points in space. We can proceed as follows: At some initial point in space-time x_0^μ we specify the duration of a physical event in terms of an invariant, measurable proper-time interval $d\tau(x_0^\mu)$. The g_{00} element of the metric tensor then allows us to relate this to a *corresponding* coordinate-time interval dt by means of the defining relation

(4.93) $$d\tau(x_0^\mu) = \sqrt{g_{00}(x_0^\mu)}\, dt$$

This dt, being an integrable coordinate-time interval, has unique meaning throughout space: to $d\tau(x_0^\mu)$ there corresponds only one value of dt which is the same throughout space. However, at a distant point x^μ this dt corresponds to quite another proper-time interval $d\tau(x^\mu)$, which is given by a relation analogous to (4.93):

(4.94) $$d\tau(x^\mu) = \sqrt{g_{00}(x^\mu)}\, dt$$

Thus the two proper-time intervals $d\tau(x_0^\mu)$ and $d\tau(x^\mu)$, which have only local meaning but correspond to the same coordinate-time interval dt, are related to each other by

(4.95) $$\frac{d\tau(x^\mu)}{d\tau(x_0^\mu)} = \left(\frac{g_{00}(x^\mu)}{g_{00}(x_0^\mu)}\right)^{1/2}$$

as follows directly from (4.93) and (4.94). This relation between corresponding proper-time intervals is the basis for calculations of the red shift, and we shall return to it in Sec. 4.4.

Let us now move on to the second basic problem of this section: the equations of motion for a particle in a four-dimensional Riemann space.

We showed in Chap. 2 that the force-free motion of a classical particle or a system of classical particles can be expressed in differential geometric form. Specifically, we showed that if x^i are the generalized coordinates of a system of n degrees of freedom, and if the kinetic energy of the system is of the form

(4.96) $$T = \frac{1}{2}\sum_{\substack{i=1\\k=1}}^{n} g_{ik}\frac{dx^i}{dt}\frac{dx^k}{dt}$$

(where g_{ik} is independent of time), then the force-free motion of the system occurs along the extremal curves or configuration-space geodesics of the variational problem

(4.97) $$\delta\int_{t_i}^{t_f} T\,dt = \delta\int_{t_i}^{t_f}\frac{1}{2}\sum_{\substack{i=1\\k=1}}^{n} g_{ik}\frac{dx^i}{dt}\frac{dx^k}{dt}\,dt$$

In terms of the line element,

(4.98) $$ds^2 = g_{ik}\,dx^i\,dx^k$$

this is expressible in differential geometric form as

(4.99) $$\delta\int_{t_i}^{t_f} T\,dt = \delta\int_{t_i}^{t_f}\frac{1}{2}\left(\frac{ds}{dt}\right)^2 dt = 0$$

For the mechanical developments in Chap. 2, we treated time as an invariant scalar and not as a coordinate. In relativity theory, on the other hand, we wish to treat time as an additional coordinate and use a line element involving four coordinate intervals. To pursue this further, let us consider a line element very similar to the classical form (4.98) but involving also a time interval dt·

(4.100) $$ds^2 = c^2\,dt^2 - g_{ik}\,dx^i\,dx^k$$

The line element corresponding to the Lorentz metric expressed in an inertial system is of precisely this form. We shall now show that the variational problem using this line element,

(4.101) $$\delta\int_{s_i}^{s_f} ds = \delta\int_{s_i}^{s_f}\left[c^2\left(\frac{dt}{ds}\right)^2 - g_{ik}\frac{dx^i}{ds}\frac{dx^k}{ds}\right]^{1/2} ds = 0$$

reduces to the nonrelativistic form (4.99) and therefore is an invariant representation of the force-free motion of a classical dynamical system. Since we are using s as parameter, we can just as well square the integrand (as we discussed in Sec. 2.3) and consider the equivalent variational problem

$$\delta \int \left\{ c^2 \left(\frac{dt}{ds}\right)^2 - g_{ik} \frac{dx^i}{ds}\frac{dx^k}{ds} \right\} ds = 0 \tag{4.102}$$

The Euler-Lagrange equations for this variational problem are

$$\frac{\partial L}{\partial x^\mu} - \frac{d}{ds}\left(\frac{\partial L}{\partial \dot{x}^\mu}\right) = 0 \tag{4.103}$$

where L is the integrand and $\dot{x}^\mu = dx^\mu/ds$. For $x^\mu = ct$, we have the equation

$$\frac{d}{ds}\left(2c\frac{dt}{ds}\right) = 0 \tag{4.104}$$

so that

$$\frac{dt}{ds} = \text{const} \tag{4.105}$$

By multiplying the coordinate time (which is just a marker) by an arbitrary constant we do not alter any physics, so we can choose the constant in Eq. (4.105) to be 1. This gives

$$dt = ds \tag{4.106}$$

allowing us to replace ds in the variational problem (4.101) by dt. The problem then becomes

$$\delta \int_{t_i}^{t_f} \left[c^2 - g_{ik}\frac{dx^i}{dt}\frac{dx^k}{dt} \right] dt = -\delta \int_{t_i}^{t_f} \left[g_{ik}\frac{dx^i}{dt}\frac{dx^k}{dt} \right] dt = 0 \tag{4.107}$$

which is, as we wished to show, the same as the nonrelativistic form (4.99).

Let us investigate the results of tentatively applying the geodesic equations as equations of motion in the example of the rotating coordinate system in which the line element is somewhat more complicated than (4.100). Using the form of the line element given in (4.83), the variational problem for the geodesics is

$$0 = \delta \int ds = \delta \int \left\{ (c^2 - r^2w^2)\left(\frac{dt}{ds}\right)^2 - \left[\left(\frac{dr}{ds}\right)^2 + r^2\left(\frac{d\varphi}{ds}\right)^2 + 2wr^2\frac{d\varphi}{ds}\frac{dt}{ds} + \left(\frac{dz}{ds}\right)^2\right]\right\}^{1/2} ds \tag{4.108}$$

If we again use s as a parameter, the variational problem can also be written in the equivalent form

$$0 = \delta \int \left\{ (c^2 - r^2w^2)\left(\frac{dt}{ds}\right)^2 - \left[\left(\frac{dr}{ds}\right)^2 + r^2\left(\frac{d\varphi}{ds}\right)^2 + 2wr^2\frac{d\varphi}{ds}\frac{dt}{ds} + \left(\frac{dz}{ds}\right)^2\right]\right\} ds \tag{4.109}$$

Accordingly, the Euler-Lagrange equation for $x^\mu = r$ is

$$-2rw^2\dot{t}^2 - 2r\dot{\varphi}^2 - 4wr\dot{\varphi}\dot{t} = -2\ddot{r} \tag{4.110}$$

Rearrangement gives

$$\ddot{r} = rw^2\dot{t}^2 + r\dot{\varphi}^2 + 2wr\dot{\varphi}\dot{t} \tag{4.111}$$

Similarly, the Euler-Lagrange equation for $x^\mu = z$ is

$$\ddot{z} = 0 \tag{4.112}$$

and that for $x^\mu = \varphi$ is

$$\frac{d}{ds}[r^2\dot{\varphi} + wr^2\dot{t}] = 0 \tag{4.113}$$

This last differential equation yields immediately

$$r^2\dot{\varphi} + wr^2\dot{t} = \text{const} \tag{4.114}$$

Finally, the Euler-Lagrange equation for $x^\mu = t$ is

$$\frac{d}{ds}[(c^2 - w^2r^2)\dot{t} - wr^2\dot{\varphi}] = 0 \tag{4.115}$$

which is also solved at once to give

$$(c^2 - r^2w^2)\dot{t} - wr^2\dot{\varphi} = \text{const} \tag{4.116}$$

These four Euler-Lagrange equations can now be shown to yield the familiar fictitious forces associated with a rotating system.

To show this, let us begin by multiplying (4.114) by the constant w:

$$wr^2\dot{\varphi} + w^2r^2\dot{t} = \text{const} \tag{4.117}$$

Adding this to (4.116), we have

$$c^2\dot{t} = \text{const} \tag{4.118}$$

By suitably stretching or compressing the time scale, which has no physical significance, we can make the constant in (4.118) equal to c; then

$$\dot{t} = \frac{1}{c} \tag{4.119}$$

To display the centrifugal force we consider a momentary radial motion in the rotating system; that is, we set $\dot{\varphi} = 0$ and insert (4.119) in Eq. (4.111) to get

$$\ddot{r} = \frac{rw^2}{c^2} \tag{4.120}$$

Using the proper-time interval of the particle in the barred system $d\tau = ds/c$, we then have

$$\frac{d^2r}{d\tau^2} = rw^2 \tag{4.121}$$

which we recognize as the familiar classical expression for centrifugal acceleration, except that the proper-time interval $d\tau$ replaces the coordinate-time interval dt.

Similarly, we can obtain the Coriolis acceleration by differentiating (4.114) with respect to s and taking $\dot{\varphi}$ to be instantaneously zero:

$$\begin{aligned} 2\dot{r}\dot{\varphi} + r\ddot{\varphi} + \frac{2w\dot{r}}{c} &= 0 \\ r\ddot{\varphi} + 2\frac{w\dot{r}}{c} &= 0 \end{aligned} \tag{4.122}$$

In terms of the proper-time interval $d\tau$ this may be written as

$$r\frac{d^2\varphi}{d\tau^2} + 2w\frac{dr}{d\tau} = 0 \tag{4.123}$$

which we also recognize as the classical Coriolis acceleration, again with the proper-time interval $d\tau$ replacing the coordinate-time interval dt.

From the above results it appears reasonable at this point to adopt tentatively the geodesic equations as the equations of motion in a four-

dimensional Riemann space—at least in the absence of forces. This has indeed proved to agree with classical theory for the example of the inertial system which we considered in the preceding paragraphs; furthermore, we obtained a reasonable generalization of the familiar Coriolis and centrifugal accelerations when we applied the geodesic equations to the example of the rotating coordinate system. In addition, the variational problem $\delta\int ds = 0$ is an invariant expression and provides an elegant and invariant characterization of the path of a particle. In the next section we shall investigate further the link between geometry and physics provided by the geodesic equations by an application to a more general metric space than the Lorentz form we considered in this section. This investigation will provide further justification for the use of the geodesic equations of motion.

4.3 Gravity as a Metric Phenomenon

In the last section we found that using the Lorentz metric and the geodesic equation as equations of motion, we obtained, in the examples considered, a reasonable description of the force-free motion of a particle in a four-dimensional Riemann space. In this section we shall attempt to show by an approximation procedure that the effect of a gravitational field of force can be described by again using the geodesic equation of motion and by allowing the metric tensor to differ "somewhat" from the Lorentz metric. The success of this procedure will serve as justification for linking the physical force of gravity with the non-Lorentzian nature of space and for using the geodesic equations as equations of motion for particles in a gravitational field.

Since the geometry of the real world is Euclidean so far as any ordinary physical measurement is concerned, the space-time metric must be very close indeed to Lorentzian. However, it is just the very minute departure from the Lorentzian metric which we desire to show is the agent of gravitational effects. Thus we shall consider a *time-independent* metric tensor of the form

$$g_{\mu\nu} = \eta_{\mu\nu} + \epsilon\gamma_{\mu\nu} \qquad \epsilon = \text{small constant} \tag{4.124}$$

where $\eta_{\mu\nu}$ is the Lorentz metric tensor and $\epsilon\gamma_{\mu\nu}$ represents a very small time-independent perturbation which is due to the presence of a gravitating body and goes to zero very far from the body. We term this $g_{\mu\nu}$ a nearly Lorentzian metric tensor.

To show that the $\epsilon\gamma_{\mu\nu}$ term is indeed the agent of gravitational forces, we shall apply the geodesic equations of motion to a Riemann space with the above metric. Furthermore, to make a close connection with classical theory, we shall suppose that the velocity of the particle along the geodesic is much less than c, or equivalently, that $\beta = v/c$ is very small; then in our approximate calculations we shall retain only first-order terms in ϵ and β, dropping terms of order ϵ^2, β^2, $\epsilon\beta$, and higher.

Using the coordinates of special relativity and the nearly Lorentzian metric tensor (4.124), we immediately obtain the line element

$$ds^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 + \epsilon\gamma_{\mu\nu}\, dx^\mu\, dx^\nu \tag{4.125}$$

Thus

$$\left(\frac{ds}{dt}\right)^2 = c^2 - v^2 + \epsilon\gamma_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt} = c^2\left(1 - \beta^2 + \epsilon\gamma_{\mu\nu}\frac{dx^\mu}{dx^0}\frac{dx^\nu}{dx^0}\right) \tag{4.126}$$

To first order in ϵ and β this is

$$\left(\frac{ds}{dt}\right)^2 \cong c^2(1 + \epsilon\gamma_{00}) \tag{4.127}$$

We next apply the same approximation to the differential equations of a geodesic:

$$\frac{d^2x^\alpha}{ds^2} + \left\{\begin{matrix}\alpha\\ \eta\ \tau\end{matrix}\right\}\frac{dx^\eta}{ds}\frac{dx^\tau}{ds} = 0 \tag{4.128}$$

Consider the second term on the left side of this equation. Since $\eta_{\mu\nu}$ is constant in space-time, it is evident from the form of the metric (4.125) that each Christoffel symbol contains a factor ϵ. Using the expression (4.127) for $(ds/dt)^2$, we can write

$$\frac{dx^\eta}{ds}\frac{dx^\tau}{ds} = \frac{dx^\eta}{dt}\frac{dx^\tau}{dt}\left(\frac{dt}{ds}\right)^2 = \frac{dx^\eta}{dt}\frac{dx^\tau}{dt}\frac{1}{c^2(1 + \epsilon\gamma_{00})} \tag{4.129}$$

If neither η nor τ is zero, this can be written in terms of the τ and η components of the velocity v as

$$\frac{dx^\eta}{dt}\frac{dx^\tau}{dt}\frac{1}{c^2(1 + \epsilon\gamma_{00})} = \frac{v^\eta}{c}\frac{v^\tau}{c}\frac{1}{(1 + \epsilon\gamma_{00})} \tag{4.130}$$

which is of order β^2. If only one of the indices η or τ is not zero, the expression is clearly of order β. Thus, unless $\eta = \tau = 0$, the product $\left\{\begin{matrix}\alpha\\ \eta\ \tau\end{matrix}\right\}\frac{dx^\eta}{ds}\frac{dx^\tau}{ds}$ is of order $\epsilon\beta$ or higher, and is to be neglected in our approx-

imation scheme. The geodesic equations to first order in ϵ and β are then

$$\frac{d^2x^\alpha}{ds^2} + \left\{\begin{matrix}\alpha\\ 0\ 0\end{matrix}\right\}\left(\frac{dx^0}{ds}\right)^2 = 0 \tag{4.131}$$

By virtue of (4.127) these may also be written within our degree of approximation as

$$\frac{d^2x^\alpha/dt^2}{c^2(1 + \epsilon\gamma_{00})} + \frac{\left\{\begin{matrix}\alpha\\ 0\ 0\end{matrix}\right\}}{1 + \epsilon\gamma_{00}} = 0 \tag{4.132}$$

since all terms neglected in the transition from (4.131) to (4.132) contain factors $\epsilon\beta$. Equivalently,

$$\frac{d^2x^\alpha}{dt^2} + \left\{\begin{matrix}\alpha\\ 0\ 0\end{matrix}\right\}c^2 = 0 \tag{4.133}$$

We do not need the differential equation for $x^0 = ct$, and indeed we shall have to test the comp tibility $\left\{\begin{matrix}0\\ 0\ 0\end{matrix}\right\} = 0$.

In order to simplify this approximate equation further, we shall calculate the Christoffel symbol $\left\{\begin{matrix}\alpha\\ 0\ 0\end{matrix}\right\}$ explicitly. By definition the corresponding Christoffel symbol of the first kind is

$$[00, \lambda] = \tfrac{1}{2}(g_{0\lambda|0} + g_{\lambda 0|0} - g_{00|\lambda}) \tag{4.134}$$

Since, by our assumption of a time-independent metric, $g_{\mu\nu}$ is independent of x^0 and $\eta_{\mu\nu}$ is a constant in all the variables x^μ, this becomes

$$[00, \lambda] = -\tfrac{1}{2}g_{00|\lambda} = -\tfrac{1}{2}\epsilon\gamma_{00|\lambda} \tag{4.135}$$

Raising the index λ to obtain the Christoffel symbol of the second kind and ignoring terms of order ϵ^2, we then obtain

$$\left\{\begin{matrix}\alpha\\ 0\ 0\end{matrix}\right\} = g^{\alpha\lambda}[00, \lambda] = -\tfrac{1}{2}g^{\alpha\lambda}\epsilon\gamma_{00|\lambda} = -\tfrac{1}{2}g^{(L)\alpha\lambda}\epsilon\gamma_{00|\lambda} \tag{4.136}$$

Separating time and space components, we obtain for $\alpha = 0$

$$\left\{\begin{matrix}0\\ 0\ 0\end{matrix}\right\} = -\tfrac{1}{2}\epsilon\gamma_{00|0} = 0 \tag{4.137}$$

because of our assumption of time independence. This is consistent with the definition $x^0 = ct$ and the validity of (4.133) for $\alpha = 0$; we recognize that our two physical assumptions, (1) time independence of the metric and (2) smallness of velocities in our coordinate system, are interdependent. This is the reason for the compatibility of (4.137) and (4.133). For $\alpha = i = 1, 2, 3$, we have

$$\left\{ {i \atop 0\ 0} \right\} = \tfrac{1}{2}\epsilon\gamma_{00|i} \tag{4.138}$$

Using this approximation for the Christoffel symbol, we can write the geodesic equations (4.133) as

$$\frac{d^2x^i}{dt^2} = -\frac{c^2}{2}\epsilon\gamma_{00|i} \tag{4.139}$$

In three-dimensional vector notation this is

$$\frac{d^2\mathbf{x}}{dt^2} = -\frac{c^2}{2}\epsilon\nabla\gamma_{00} \tag{4.140}$$

This is simply Newton's equation of motion in a classical gravitational field derived from a scalar potential if we identify the scalar potential as

$$\varphi = \frac{c^2}{2}\epsilon\gamma_{00} \tag{4.141}$$

Conversely, given the classical potential φ, the motion of a particle will be along a four-dimensional geodesic if the g_{00} term of the metric tensor has the form

$$g_{00} = 1 + \frac{2\varphi}{c^2} \tag{4.142}$$

The other components do not enter in our approximation, except through the assumption that they are time-independent and nearly Lorentzian.

Let us summarize the preceding results: If we ignore second-order terms in ϵ and β (the weak field and low velocity limit), then the geodesic equation (a purely geometric relation) is equivalent to Newton's equation (4.140) (a purely mechanistic relation), provided the g_{00} of the metric tensor satisfies the relation (4.142). This equivalence provides rather good justification for the use of a non-Lorentzian metric to describe a gravitational field and for the use of the geodesic equations of motion in

the resulting Riemann space. Furthermore, the approximate relation (4.142) for the g_{00} term of the metric tensor will itself be useful for several reasons: (1) It will allow us to relate certain constants which appear formally in the later developments of the relativity theory to familiar classical quantities, such as mass and the gravitational constant. (2) We may check, when we solve the gravitational field equations of Einstein in Chap. 8, that the g_{00} component is consistent with our approximate result. (3) We can predict the red shift of spectral lines in a gravitational field, using only the above results in the framework of special relativity. This will be the subject of the next section.

4.4 The Red Shift

Without going beyond the preliminary notions we have developed in this chapter, we can predict a rather interesting effect of the gravitational field: the slowing down of time in the field and the consequent red shift of spectral lines emitted by atoms located on massive bodies. The effect has been tested by experiment and been rather well verified; we thus have experimental justification for the basic theoretical concepts we have set forth in this chapter.

Consider, for example, a light wave emitted on the sun and received on the earth. Let the gravitational potential at the surface of the sun be φ_s. Using (4.94) and the approximate g_{00} given in Eq. (4.142), proper-time intervals are related to coordinate-time intervals by the equation

$$d\tau_s = \sqrt{g_{00}(x_s^\mu)}\,dt = \left(1 + \frac{2\varphi_s}{c^2}\right)^{1/2} dt \tag{4.143}$$

Similarly, on the earth, proper-time intervals are related to coordinate-time intervals by

$$d\tau_e = \sqrt{g_{00}(x_e^\mu)}\,dt = \left(1 + \frac{2\varphi_e}{c^2}\right)^{1/2} dt \tag{4.144}$$

where φ_e is the value of the gravitational potential on the earth. Suppose now n waves of frequency ν_0 are emitted in proper time $\Delta\tau_s$ from an atom on the sun. Then

$$n = \nu_0\,\Delta\tau_s \tag{4.145}$$

On the earth one certainly receives n waves, but the frequency and time duration of the wave train have changed. Using a frequency-duration relation for the earth analogous to (4.145) for the sun,

$$n = \nu_e \, \Delta\tau_e \tag{4.146}$$

we obtain, since n is a constant,

$$\nu_0 \, \Delta\tau_s = \nu_e \, \Delta\tau_e \tag{4.147}$$

Thus

$$\nu_e = \nu_0 \frac{\Delta\tau_s}{\Delta\tau_e} \tag{4.148}$$

From (4.143) the coordinate-time duration of the wave corresponding to $\Delta\tau_s$ is

$$\Delta t = \frac{\Delta\tau_s}{\sqrt{g_{00}(x_s^\mu)}} = \frac{\Delta\tau_s}{\sqrt{1 + 2\varphi_s/c^2}} \tag{4.149}$$

We suppose that the coordinate-time duration of the wave Δt is the *same* on the earth as on the sun. [See Sec. 4.2 on proper-time and coordinate-time, especially (4.95).] Equation (4.144) then gives

$$\Delta t = \frac{\Delta\tau_e}{\sqrt{g_{00}(x_e^\mu)}} = \frac{\Delta\tau_e}{\sqrt{1 + 2\varphi_e/c^2}} \tag{4.150}$$

By virtue of this and (4.149), we then have

$$\frac{\Delta\tau_s}{\Delta\tau_e} = \frac{\sqrt{g_{00}(x_s^\mu)}}{\sqrt{g_{00}(x_e^\mu)}} = \left(\frac{1 + 2\varphi_s/c^2}{1 + 2\varphi_e/c^2}\right)^{1/2} \tag{4.151}$$

Substitution of this into (4.148) gives

$$\nu_e = \nu_0 \frac{\sqrt{g_{00}(x_s^\mu)}}{\sqrt{g_{00}(x_e^\mu)}} = \nu_0 \left(\frac{1 + 2\varphi_s/c^2}{1 + 2\varphi_e/c^2}\right)^{1/2} \tag{4.152}$$

Expanding to first order in the small quantities φ_s/c^2 and φ_e/c^2, we obtain

$$\frac{\nu_e - \nu_0}{\nu_0} = \frac{\varphi_s - \varphi_e}{c^2} \tag{4.153}$$

or in briefer notation,

$$\frac{\Delta\nu}{\nu_0} = \frac{\Delta\varphi}{c^2} \tag{4.154}$$

Since the sun is at a large negative potential relative to the earth, we see that $\Delta\varphi$ is negative. Thus the frequency of light *decreases* as it leaves the sun, and when it is received on earth, we see a shift toward the red end of the spectrum. It is as if the atoms of the sun were vibrating in slow motion when we view them from the earth. Of course, there is nothing special about using the earth and sun as the two points considered, and we can just as well use two points at different heights on the earth if our measurement is precise enough to detect the correspondingly small shift.

Alternative derivations of the red shift formula. Using the equivalence principle in a very direct way, Einstein made the first derivation of the red shift by considering a system rotating at constant angular velocity w such as we used in Sec. 4.2. The linear velocity of a fixed point at radius r is then $v = wr$, and the centrifugal potential is

$$\varphi_{\text{cent}} = -\tfrac{1}{2}w^2r^2 \tag{4.155}$$

As obtained in Sec. 4.2, the relation between time intervals on the rotating system and external inertial intervals is, in agreement with special relativity,

$$d\tau = \left(1 - \frac{v^2}{c^2}\right)^{1/2} dt = \left(1 - \frac{w^2r^2}{c^2}\right)^{1/2} dt = \left(1 + \frac{2\varphi_{\text{cent}}}{c^2}\right)^{1/2} dt \tag{4.156}$$

But by the principle of equivalence, the centrifugal potential should be experimentally indistinguishable in the small from a gravitational potential, and the above relation should hold equally well for a gravitational potential φ. Proceeding as before, we then obtain the red shift formula from the above, since (4.156) and (4.143) are formally identical.

For the case of the terrestrial red shift, where the points under consideration are close together, another derivation is possible. Let us replace the gravitational field by an acceleration of the coordinate system. That is, instead of placing our apparatus in the earth's field with potential gradient $g = 980$ cm/sec^2, we put it in an elevator or rocket accelerating at g in free space, which (according to the equivalence principle) will give identical physical results (Fig. 4.1). The time it takes a light wave to travel between sender and receiver is roughly d/c. But in this time the receiver has increased its velocity over that of the sender when the wave was emitted by an amount $(d/c)g$. Thus there is a consequent Doppler shift corresponding to $\Delta v/c = (d/c^2)g$. The frequency shift $\Delta\nu/\nu = -\Delta v/c$ is then

$$\frac{\Delta\nu}{\nu} = -\frac{gd}{c^2} \tag{4.157}$$

If we note that $-gd$ is the change in the gravitational potential $\Delta\varphi$ over a distance d, we can write as before $\Delta\nu/\nu = \Delta\varphi/c^2$.

It is important to note that the two alternative derivations of the red shift which have been presented here rely on the validity of the *principle of equivalence*, but not on the field equations of the general theory of relativity; the derivation of the red shift does not require the use of Einstein's equations, which will be introduced in the next chapter.

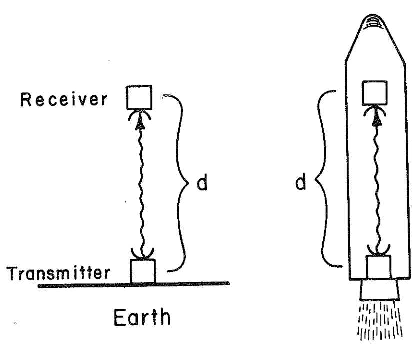
Yet a third derivation of the red shift is possible which does not rely explicitly on the principle of equivalence but only on the mass-energy relation $E = mc^2$ and the interpretation of light as energetic quanta with an effective mass m. According to the mass-energy relation, a light quantum of energy $E = h\nu$ will have effective mass

(4.158) $$m = \frac{E}{c^2} = \frac{h\nu}{c^2} \qquad h = \text{Planck's constant}$$

Accordingly, the sum of potential and kinetic energy of the quantum at a point where the gravitational potential is φ will be $h\nu + m\varphi$. Thus, when the quantum moves through a potential difference $\Delta\varphi$, say, between the sun (s) and the earth (e), we have the energy-balance equation

(4.159) $$h\nu_e + m\varphi_e = h\nu_0 + m\varphi_s$$

Fig. 4.1



Using the effective mass relation (4.158), we have

(4.160) $$h\nu_e - h\nu_0 = m(\varphi_s - \varphi_e) = \frac{h\nu_0}{c^2}(\varphi_s - \varphi_e)$$

which gives, upon canceling the h, the previous result (4.154):

(4.161) $$\frac{\Delta\nu}{\nu_0} = \frac{\Delta\varphi}{c^2}$$

Experimental tests of the red shift. Historically, the initial interest in the red shift equation (4.154) centered about its experimental verification through spectral measurements of the sun and other stars; this would be a rather direct verification of the principle of equivalence. Such astronomical measurements are, however, difficult to make, and their interpretation is apt to be ambiguous, because of the presence of nongravitational effects on a stellar surface such as Doppler shifts in the high-temperature gas, intense electromagnetic radiation fields, possible high electric fields due to gas ionization, vertical currents in the stellar gas, high gas pressure, etc. As a result it has been difficult to isolate accurately that part of the observed spectral shift which is attributable to the gravitational red shift. We shall mention only one recent result. Brault and Dicke (Brault, 1962; Dicke, 1964) used direct electronic techniques without appealing to the usual photographic-plate techniques of astronomers. They measured the displacement of the center of the broad D_1 line of sodium as a function of the radial distance across the solar disk. They found to within 5 per cent accuracy that the shift was constant (after corrections for the small line asymmetry) and agreed with the theoretical value.

Since it has been impossible to place complete confidence in astronomical measurements, there has naturally been considerable interest in the possibility of a terrestrial test of the red shift. This is a very difficult task, for the expected shift over a vertical distance of, say, 100 ft, is only of the order of 10^{-15}. Fortunately, the discovery of the Mössbauer effect in 1958 (Mössbauer, 1958a and b, 1959) gave a method of producing and detecting gamma rays which are monochromatic to 1 part in 10^{12} and made a terrestrial test feasible.

In general, it is difficult to measure very small shifts in the gamma-ray spectral lines of nuclei; this is because a nucleus in a crystal is usually able to recoil when emitting a gamma ray, and one therefore observes a Doppler effect in the observed spectra. Mössbauer found that in some crystals such as Fe^{57} the whole crystal instead of one nucleus can take up the recoil momentum. This makes the emission effectively recoilless, and one can observe very narrow spectral lines. Indeed, in Fe^{57}, there is a line near 14.4 kev with a fractional width of 10^{-12}. If a radioactive emitting sample of Fe^{57} is placed near a thin nonradioactive absorbing sheet of the same material, one then finds that a large fraction of the gamma radiation falling on the sheet is *resonantly absorbed*. If, however, either the absorber or emitter is moved at a velocity of only a few millimeters per second, the consequent Doppler shift becomes as large as the line width and the resonant absorption falls off very rapidly. This gives one a very accurate tool for the detection of quite small frequency shifts.

In the red shift experiment (Pound and Rebka, 1960; and Pound and Snider, 1964) emitter and absorber were placed at opposite ends of a vertical 72-ft tower. Gamma rays emitted at the bottom then suffered a gravitational red shift in traveling to the absorber at the top and, as a result, were less favorably absorbed. By moving the emitter upward at a small velocity, a compensating Doppler shift was produced which restored resonant absorption. A measurement of the emitter velocity then allowed a calculation of the ratio $\Delta\nu/\nu$. The experimental result obtained is 0.997 ± 0.008 times the predicted shift of 4.92×10^{-15}. This result represents verification of the correctness of the red shift equation (4.154) to better than 1 per cent.

This ends for the present the consideration of how the gravitational field, in the form of a non-Lorentzian metric, influences the matter (or light) in its vicinity. In the next chapter we shall consider the converse problem, how the matter influences the metric structure of space, and begin the discussion of the gravitational field equations in free space.

Exercises

4.1 Show that in Lorentz space

$$F^{\sigma\tau}F_{\sigma\tau} = 2(\mathbf{H}^2 - \mathbf{E}^2) \qquad {}^*F^{\alpha\beta}F_{\alpha\beta} = 4\,\mathbf{E}\cdot\mathbf{H}$$

Are there any more bilinear invariants that can be constructed from the components of the electromagnetic field? What are the values of the above invariants for a plane wave?

4.2 Define the following complex tensor from the Maxwell tensor and its dual tensor

$$\omega_{\mu\nu} = F_{\mu\nu} + i({}^*F_{\mu\nu})$$

Express Maxwell's equations in terms of $\omega_{\mu\nu}$.

4.3 The tensor introduced above has the remarkable property that the bilinear invariants of electromagnetic theory discussed in Exercise 4.1 are simply accommodated in a single complex invariant. To see this show that

$$\omega^{\alpha\beta}\omega_{\alpha\beta} = 2(F_{\sigma\tau}F^{\sigma\tau}) + i({}^*F^{\alpha\beta}F_{\alpha\beta})$$

$$\omega^{\alpha\beta}\bar{\omega}_{\alpha\beta} = 0$$

where $\bar{\omega}_{\alpha\beta}$ is the complex conjugate of $\omega_{\alpha\beta}$. What is $\bar{\omega}^{\alpha\beta}\bar{\omega}_{\alpha\beta}$?

4.4 Consider Maxwell's equations in special relativity. Show that by a gauge transformation one may assure that

$$A^{\mu}{}_{|\mu} = 0$$

This is termed the Lorentz gauge condition. Using such a gauge, show that A^μ satisfies the equation

$$\Box^2 A^\mu \equiv A^\mu{}_{|\alpha|\beta}g^{\alpha\beta} = s^\mu$$

4.5 In free space show that $F^{\mu\nu}$ satisfies the wave equation whereas A^μ satisfies it only if the Lorentz condition is imposed. (Work in Lorentz space: is the exercise true in an arbitrary Riemann space?)

4.6 The coordinate transformation considered in special relativity is linear. What is it explicitly for motion in the x direction? What are the transformation laws for contravariant and covariant special relativistic four-vectors? What are they for second-rank tensors? Obtain the transformation laws for the electromagnetic fields under the Lorentz transformation. State them in the form

$$E'_{\parallel} = F(E,B) \qquad H'_{\parallel} = K(E,B)$$

$$E'_{\perp} = G(E,B) \qquad H'_{\perp} = J(E,B)$$

where $\parallel$ means parallel to the direction of relative motion and $\perp$ means perpendicular to the direction of relative motion.

4.7 In a weak gravitational field represented by a first-order metric with $g_{00} = 1 - 2\varphi/c^2$ show that the Lorentz time-dilation factor should be replaced by $(1 - v^2/c^2 - 2\varphi/c^2)^{-1/2}$.

Problems

A useful reference for these problems is Vishveshwara, 1968.

4.1 Let k^α represent a field of null geodesics; i.e., the lines with tangent vectors $k^\alpha = dx^\alpha/dq$ are geodesics, and $k^\mu k_\mu = 0$. Show that this implies that $k^\alpha{}_{\|\beta}k^\beta = 0$ if q is one of the distinguished parameters discussed in Sec. 2.3. Show moreover that by a suitable normalization k^α can be interpreted as the wave vector of a photon. See Sec. 6.5.

4.2 A stationary metric has, by definition, a timelike Killing vector ξ^μ, which can be normalized to yield a unit vector $u^\alpha = \xi^\alpha/(\xi^\mu\xi_\mu)^{1/2}$. Show that u^α may be interpreted as the four-velocity of an observer at rest. Show that such an observer will observe the photons discussed in Prob. 4.1 to have frequencies $u^\alpha k_\alpha$.

4.3 Along a null geodesic line show that $(d/dq)\,(k_\alpha\xi^\alpha) = 0$ and that the red shift of a photon emitted at s and observed at o by stationary observers may be written

$$\frac{\nu_o}{\nu_s} = \frac{(k_\alpha u^\alpha)_o}{(k_\alpha u^\alpha)_s} = \frac{(\xi^\mu\xi_\mu)_s{}^{1/2}}{(\xi^\beta\xi_\beta)_o{}^{1/2}}$$

This covariant statement is equivalent to (4.152).

Bibliography

Textbooks on special relativity:

Aharoni, J. (1959): "The Special Theory of Relativity," Oxford.
Bergmann, P. G. (1942): "An Introduction to the Theory of Relativity," New York.
Bergmann, P. G. (1962): The Special Theory of Relativity, in Encyclopedia of Physics, vol. 4, Berlin-Göttingen-Heidelberg, pp. 108–202.
Laue, M. von (1955): "Die Relativitätstheorie," vol. 1, "Spezielle Relativitätstheorie," Brunswick.
Papapetrou, A. (1955): "Spezielle Relativitätstheorie," Berlin.
Sommerfeld, A. (1952): "Lectures on Theoretical Physics," vol. 3, "Electrodynamics," New York.
Synge, J. L. (1956): "Relativity: The Special Theory," Amsterdam.

On characteristics:

Courant, R., and D. Hilbert (1962): "Methods of Mathematical Physics," vol. 2, "Partial Differential Equations," New York.
Fock, V. (1959): "The Theory of Space, Time and Gravitation," New York.

Experimental tests of red shift:

Adam, M. G. (1955): Interferometric Measurement of Wave Lengths, II, *Monthly Notices Roy. Astron. Soc.*, **115**:405–421.
Blamont, J. E., and F. Roddier (1961): Precise Observation of the Profile of the Strontium Resonance Line: Evidence for the Gravitational Red Shift on the Sun, *Phys. Rev. Letters*, **7**:437–439.
Brault, J. (1962): Thesis, Princeton University, Princeton, N.J.
Cranshaw, T. E., J. P. Schiffer, and A. B. Whitehead (1960): Measurement of the Gravitational Red Shift Using the Mössbauer Effect in Fe^{57}, *Phys. Rev. Letters*, **4**:163–164.
Dicke, R. H. (1964): Experimental Relativity, in B. Dewitt (ed.), "Relativity, Groups and Topology," New York.
Finlay-Freundlich, E. (1954): Red Shifts in the Spectra of Celestial Bodies, *Philos. Mag.*, **45**:303–319.
Frauenfelder, H. (1962): "The Mössbauer Effect," New York.
Hay, H. J., J. P. Schiffer, T. E. Cranshaw, and P. A. Egelstaff (1960): Measurement of the Red Shift in an Accelerated System Using the Mössbauer Effect in Fe^{57}, *Phys. Rev. Letters*, **4**:165–166.
Mössbauer, R. L. (1958a): Kernresonanzfluoreszenz von Gammastrahlung in Ir^{191}, *Z. Physik*, **151**:124–143.
Mössbauer, R. L. (1958b): Kernresonanzabsorption von Gammastrahlung in Ir^{191}, *Naturwissenschaften*, **45**:538–539.
Mössbauer, R. L. (1959): Kernresonanzabsorption von γ-Strahlung in Ir^{191}, *Z. Naturforsch.*, **14a**:211–216.
Pound, R. V., and G. A. Rebka (1959): Gravitational Red Shift in Nuclear Resonance, *Phys. Rev. Letters*, **3**:439–441.
Pound, R. V., and G. A. Rebka (1960): Apparent Weight of Photons, *Phys. Rev. Letters*, **4**:337–341.
Pound, R. V., and J. L. Snider (1964): Effect of Gravity on Nuclear Resonance, *Phys. Rev. Letters*, **13**:539–540.
Schröter, E. H. (1956): Deutung der Rotverschiebung . . . der Sonnenatmosphäre, *Z. Astrophys.*, **41**:141–181.
Vishveshwara, C. V. (1968): Generalization of the "Schwarzschild Surface" to Arbitrary Static and Stationary Metrics, *J. Math. Phys.*, **9**:1319.

CHAPTER

5

The Gravitational Field Equations in Free Space

From our investigations in the preceding chapter we can state the following two conclusions: (1) In the absence of forces, mechanical phenomena can be suitably described in a differential geometric framework by using a Lorentz metric and a geodesic equation of motion. (2) The effect of gravitational forces can be included in the differential geometric framework by using a non-Lorentzian metric and retaining the geodesic equation of motion; the metric in this case differs only slightly from the Lorentz metric in regions where the gravitational potential is nonzero, and is Lorentzian where the gravitational potential is zero. The program of this chapter is to obtain field equations for the metric tensor in matter-free regions of space; that is, we wish to answer the question of how matter affects the metric structure of the free space in its vicinity.

5.1 Criteria for the Field Equations

Our first criterion, as could be expected from our entire approach, is that the field equations (and the whole theory!) be phrased in covariant tensor form. This is Einstein's well-known *principle of covariance*, and is seen to be desirable for the following reasons:

1. The principle of equivalence implies that accelerated systems must be considered to be quite as respectable as inertial systems, so we demand that physical laws do not distinguish between the two. This will clearly be so if the laws are tensor laws, for then the system of coordinates does not enter the equations at all.

2. From Chap. 4 we know that, in a four-dimensional tensor formulation, both fictitious forces and gravitational forces appear as Christoffel symbols in the geodesic equations of motion. The fact that they appear mathematically alike is a very desirable characteristic of the tensor approach, for according to the equivalence principle, gravitational and fictitious (inertial) forces are indistinguishable "in the small."

3. It was pointed out first by Kretschmann (Kretschmann, 1917) that on purely mathematical grounds any (tentative) physical law expressed in a particular coordinate system can be brought into a covariant tensor form. However, in physics, one wants to keep the number of quantities (observables) entering the equations to a minimum; this will single out certain laws for which no new formal tensor quantities will have to be introduced in the process of making the formulas covariant. The principle of covariance thereby provides a purely formal but successful guide to choosing physical laws on the basis of elegance and simplicity. Applying it to the problem of gravity gives the first criterion: *(a) The gravitational equations should be phrased in covariant tensor form.*

The classical theory of gravitation in free space can be basically stated in two equations, namely, Newton's second law in a gravitational field with a potential φ,

$$\frac{d^2x^i}{dt^2} = -\frac{\partial\varphi}{\partial x^i} \tag{5.1}$$

and Laplace's equation for the potential:

$$\sum_{i=1}^{3} \frac{\partial^2\varphi}{\partial x^{i2}} = 0 \tag{5.2}$$

(The Latin indices run from 1 to 3 by our standing convention.) Equation (5.1) tells how matter moves in a given gravitational field. Equation (5.2) describes how matter (or rather its absence) determines the gravitational field; indeed, (5.2) is a special case of the Poisson-Laplace equation $\nabla^2\varphi = -4\pi\rho$. We saw in Sec. 4.3 that the geodesic equation of motion which we desire to use in the relativity theory reduces to the classical law (5.1) *if* we identify $g_{00} \cong 1 + 2\varphi/c^2$. This provides us with a second clue to the nature of the relativity field equations; indeed, since Laplace's equation involves the second derivatives of φ, we might expect that the general relativistic field equations will involve the second derivatives of g_{00}. Because of the tensor form of the equations, this implies that the second derivatives of all the components of $g_{\mu\nu}$ should appear. Our inclination to look for second-order equations is strengthened also by the fact that most of the differential field equations of classical physics are of second order. Thus our second criterion is: *(b) The field equations should be of second order in the components of the metric tensor.*

We found in Chap. 4 that, in the absence of forces, we can describe mechanics using a Lorentz metric, so we adopt as a third criterion: *(c) For the case where all space is empty of matter* (and there is consequently no gravitational field), *the field equation must admit the Lorentz metric as a particular solution.* This last criterion will be somewhat modified in Chap. 10, when we consider large-scale cosmological problems.

Our fourth criterion is intended to guarantee a unique solution of the field equations. If we have a differential equation of the form

$$F(y^{(n)}, \ldots, y, x) = 0$$

which can be solved for $y^{(n)}$, it is clear that the differential equation can have a unique solution only if $y^{(n)}$ is uniquely determined by $y^{(n-1)}$, $y^{(n-2)}$, etc. This will surely be the case if $y^{(n)}$ enters linearly; indeed, if the function F is algebraic, linearity in $y^{(n)}$ is a necessary and sufficient condition for uniqueness of a solution. Thus, for the field equations, we ask that the second derivatives of $g_{\mu\nu}$ enter linearly; that is, the equations should be quasi-linear. The final criterion can then be stated: *(d) The field equations should be quasi-linear.*

Other sets of guiding criteria for developing the field equations are of course possible, but the set of four criteria we have stated above are particularly simple, and will prove to be an adequate guide in the development of the following sections.

5.2 The Riemann Curvature Tensor

The first guide we shall invoke is criterion (*c*). Since the Lorentz metric must be one very important solution of the gravitational field equations, we shall try to find a necessary condition, expressed in the form of a differential field equation, for a space to have a Lorentz metric.

Let us note first an important property of the Lorentz metric: in the coordinates of special relativity its components are *constant over all space.* This property will be essential in providing our criteria for a space to have a Lorentz metric. If a space has a Lorentz metric, then by assumption there exists a system where the metric is constant and has the Lorentz form. That is, there exists a geodesic system *in the large.* All the Christoffel symbols are *everywhere* zero in that system; consequently the ordinary derivatives and the covariant derivatives of any vector ξ^α are equal:

$$\xi^\alpha{}_{\|\beta} = \xi^\alpha{}_{|\beta} \tag{5.3}$$

Indeed, it is evident that covariant and ordinary derivatives of *all* orders are equal if and only if the Christoffel symbols are *everywhere* zero, so we also have

$$\xi^\alpha{}_{\|\beta\|\gamma} = \xi^\alpha{}_{|\beta|\gamma} \tag{5.4}$$

and similarly for derivatives of all orders. Since, for any vector field, the order of ordinary differentiation is irrelevant, it follows that the order of covariant differentiation is also irrelevant. That is,

$$\xi^\alpha{}_{\|\beta\|\gamma} - \xi^\alpha{}_{\|\gamma\|\beta} = 0 \tag{5.5}$$

But this is a tensor equation; since it is true in one system, it must be true in *all systems*, not just the geodesic system. Thus we have the rule that *when* a space admits a Lorentz metric, (5.5) holds. The differential equation (5.5) is consequently a necessary condition for a space to admit a Lorentz metric. A little manipulation will put this equation in more useful form. For convenience, let us define a tensor $t^\alpha{}_\beta$:

$$t^\alpha{}_\beta = \xi^\alpha{}_{\|\beta} = \xi^\alpha{}_{|\beta} + \begin{Bmatrix} \alpha \\ \beta\ \eta \end{Bmatrix} \xi^\eta \tag{5.6}$$

Then

$$t^\alpha{}_{\beta\|\gamma} = \xi^\alpha{}_{\|\beta\|\gamma} = t^\alpha{}_{\beta|\gamma} + \begin{Bmatrix} \alpha \\ \tau\ \gamma \end{Bmatrix} t^\tau{}_\beta - \begin{Bmatrix} \lambda \\ \beta\ \gamma \end{Bmatrix} t^\alpha{}_\lambda \tag{5.6'}$$

Inserting (5.6) into (5.6′), we obtain

$$\xi^\alpha{}_{\|\beta\|\gamma} = \xi^\alpha{}_{|\beta|\gamma} + \begin{Bmatrix} \alpha \\ \beta\ \eta \end{Bmatrix}_{|\gamma} \xi^\eta + \begin{Bmatrix} \alpha \\ \beta\ \eta \end{Bmatrix} \xi^\eta{}_{|\gamma} + \begin{Bmatrix} \alpha \\ \tau\ \gamma \end{Bmatrix} \xi^\tau{}_{|\beta} + \begin{Bmatrix} \alpha \\ \tau\ \gamma \end{Bmatrix} \begin{Bmatrix} \tau \\ \beta\ \eta \end{Bmatrix} \xi^\eta - \begin{Bmatrix} \lambda \\ \beta\ \gamma \end{Bmatrix} t^\alpha{}_\lambda \tag{5.7}$$

Interchanging β and γ, we obtain a similar expression.

$$\xi^\alpha{}_{\|\gamma\|\beta} = \xi^\alpha{}_{|\gamma|\beta} + \begin{Bmatrix} \alpha \\ \gamma\ \eta \end{Bmatrix}_{|\beta} \xi^\eta + \begin{Bmatrix} \alpha \\ \gamma\ \eta \end{Bmatrix} \xi^\eta{}_{|\beta} + \begin{Bmatrix} \alpha \\ \eta\ \beta \end{Bmatrix} \xi^\eta{}_{|\gamma} + \begin{Bmatrix} \alpha \\ \tau\ \beta \end{Bmatrix} \begin{Bmatrix} \tau \\ \gamma\ \eta \end{Bmatrix} \xi^\eta - \begin{Bmatrix} \lambda \\ \gamma\ \beta \end{Bmatrix} t^\alpha{}_\lambda \tag{5.8}$$

The difference of the two expressions then gives the left side of (5.5) in the form

$$\xi^\alpha{}_{\|\beta\|\gamma} - \xi^\alpha{}_{\|\gamma\|\beta} = \begin{Bmatrix} \alpha \\ \beta\ \eta \end{Bmatrix}_{|\gamma} \xi^\eta - \begin{Bmatrix} \alpha \\ \gamma\ \eta \end{Bmatrix}_{|\beta} \xi^\eta + \begin{Bmatrix} \alpha \\ \tau\ \gamma \end{Bmatrix} \begin{Bmatrix} \tau \\ \beta\ \eta \end{Bmatrix} \xi^\eta \tag{5.9}$$

$$- \begin{Bmatrix} \alpha \\ \tau\ \beta \end{Bmatrix} \begin{Bmatrix} \tau \\ \gamma\ \eta \end{Bmatrix} \xi^\eta$$

$$= \left[\begin{Bmatrix} \alpha \\ \beta\ \eta \end{Bmatrix}_{|\gamma} - \begin{Bmatrix} \alpha \\ \gamma\ \eta \end{Bmatrix}_{|\beta} + \begin{Bmatrix} \alpha \\ \tau\ \gamma \end{Bmatrix} \begin{Bmatrix} \tau \\ \beta\ \eta \end{Bmatrix} - \begin{Bmatrix} \alpha \\ \tau\ \beta \end{Bmatrix} \begin{Bmatrix} \tau \\ \gamma\ \eta \end{Bmatrix} \right] \xi^\eta$$

The object in braces must be a tensor by the quotient theorem. It is known as the *Riemann curvature tensor*, and as we shall see, it plays a central role in the geometric structure of a Riemann space. We denote it by $R^\alpha{}_{\eta\beta\gamma}$:

$$R^\alpha{}_{\eta\beta\gamma} = \begin{Bmatrix} \alpha \\ \beta\ \eta \end{Bmatrix}_{|\gamma} - \begin{Bmatrix} \alpha \\ \eta\ \gamma \end{Bmatrix}_{|\beta} + \begin{Bmatrix} \alpha \\ \tau\ \gamma \end{Bmatrix} \begin{Bmatrix} \tau \\ \beta\ \eta \end{Bmatrix} - \begin{Bmatrix} \alpha \\ \tau\ \beta \end{Bmatrix} \begin{Bmatrix} \tau \\ \gamma\ \eta \end{Bmatrix} \tag{5.10}$$

Note that, although we have introduced $R^\alpha{}_{\eta\beta\gamma}$ for a Riemann space, it is evident that our derivation holds also in a general affine space, since it involves only the coefficients of connection, and not the metric tensor itself. In the more general case, one need only replace $-\begin{Bmatrix} \alpha \\ \beta\ \eta \end{Bmatrix}$ by $\Gamma^\alpha{}_{\beta\eta}$. However, as soon as we lower an index to form the tensor $R_{\alpha\eta\beta\gamma}$, we commit ourselves to a metric space.

In terms of the Riemann tensor, Eq. (5.9) can be written

$$\xi^\alpha{}_{\|\beta\|\gamma} - \xi^\alpha{}_{\|\gamma\|\beta} = R^\alpha{}_{\eta\beta\gamma}\xi^\eta \tag{5.11}$$

Thus the necessary condition (5.5) that a Riemann space admit a Lorentz metric can be written

$$R^\alpha{}_{\eta\beta\gamma}\xi^\eta = \xi^\alpha{}_{\|\beta\|\gamma} - \xi^\alpha{}_{\|\gamma\|\beta} = 0 \tag{5.12}$$

Since ξ^η is an arbitrary vector, this implies the following simple property, which is a necessary condition for a space to have a Lorentz metric:

$$R^\alpha{}_{\eta\beta\gamma} = 0 \tag{5.13}$$

For simplicity we term a space *flat* if its *Riemann curvature tensor vanishes*. Equation (5.13) then states that *a space with a Lorentz metric is flat*. The converse statement, that a (physically acceptable) space has a Lorentz metric if it is flat, will be proved in Sec. 5.6, after we have further investigated the properties of the Riemann tensor.

Equation (5.13) now represents *a field equation for flat and gravity-free space* since the Lorentz metric is certainly a solution. Furthermore, since it involves the first derivatives of the Christoffel symbols linearly, it is second-order and quasi-linear in the metric tensor. It thus satisfies the field-equation criteria (*b*) and (*d*). This special field equation for *gravity-free* space will be a very helpful guide in obtaining the general field equations of gravity in a *nonflat* space.

In the process of achieving the primary aim of this section, the necessary condition for a space to have a Lorentz metric, we have also obtained Eq. (5.11), which is interesting and important in its own right. It is also possible to obtain a similar result for higher-rank tensors. Indeed, for a tensor of the form $\xi^\alpha\eta^\delta$, we have, by the product rule for covariant derivatives,

$$(\xi^\alpha\eta^\delta)_{\|\beta} = \xi^\alpha{}_{\|\beta}\eta^\delta + \xi^\alpha\eta^\delta{}_{\|\beta} \tag{5.14}$$

and

$$(\xi^\alpha\eta^\delta)_{\|\beta\|\gamma} = \xi^\alpha{}_{\|\beta\|\gamma}\eta^\delta + \xi^\alpha\eta^\delta{}_{\|\beta\|\gamma} + \xi^\alpha{}_{\|\beta}\eta^\delta{}_{\|\gamma} + \xi^\alpha{}_{\|\gamma}\eta^\delta{}_{\|\beta} \tag{5.15}$$

Interchanging the indices β and γ gives

$$(\xi^\alpha\eta^\delta)_{\|\gamma\|\beta} = \xi^\alpha{}_{\|\gamma\|\beta}\eta^\delta + \xi^\alpha\eta^\delta{}_{\|\gamma\|\beta} + \xi^\alpha{}_{\|\gamma}\eta^\delta{}_{\|\beta} + \xi^\alpha{}_{\|\beta}\eta^\delta{}_{\|\gamma} \tag{5.16}$$

The difference of the two expressions is then

$$(\xi^\alpha\eta^\delta)_{\|\beta\|\gamma} - (\xi^\alpha\eta^\delta)_{\|\gamma\|\beta} = (\xi^\alpha{}_{\|\beta\|\gamma} - \xi^\alpha{}_{\|\gamma\|\beta})\eta^\delta + (\eta^\delta{}_{\|\beta\|\gamma} - \eta^\delta{}_{\|\gamma\|\beta})\xi^\alpha \tag{5.17}$$

Inserting (5.11) into (5.17), we have

$$(\xi^\alpha\eta^\delta)_{\|\beta\|\gamma} - (\xi^\alpha\eta^\delta)_{\|\gamma\|\beta} = R^\alpha{}_{\tau\beta\gamma}\xi^\tau\eta^\delta + R^\delta{}_{\tau\beta\gamma}\xi^\alpha\eta^\tau \tag{5.18}$$

which is the result analogous to (5.11). However, we know from Chap. 2 that *any* second-rank tensor can be written as a linear combination of such products. Thus, for any second-rank tensor $T^{\alpha\delta}$, we must have

$$T^{\alpha\delta}{}_{\|\beta\|\gamma} - T^{\alpha\delta}{}_{\|\gamma\|\beta} = R^\alpha{}_{\tau\beta\gamma}T^{\tau\delta} + R^\delta{}_{\tau\beta\gamma}T^{\alpha\tau} \tag{5.19}$$

The generalization of this process to yet higher rank tensors is evident.

Before ending this section we also wish to note the immediate consequence of (5.11):

$$\xi_{\alpha\|\beta\|\gamma} - \xi_{\alpha\|\gamma\|\beta} = R_{\alpha\rho\beta\gamma}\xi^\rho \tag{5.11$'$}$$

which is useful in covariant differentiation of covariant vectors. This formula makes sense only in a metric space, as we remarked after Eq. (5.10). On making the same calculations for the interchange of derivatives of covariant vectors, as we did for contravariant vectors, we would have obtained

$$\xi_{\alpha\|\beta\|\gamma} - \xi_{\alpha\|\gamma\|\beta} = -R^\eta{}_{\alpha\beta\gamma}\xi_\eta \tag{5.11$''$}$$

a formula valid in the general affine case.

5.3 Symmetry Properties of the Riemann Tensor

Unfortunately, the Riemann tensor is a rather cumbersome tensor, with 4^4, or 256, components. However, the number of *independent* components is much less than this because of various symmetry relations. By inspection of (5.10), we see that it is antisymmetric in the third and fourth indices, β and γ. Thus the $\beta\gamma$ subblock has only 6, instead of 16, *independent* components. Combined with the 16 components of the $\alpha\eta$ subblock, this reduces the number of independent components to at most 96.

Consider now an arbitrary vector field ξ and the square of its length

$$\varphi = g_{\tau\lambda}\xi^\tau\xi^\lambda \tag{5.20}$$

Since φ is a scalar, its ordinary and covariant derivatives are the same, so we have

$$\varphi_{\|\beta} = \varphi_{|\beta} \tag{5.21}$$

Antisymmetrizing $\varphi_{\|\lambda\|\tau}$, we have, by the results of Chap. 3,

$$\{\varphi_{|\beta|\gamma}\} = \varphi_{\|\beta\|\gamma} - \varphi_{\|\gamma\|\beta} = \varphi_{|\beta|\gamma} - \varphi_{|\gamma|\beta} = 0 \tag{5.22}$$

Now since

$$\varphi_{\|\beta} = g_{\tau\lambda}\xi^\tau{}_{\|\beta}\xi^\lambda + g_{\tau\lambda}\xi^\tau\xi^\lambda{}_{\|\beta} = 2g_{\tau\lambda}\xi^\tau\xi^\lambda{}_{\|\beta} \tag{5.23}$$

and

$$\varphi_{\|\beta\|\gamma} = 2g_{\tau\lambda}\xi^\tau{}_{\|\gamma}\xi^\lambda{}_{\|\beta} + 2g_{\tau\lambda}\xi^\tau\xi^\lambda{}_{\|\beta\|\gamma} \tag{5.24}$$

we have also

$$\varphi_{\|\beta\|\gamma} - \varphi_{\|\gamma\|\beta} = 2g_{\tau\lambda}\xi^{\tau}(\xi^{\lambda}{}_{\|\beta\|\gamma} - \xi^{\lambda}{}_{\|\gamma\|\beta}) = 0 \tag{5.25}$$

Using (5.11), this can be stated in terms of the Riemann tensor as

$$\varphi_{\|\beta\|\gamma} - \varphi_{\|\gamma\|\beta} = 2\xi_{\alpha}R^{\alpha}{}_{\eta\beta\gamma}\xi^{\eta} = 0 \tag{5.26}$$

Thus, by (5.22) and (5.26), we can assert, for an arbitrary vector field $\boldsymbol{\xi}$, that

$$R_{\alpha\eta\beta\gamma}\xi^{\eta}\xi^{\alpha} = 0 \tag{5.27}$$

Let us now choose at a given point of our Riemann space the arbitrary vector $\boldsymbol{\xi}$ to be a unit vector with the α_0 component equal to 1 and all other components zero. Then, noting that α_0 is *not* a summation index, we have

$$R_{\alpha_0\alpha_0\beta\gamma} = 0 \tag{5.28}$$

That is, the diagonal terms of the subblock are zero. We can express the summation in (5.27) as

$$R_{\alpha\eta\beta\gamma}\xi^{\alpha}\xi^{\eta} = \tfrac{1}{2}(R_{\alpha\eta\beta\gamma} + R_{\eta\alpha\beta\gamma})\xi^{\alpha}\xi^{\eta} = 0 \tag{5.29}$$

Now we choose ξ to have two nonzero components, ξ^{μ_0} and ξ^{ν_0}, both equal to 1. Since the diagonal terms are zero according to (5.28), we have

$$R_{\mu_0\nu_0\beta\gamma} + R_{\nu_0\mu_0\beta\gamma} = 0 \tag{5.30}$$

Since μ_0 and ν_0 are any two components, we have the resultant antisymmetry relation

$$R_{\alpha\eta\beta\gamma} = -R_{\eta\alpha\beta\gamma} \tag{5.31}$$

That is, the Riemann tensor is antisymmetric in the first and second indices. We now have at most 6 independent components in the $\alpha\eta$ subblock, so together with the 6 in the $\beta\gamma$ subblock, there remain at most 36 independent components.

To obtain the final symmetry property, consider an antisymmetrized tensor formed from an arbitrary vector ξ_α:

$$\{\xi_{\alpha\|\beta\|\gamma} - \xi_{\beta\|\alpha\|\gamma}\} = \{\{\xi_{\alpha|\beta}\}_{\|\gamma}\} \tag{5.32}$$

By the general property of antisymmetrized tensors (Chap. 3), we may replace the covariant derivatives by ordinary derivatives, so we have

$$\{\xi_{\alpha\|\beta\|\gamma} - \xi_{\beta\|\alpha\|\gamma}\} = \{\{\xi_{\alpha|\beta}\}_{|\gamma}\} \tag{5.33}$$

Furthermore, we know that an exact antisymmetric tensor is closed, so that the repeated application of differentiation with antisymmetrization always produces a null tensor, and the above expression must be identically zero:

$$\{\xi_{\alpha\|\beta\|\gamma} - \xi_{\beta\|\alpha\|\gamma}\} = 0 \tag{5.34}$$

In the foregoing we have repeated the same considerations which led to the preceding antisymmetry law (5.31); only we started with the vector ξ_α instead of the scalar φ. On the other hand, the indices of the second term in the above null expression may be cyclically permuted without changing the expression. This gives

$$\{\xi_{\alpha\|\beta\|\gamma} - \xi_{\alpha\|\gamma\|\beta}\} = 0 \tag{5.35}$$

By virtue of Eq. (5.11) we may write this as

$$\{\xi_{\alpha\|\beta\|\gamma} - \xi_{\alpha\|\gamma\|\beta}\} = \{R_{\alpha\eta\beta\gamma}\xi^{\eta}\} = 0 \tag{5.36}$$

It is convenient to introduce here the notational convention

$$\{R_{\alpha\eta\beta\gamma}\xi^{\eta}\}_{(\alpha,\beta,\gamma)} = 0 \tag{5.37}$$

which means that only the indices α, β, and γ are included in the antisymmetrization, not the summation index η. Since η does not enter the antisymmetrization process, we can take ξ^η outside the brackets.

$$\{R_{\alpha\eta\beta\gamma}\xi^{\eta}\}_{(\alpha,\beta,\gamma)} = \{R_{\alpha\eta\beta\gamma}\}_{(\alpha,\beta,\gamma)}\xi^{\eta} = 0 \tag{5.38}$$

Thus, since ξ^η is an arbitrary vector, we have

$$\{R_{\alpha\eta\beta\gamma}\}_{(\alpha,\beta,\gamma)} = 0 \tag{5.39}$$

Relabeling the indices, we have also the relation

$$\{R_{\eta\alpha\beta\gamma}\}_{(\eta,\beta,\gamma)} = 0 \tag{5.40}$$

Let us write out these last two relations, (5.39) and (5.40); making use of the antisymmetry in the first and second indices and in the third and fourth indices, we obtain

(5.41) $$R_{\alpha\eta\beta\gamma} + R_{\beta\eta\gamma\alpha} + R_{\gamma\eta\alpha\beta} = 0$$

(5.42) $$R_{\eta\alpha\beta\gamma} + R_{\beta\alpha\gamma\eta} + R_{\gamma\alpha\eta\beta} = 0$$

Subtracting (5.42) from (5.41), remembering the antisymmetry in the first two indices, we get

(5.43) $$2R_{\alpha\eta\beta\gamma} + R_{\beta\eta\gamma\alpha} + R_{\gamma\eta\alpha\beta} - R_{\beta\alpha\gamma\eta} - R_{\gamma\alpha\eta\beta} = 0$$

or equivalently,

(5.44) $$2R_{\alpha\eta\beta\gamma} - R_{\eta\beta\gamma\alpha} - R_{\eta\gamma\alpha\beta} - R_{\alpha\beta\eta\gamma} - R_{\alpha\gamma\beta\eta} = 0$$

Regrouping terms, we can write this as

(5.45) $$2R_{\alpha\eta\beta\gamma} - (R_{\eta\beta\gamma\alpha} + R_{\alpha\beta\eta\gamma}) - (R_{\eta\gamma\alpha\beta} + R_{\alpha\gamma\beta\eta}) = 0$$

This can be simplified if we relabel the indices in (5.41) and (5.42) to give the pair of relations

(5.46) $$R_{\eta\beta\gamma\alpha} + R_{\alpha\beta\eta\gamma} = -R_{\gamma\beta\alpha\eta}$$

(5.47) $$R_{\eta\gamma\alpha\beta} + R_{\alpha\gamma\beta\eta} = -R_{\beta\gamma\eta\alpha}$$

which we then may substitute in (5.45) to get

(5.48) $$2R_{\alpha\eta\beta\gamma} + R_{\gamma\beta\alpha\eta} + R_{\beta\gamma\eta\alpha} = 0$$

But the Riemann tensor is antisymmetric in the first two and in the last two indices, so the last two terms of (5.48) are equal. Thus

(5.49) $$2(R_{\alpha\eta\beta\gamma} + R_{\gamma\beta\alpha\eta}) = 2(R_{\alpha\eta\beta\gamma} - R_{\beta\gamma\alpha\eta}) = 0$$

We then obtain, finally,

(5.50) $$R_{\alpha\eta\beta\gamma} = R_{\beta\gamma\alpha\eta}$$

which is our final symmetry property. Thus we see that $R_{\alpha\eta\beta\gamma}$ is *antisymmetric* in each of the index pairs $\alpha\eta$ and $\beta\gamma$, but is *symmetric* under interchange of the pairs. If we treat the first index pair as one index capable of assuming six values, and similarly for the second pair, we take

account of the antisymmetry in those pairs. Then the symmetry under interchange of the pairs implies that we have left at most the same number of independent components as a symmetric 6×6 matrix, that is, 21 independent components. Let us repeat in summary the symmetry properties we have obtained:

(5.51) $$\begin{aligned} R_{\alpha\eta\beta\gamma} &= -R_{\alpha\eta\gamma\beta} \\ R_{\alpha\eta\beta\gamma} &= -R_{\eta\alpha\beta\gamma} \\ R_{\alpha\eta\beta\gamma} &= R_{\beta\gamma\alpha\eta} \end{aligned}$$

There is, actually, one more symmetry property contained in (5.39) which is not contained in the set (5.51) and which reduces the number of independent components ultimately to 20. The additional condition can be expressed as follows: Let us consider the component R_{0123} of the Riemann tensor. Under all possible permutations of the indices 0, 1, 2, 3, we obtain 4!, or 24, formally different components. However, when applying (5.51) to R_{0123}, the indices 0 and 1, on the one hand, and 2 and 3, on the other hand, remain adjacent, and we can identify all components as multiples of the basic three

(5.52) $$R_{0123},\ R_{0231},\ R_{0312}$$

But from the cyclic symmetry (5.42) it follows that

(5.53) $$R_{1023} + R_{2031} + R_{3012} = 0$$

which establishes a relation among the components (5.52) and is therefore not contained in (5.51). On the other hand, (5.39) is important in its own right and will be useful in the following sections, so we repeat it also in our summary of symmetry properties:

(5.54) $$\{R_{\alpha\eta\beta\gamma}\}_{(\alpha,\beta,\gamma)} = 0$$

5.4 The Bianchi Identities

In the preceding section we obtained several useful algebraic symmetry relations for the Riemann curvature tensor $R_{\alpha\eta\beta\gamma}$, which are summarized in (5.51) and (5.54). In this section we shall use these relations to obtain a set of symmetry relations on the covariant derivatives of the Riemann tensor.

Lowering the index α in Eq. (5.11) gives

$$\xi_{\alpha\|\beta\|\gamma} - \xi_{\alpha\|\gamma\|\beta} = R_{\alpha\eta\beta\gamma}\xi^{\eta} \tag{5.55}$$

Using the product rule for covariant derivatives, we then obtain

$$(\xi_{\alpha\|\beta\|\gamma} - \xi_{\alpha\|\gamma\|\beta})_{\|\delta} = R_{\alpha\eta\beta\gamma\|\delta}\xi^{\eta} + R_{\alpha\eta\beta\gamma}\xi^{\eta}{}_{\|\delta} \tag{5.56}$$

Let us antisymmetrize this with respect to the indices β, γ, and δ:

$$\{\xi_{\alpha\|\beta\|\gamma\|\delta} - \xi_{\alpha\|\gamma\|\beta\|\delta}\}_{(\beta,\gamma,\delta)} = \{R_{\alpha\eta\beta\gamma\|\delta}\}_{(\beta,\gamma,\delta)}\xi^{\eta} + \{R_{\alpha\eta\beta\gamma}\xi^{\eta}{}_{\|\delta}\}_{(\beta,\gamma,\delta)} \tag{5.57}$$

By the nature of the process of antisymmetrization, we may make an even permutation of the indices β, γ, and δ on the left side of the above equation to obtain

$$\{(\xi_{\alpha\|\delta})_{\|\beta\|\gamma} - (\xi_{\alpha\|\delta})_{\|\gamma\|\beta}\}_{(\beta,\gamma,\delta)} = \{R_{\alpha\eta\beta\gamma\|\delta}\}_{(\beta,\gamma,\delta)}\xi^{\eta} + \{R_{\alpha\eta\beta\gamma}\xi^{\eta}{}_{\|\delta}\}_{(\beta,\gamma,\delta)} \tag{5.58}$$

Note that the tensor $\xi_{\alpha\|\delta}$ appears in both terms of the left side. Applying Eq. (5.19) to this tensor, we have

$$(\xi_{\alpha\|\delta})_{\|\beta\|\gamma} - (\xi_{\alpha\|\delta})_{\|\gamma\|\beta} = R_{\alpha\eta\beta\gamma}\xi^{\eta}{}_{\|\delta} + R_{\delta\eta\beta\gamma}\xi_{\alpha}{}^{\|\eta} \tag{5.59}$$

where $\xi_{\alpha}{}^{\|\eta} = g^{\eta\nu}\xi_{\alpha\|\nu}$ is the contravariant form of $\xi_{\alpha\|\eta}$. Antisymmetrizing, we obtain an alternative form for the left side of (5.58).

$$\{(\xi_{\alpha\|\delta})_{\|\beta\|\gamma} - (\xi_{\alpha\|\delta})_{\|\gamma\|\beta}\}_{(\beta,\gamma,\delta)} = \{R_{\alpha\eta\beta\gamma}\xi^{\eta}{}_{\|\delta}\}_{(\beta,\gamma,\delta)} + \{R_{\delta\eta\beta\gamma}\}_{(\beta,\gamma,\delta)}\xi_{\alpha}{}^{\|\eta} \tag{5.60}$$

Comparing (5.58) and (5.60), we then have

$$\{R_{\alpha\eta\beta\gamma\|\delta}\}_{(\beta,\gamma,\delta)}\xi^{\eta} = \{R_{\delta\eta\beta\gamma}\}_{(\beta,\gamma,\delta)}\xi_{\alpha}{}^{\|\eta} \tag{5.61}$$

However, the right side of the above is zero by virtue of the symmetry relation (5.39), so since ξ^{η} is an arbitrary vector, we have finally

$$\{R_{\alpha\eta\beta\gamma\|\delta}\}_{(\beta,\gamma,\delta)} = 0 \tag{5.62}$$

The symmetry relations contained in (5.62) are known as the Bianchi identities. They will be quite useful when we investigate the gravitational field equations in Sec. 5.8.

One could expect a priori that there should exist numerous relations between the 20 components of the Riemann tensor. This is because it is constructed out of the 10 components of the metric tensor, and a large number of consistency conditions are clearly necessary in order that a tensor with the symmetries obtained in Sec. 5.3 be of such form.

5.5 Integrability and the Riemann Tensor

By our definition of parallel displacement, the change in the α component of a vector ξ^{α} under parallel displacement is given by

$$d\xi^{\alpha} = -\begin{Bmatrix} \alpha \\ \beta \;\; \gamma \end{Bmatrix} \xi^{\beta}\, dx^{\gamma} \tag{5.63}$$

If we now are given a curve between two points and the value of ξ^{α} at one of the points, the above equation enables us to compute ξ^{α} at the other point; however, if we displace the same vector along a *different* curve connecting the same two points, we have no reason to expect to arrive at the same value of ξ^{α} at the endpoint. Equivalently, if we displace a vector parallel to itself around a *closed path*, we have no reason to expect the ξ^{α} to return to their initial values. We shall be concerned in this section with the question of how the change in parallel displaced vectors depends upon the path taken. That is, we shall study the dependence of the functional

$$I^{\alpha}(\Gamma) = \int_{\Gamma} d\xi^{\alpha} \tag{5.64}$$

on the path Γ.

Let us first consider the simplest and most important case. We assume that we start with an arbitrary vector at an arbitrarily given point and construct from it by parallel displacement a uniquely determined vector field in some finite neighborhood. This is clearly possible only if the result of the parallel displacement is independent of the path used to reach the final point considered. We thereby generate a vector field:

$$\xi^{\alpha} = \xi^{\alpha}(x^{\mu}) \tag{5.65}$$

The law of parallel displacement (5.63) then states that the derivatives of ξ^{α} are given by

$$\xi^{\alpha}{}_{|\gamma} = \frac{\partial \xi^{\alpha}}{\partial x^{\gamma}} = -\begin{Bmatrix} \alpha \\ \beta \;\; \gamma \end{Bmatrix} \xi^{\beta} \tag{5.66}$$

Note that this implies that the covariant derivatives of ξ^{α} are zero. Furthermore, the order of ordinary differentiation of a vector field is irrelevant; that is,

$$\xi^{\alpha}{}_{|\gamma|\delta} = \xi^{\alpha}{}_{|\delta|\gamma} \tag{5.67}$$

We see that we must have, according to (5.66),

$$\left(\left\{{\alpha \atop \beta\ \gamma}\right\}\xi^\beta\right)_{|\delta} = \left(\left\{{\alpha \atop \beta\ \delta}\right\}\xi^\beta\right)_{|\gamma} \tag{5.68}$$

This can be put in more meaningful form by applying the product rule for ordinary derivatives,

$$\left\{{\alpha \atop \beta\ \gamma}\right\}_{|\delta}\xi^\beta + \left\{{\alpha \atop \beta\ \gamma}\right\}\xi^\beta{}_{|\delta} = \left\{{\alpha \atop \beta\ \delta}\right\}_{|\gamma}\xi^\beta + \left\{{\alpha \atop \beta\ \delta}\right\}\xi^\beta{}_{|\gamma} \tag{5.69}$$

and substituting for the first derivatives of ξ^β from (5.66). We obtain

$$\left\{{\alpha \atop \beta\ \gamma}\right\}_{|\delta}\xi^\beta - \left\{{\alpha \atop \beta\ \gamma}\right\}\left\{{\beta \atop \tau\ \delta}\right\}\xi^\tau = \left\{{\alpha \atop \beta\ \delta}\right\}_{|\gamma}\xi^\beta - \left\{{\alpha \atop \beta\ \delta}\right\}\left\{{\beta \atop \tau\ \gamma}\right\}\xi^\tau \tag{5.70}$$

Rearranging terms and relabeling indices, we then get

$$\left[\left\{{\alpha \atop \beta\ \gamma}\right\}_{|\delta} - \left\{{\alpha \atop \beta\ \delta}\right\}_{|\gamma} + \left\{{\alpha \atop \tau\ \delta}\right\}\left\{{\tau \atop \beta\ \gamma}\right\} - \left\{{\alpha \atop \tau\ \gamma}\right\}\left\{{\tau \atop \beta\ \delta}\right\}\right]\xi^\beta = 0 \tag{5.71}$$

or by definition of the Riemann tensor,

$$R^\alpha{}_{\beta\gamma\delta}\xi^\beta = 0 \tag{5.72}$$

Since ξ^β can be arbitrarily chosen at any single given point, we must then have

$$R^\alpha{}_{\beta\gamma\delta} = 0 \qquad \text{(condition for integrability)} \tag{5.73}$$

Thus we see that a space must be flat (have a null Riemann curvature tensor) if we are able to establish in it a vector field by parallel displacement, starting with an arbitrary value at an arbitrary point of the space.

Let us next investigate the case of a space whose Riemann curvature tensor is not necessarily zero. We shall study the effect of a nonzero Riemann tensor on the parallel displacement of a vector. Consider two paths between some initial point P_i and a final point P_f. One path consists of a displacement along the vector **dx**, followed by a displacement along $\mathbf{d\hat{x}}$, and the other is along the same vectors in reverse order as illustrated in Fig. 5.1. We shall compute the change of ξ^α along

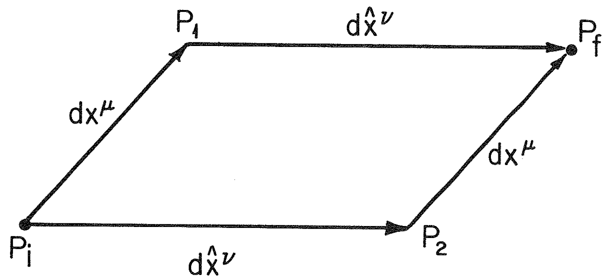


Fig. 5.1

both paths and compare the results. By the law of parallel displacement, the change of ξ^α between P_i and P_1 (along the first path) is

$$d\xi^\alpha(P_i,P_1) = -\left\{{\alpha \atop \eta\ \rho}\right\}\xi^\rho\,dx^\eta \tag{5.74}$$

(Unless explicitly noted, the vector ξ^α and the Christoffel symbols are always evaluated at P_i as in the right side above.) At P_1 the displaced vector is then given by

$$\xi^\alpha(P_1) = \xi^\alpha - \left\{{\alpha \atop \eta\ \rho}\right\}\xi^\rho\,dx^\eta \tag{5.75}$$

The Christoffel symbols at P_1 are, to first order in a Taylor series expansion in the vector **dx**,

$$\left\{{\alpha \atop \gamma\ \beta}\right\}_{P_1} = \left\{{\alpha \atop \gamma\ \beta}\right\} + \left\{{\alpha \atop \gamma\ \beta}\right\}_{|\eta}dx^\eta \tag{5.76}$$

Next applying the law of parallel displacement to the vector $\xi^\alpha(P_1)$, using the Christoffel symbols at P_1 given by (5.76), we obtain the change in ξ^α between P_1 and P_f to second order in the displacement vectors **dx** and $\mathbf{d\hat{x}}$:

$$d\xi^\alpha(P_1,P_f) = -\left[\left\{{\alpha \atop \gamma\ \beta}\right\} + \left\{{\alpha \atop \gamma\ \beta}\right\}_{|\eta}dx^\eta\right]\left[\xi^\beta - \left\{{\beta \atop \eta\ \rho}\right\}\xi^\rho\,dx^\eta\right]d\hat{x}^\gamma \tag{5.77}$$

Rearranging terms, we can write this in the form

$$d\xi^\alpha(P_1,P_f) = -\left\{{\alpha \atop \gamma\ \beta}\right\}\xi^\beta\,d\hat{x}^\gamma - \left\{{\alpha \atop \gamma\ \beta}\right\}_{|\eta}\xi^\beta\,dx^\eta\,d\hat{x}^\gamma + \left\{{\alpha \atop \gamma\ \beta}\right\}\left\{{\beta \atop \eta\ \rho}\right\}\xi^\rho\,dx^\eta\,d\hat{x}^\gamma \tag{5.78}$$

Again, as we noted above, the terms on the right side are evaluated at P_i. Thus the value of ξ^α after traversing the entire path is, to second order in the displacement vectors $\mathbf{dx}$ and $\mathbf{d\hat{x}}$,

$$(5.79)\quad \xi^\alpha(P_iP_1P_f) = \xi^\alpha + d\xi^\alpha(P_iP_1) + d\xi^\alpha(P_1P_f)$$
$$= \xi^\alpha - \begin{Bmatrix}\alpha\\ \eta\ \beta\end{Bmatrix}\xi^\beta\,dx^\eta - \begin{Bmatrix}\alpha\\ \gamma\ \beta\end{Bmatrix}\xi^\beta\,d\hat{x}^\gamma - \begin{Bmatrix}\alpha\\ \gamma\ \beta\end{Bmatrix}_{|\eta}\xi^\beta\,dx^\eta\,d\hat{x}^\gamma + \begin{Bmatrix}\alpha\\ \gamma\ \lambda\end{Bmatrix}\begin{Bmatrix}\lambda\\ \eta\ \beta\end{Bmatrix}\xi^\beta\,dx^\eta\,d\hat{x}^\gamma$$

The result along the second path, $P_i \to P_2 \to P_f$, is gotten by simply interchanging $\mathbf{dx}$ and $\mathbf{d\hat{x}}$:

$$(5.80)\quad \xi^\alpha(P_iP_2P_f) = \xi^\alpha - \begin{Bmatrix}\alpha\\ \gamma\ \beta\end{Bmatrix}\xi^\beta\,d\hat{x}^\gamma - \begin{Bmatrix}\alpha\\ \eta\ \beta\end{Bmatrix}\xi^\beta\,dx^\eta - \begin{Bmatrix}\alpha\\ \eta\ \beta\end{Bmatrix}_{|\gamma}\xi^\beta\,d\hat{x}^\gamma\,dx^\eta + \begin{Bmatrix}\alpha\\ \eta\ \lambda\end{Bmatrix}\begin{Bmatrix}\lambda\\ \gamma\ \beta\end{Bmatrix}\xi^\beta\,d\hat{x}^\gamma\,dx^\eta$$

The difference between the ξ^α obtained by parallel displacement along the two routes is therefore

$$(5.81)\quad \Delta\xi^\alpha = \begin{Bmatrix}\alpha\\ \eta\ \beta\end{Bmatrix}_{|\gamma}\xi^\beta\,d\hat{x}^\gamma\,dx^\eta - \begin{Bmatrix}\alpha\\ \gamma\ \beta\end{Bmatrix}_{|\eta}\xi^\beta\,d\hat{x}^\gamma\,dx^\eta + \begin{Bmatrix}\alpha\\ \gamma\ \lambda\end{Bmatrix}\begin{Bmatrix}\lambda\\ \eta\ \beta\end{Bmatrix}\xi^\beta\,dx^\eta\,d\hat{x}^\gamma - \begin{Bmatrix}\alpha\\ \eta\ \lambda\end{Bmatrix}\begin{Bmatrix}\lambda\\ \gamma\ \beta\end{Bmatrix}\xi^\beta\,dx^\eta\,d\hat{x}^\gamma$$

By definition of the Riemann tensor, this is precisely

$$(5.82)\qquad \Delta\xi^\alpha = R^\alpha{}_{\beta\eta\gamma}\xi^\beta\,dx^\eta\,d\hat{x}^\gamma$$

Thus the value of ξ^α at the nearby point is independent of path if and only if $R^\alpha{}_{\beta\eta\gamma} = 0$, and for a nonzero Riemann tensor the difference in final value is given by (5.82).

We may avoid the use of too many differentials while studying the integrability conditions by use of the methods of calculus of variations. Indeed, let ξ^β be an arbitrary vector field and consider the integral

$$(5.83)\qquad J^\alpha[C] = -\int_C \begin{Bmatrix}\alpha\\ \beta\ \gamma\end{Bmatrix}\xi^\beta\,dx^\gamma$$

along a curve C which connects two given points in our Riemann space, say, P_i and P_f. In general, this integral will be a function of the curve C; we wish to find out for which vector fields this integral is path-independent. We perform a variation δx^γ which vanishes at the endpoints of C and find

$$(5.84)\quad \delta J^\alpha = -\int_C\left[\begin{Bmatrix}\alpha\\ \beta\ \gamma\end{Bmatrix}_{|\rho}\xi^\beta\,\delta x^\rho\,dx^\gamma + \begin{Bmatrix}\alpha\\ \beta\ \gamma\end{Bmatrix}\xi^\beta{}_{|\rho}\,\delta x^\rho\,dx^\gamma + \begin{Bmatrix}\alpha\\ \beta\ \gamma\end{Bmatrix}\xi^\beta\delta\,dx^\gamma\right]$$

We then integrate by parts over the last term and use the fact that $\delta x^\gamma = 0$ at the endpoints; we obtain, by use of definition (5.10),

$$(5.85)\quad \delta J^\alpha = \int_C\left[\begin{Bmatrix}\alpha\\ \beta\ \gamma\end{Bmatrix}_{|\rho}\xi^\beta + \begin{Bmatrix}\alpha\\ \beta\ \gamma\end{Bmatrix}\xi^\beta{}_{|\rho}\right](\delta x^\gamma\,dx^\rho - \delta x^\rho\,dx^\gamma)$$
$$= \int_C\left[R^\alpha{}_{\beta\gamma\rho}\xi^\beta + \begin{Bmatrix}\alpha\\ \beta\ \gamma\end{Bmatrix}\xi^\beta{}_{\|\rho} - \begin{Bmatrix}\alpha\\ \beta\ \rho\end{Bmatrix}\xi^\beta{}_{\|\gamma}\right]\delta x^\gamma\,dx^\rho$$

We may even choose ξ^β in such a way that on C all its covariant derivatives vanish. In this case,

$$(5.86)\qquad \delta J^\alpha = \int_C R^\alpha{}_{\beta\gamma\rho}\xi^\beta\,\delta x^\gamma\,dx^\rho$$

which is indeed just the integral form of (5.82).

Let us assume, on the other hand, that the Riemann tensor vanishes identically. We may define a vector field by the system of linear partial differential equations

$$(5.87)\qquad \xi^\alpha{}_{|\gamma} = -\begin{Bmatrix}\alpha\\ \beta\ \gamma\end{Bmatrix}\xi^\beta$$

The condition for integrability of this system was shown above to be the vanishing of the Riemann tensor, which is now fulfilled. It follows from the general theory of such differential systems that this condition is also sufficient for the existence of a solution. Hence $R_{\alpha\beta\gamma\delta} = 0$ indeed implies the existence of a vector field with zero covariant derivative $\xi^\alpha{}_{\|\gamma} = 0$. We may prescribe the initial values of the vector ξ^α at one point arbitrarily and continue this vector into the neighborhood by means of (5.87).

Let us now summarize the results of this section. We can establish a vector field $\xi^\alpha(x)$ by parallel displacement of an arbitrary vector ξ^α from some initial point to all points in a neighborhood in the Riemann space (by an arbitrary route) if and only if the Riemann tensor of the space is identically zero. In other words, we have found that (5.87) is integrable if and only if the space is flat, that is, has a zero Riemann curvature tensor. We shall henceforth refer to such a space as an integrable space. In addition, note that the covariant derivative of a vector field formed by parallel displacement is clearly zero. Therefore such a vector field is a natural generalization of a constant vector field in Cartesian coordinates. The results of this section then indicate that such a generalized constant vector field can exist only in a flat space with a zero Riemann curvature tensor.

5.6 Pseudo-Euclidean and Flat Spaces

From the previous section we know that we can establish a vector field with a zero covariant derivative if and only if the Riemann tensor is everywhere zero. We shall show in this section that the existence of such a generalized constant vector field ensures the existence of a coordinate system where the metric tensor has constant components. If one can find such a coordinate system where the metric tensor has constant components, the space is termed by definition a pseudo-Euclidean space. Thus we can say that the goal of this section is to show that a flat space (a space with a null Riemann tensor) is also pseudo-Euclidean.

Let us then suppose that the Riemann tensor of a space is everywhere zero; in that case we can establish a generalized constant vector field by the parallel displacement of some arbitrary vector ξ_α from an initial point to any given point in space. Let us do this with the following set of four vectors $\xi_\alpha^{(\gamma)}$:

$$\xi_\alpha^{(0)} = (1,0,0,0) \qquad \xi_\alpha^{(1)} = (0,1,0,0)$$
$$\xi_\alpha^{(2)} = (0,0,1,0) \qquad \xi_\alpha^{(3)} = (0,0,0,1) \tag{5.88}$$

which we can write more simply as

$$\xi_\alpha^{(\gamma)} = \delta_\alpha^{(\gamma)} = \begin{cases} 1 & \text{for } \gamma = \alpha \\ 0 & \text{for } \gamma \neq \alpha \end{cases} \tag{5.89}$$

(Note that γ is *not* a tensor index.) These vectors then represent the value of four generalized constant vector fields at an initial point, say, P_0, in some fixed but arbitrary coordinate system x^α:

$$\xi_\alpha^{(\gamma)}(P_0) = \delta_\alpha^{(\gamma)} \tag{5.90}$$

Now since each vector field $\xi_\alpha^{(\gamma)}(x)$ has a zero covariant derivative,

$$\xi_{\alpha\|\beta}^{(\gamma)} = 0 \tag{5.91}$$

(we omit the argument x for clarity), it must also have a zero curl:

$$\xi_{\alpha\|\beta}^{(\gamma)} - \xi_{\beta\|\alpha}^{(\gamma)} = \xi_{\alpha|\beta}^{(\gamma)} - \xi_{\beta|\alpha}^{(\gamma)} = 0 \tag{5.92}$$

The zero curl implies that each $\xi_\alpha^{(\gamma)}$ then has a scalar potential $\varphi^{(\gamma)}(x^\mu)$ such that

$$\xi_\alpha^{(\gamma)} = \frac{\partial\varphi^{(\gamma)}}{\partial x^\alpha} \tag{5.93}$$

Let us now use these scalar potential functions to define a transformation to a new coordinate system $\bar{x}^\gamma$:

$$\bar{x}^\gamma = \varphi^{(\gamma)}(x^\mu) \tag{5.94}$$

Such a transformation is permissible in *some neighborhood* of P_0 since the Jacobian of the transformation at P_0 is

$$\left\| \frac{\partial\varphi^{(\gamma)}}{\partial x^\alpha} \right\| = \|\xi_\alpha^{(\gamma)}\| = 1 \tag{5.95}$$

In the following development we shall consider only the neighborhood of P_0 where the Jacobian (5.95) remains positive until we note otherwise. In the barred system Eq. (5.93) takes the form

$$\bar{\xi}_\alpha^{(\gamma)} = \frac{\partial\varphi^{(\gamma)}}{\partial\bar{x}^\alpha} \tag{5.96}$$

which, by virtue of the transformation (5.94), gives

$$\bar{\xi}_\alpha^{(0)} = (1,0,0,0) \qquad \bar{\xi}_\alpha^{(1)} = (0,1,0,0)$$
$$\bar{\xi}_\alpha^{(2)} = (0,0,1,0) \qquad \bar{\xi}_\alpha^{(3)} = (0,0,0,1) \tag{5.97}$$