We see that constancy of components in one particular coordinate system requires that, in an arbitrary coordinate system, the much more general law (2.1) holds. Specifically, we recognize that the increment of the vector components is a bilinear function of the components ξ^j of the vector and of the displacement dx^m tangent to the curve along which the transplantation takes place.

In order to generalize the above notions, let us now ignore the origin of formula (2.1) and consider only its differential form

$$d\xi^i = \Gamma^i_{mj}\, dx^m\, \xi^j \tag{2.3}$$

The Γ^i_{mj} coefficients are now considered to be any set of given functions of the coordinates and are no longer restricted to have the form (2.2), which was derived from the fact that the vector field considered had constant components in the particular original coordinate system. The vector field ξ^i is considered to be obtained from its value at one given point by the transplantation law (2.3). Equation (2.3) defines a general law for transplantation of the vector ξ^i at the point x into the quantities $\xi^i + d\xi^i$ at the point $x + dx$. It is a law of affine character; that is, it has invariant structure under a linear transformation of the coordinates.

If we now try to make this law of transplantation coordinate-invariant and demand that $\xi^i + d\xi^i$ still be a vector at the point $x + dx$, we shall be forced into certain requirements on the Γ^i_{mj} coefficients; we shall see that these requirements define a transformation law for the Γ^i_{mj} coefficients and therefore allow us to transplant a vector by infinitesimal amounts in a covariant fashion.

Proof. In the unbarred coordinate system we take for the transplantation law

$$\xi^i(x + dx) = \xi^i + d\xi^i = \xi^i + \Gamma^i_{mj}\, dx^m\, \xi^j \tag{2.4}$$

and impose the requirement that the same law hold in the barred coordinate system (that is, that the law be covariant). We further require that $\xi^i(x + dx)$ be a vector; by definition this means

$$\bar{\xi}^j(x + dx) = \xi^i(x + dx)\left(\frac{\partial \bar{x}^j}{\partial x^i}\right)_{x+dx}$$

In this expression the geometric point at which the transformation is carried out is characterized by the markers $x + dx$ (which could also be called $\bar{x} + d\bar{x}$ since the two sets of markers are in one-to-one correspondence). Writing out this expression with the help of the transplantation law (2.4), we have

$$\bar{\xi}^j + \bar{\Gamma}^j_{ms}\, d\bar{x}^m\, \bar{\xi}^s = (\xi^i + \Gamma^i_{ml}\, dx^m\, \xi^l)\left(\frac{\partial \bar{x}^j}{\partial x^i}\right)_{x+dx}$$

Expanding the last factor in a Taylor series and keeping only terms up to the first order, we obtain

$$\left(\frac{\partial \bar{x}^j}{\partial x^i}\right)_{x+dx} = \left(\frac{\partial \bar{x}^j}{\partial x^i}\right)_x + \frac{\partial^2 \bar{x}^j}{\partial x^i\, \partial x^\eta}\, dx^\eta$$

Putting this into the previous equation and relabeling the dummy indices, we find that

$$\bar{\Gamma}^j_{ms}\, d\bar{x}^m\, \bar{\xi}^s = \left(\Gamma^i_{\alpha\beta}\frac{\partial \bar{x}^j}{\partial x^i} + \frac{\partial^2 \bar{x}^j}{\partial x^\beta\, \partial x^\alpha}\right)\xi^\beta\, dx^\alpha$$

But we can also express $dx^\alpha \xi^\beta$ in terms of barred quantities as

$$\xi^\beta\, dx^\alpha = \frac{\partial x^\alpha}{\partial \bar{x}^m}\frac{\partial x^\beta}{\partial \bar{x}^s}\,\bar{\xi}^s\, d\bar{x}^m$$

and obtain

$$\bar{\Gamma}^j_{ms}\, d\bar{x}^m\, \bar{\xi}^s = \left(\frac{\partial \bar{x}^j}{\partial x^i}\frac{\partial x^\alpha}{\partial \bar{x}^m}\frac{\partial x^\beta}{\partial \bar{x}^s}\Gamma^i_{\alpha\beta} + \frac{\partial^2 \bar{x}^j}{\partial x^\alpha\, \partial x^\beta}\frac{\partial x^\alpha}{\partial \bar{x}^m}\frac{\partial x^\beta}{\partial \bar{x}^s}\right)d\bar{x}^m\, \bar{\xi}^s$$

In this equation the coefficients $d\bar{x}^m$ and $\bar{\xi}^s$ are arbitrary; therefore we must have

$$\bar{\Gamma}^j_{ms} = \frac{\partial \bar{x}^j}{\partial x^i}\frac{\partial x^\alpha}{\partial \bar{x}^m}\frac{\partial x^\beta}{\partial \bar{x}^s}\Gamma^i_{\alpha\beta} + \frac{\partial^2 \bar{x}^j}{\partial x^\alpha\, \partial x^\beta}\frac{\partial x^\alpha}{\partial \bar{x}^m}\frac{\partial x^\beta}{\partial \bar{x}^s} \tag{2.5}$$

which is the transformation law of the coefficients $\Gamma^i_{\alpha\beta}$ into which we are forced by the covariance requirement. These coefficients are called *coefficients of affine connection* or simply *connections*. Note that the transformation law (2.5) is *inhomogeneous* in the coefficients $\Gamma^i_{\alpha\beta}$; thus it is fundamentally different from a tensor transformation law and *the connections are not tensors*.

With the above transformation law for $\Gamma^i_{\alpha\beta}$ we have set up a consistent covariant definition of the transplantation of a vector in terms of the increments of its components:

$$d\xi^i = \Gamma^i_{mj}\, dx^m\, \xi^j \tag{2.6}$$

We have now derived necessary conditions for the coefficient set Γ^i_{kl} in order that the transplantation law (2.6) may yield a vector $\xi^i + d\xi^i$ at the point $x^m + dx^m$ if ξ^i is a vector at the point x^m. The next question is how far this transplantation law is sufficient and how it works in the large. For this purpose we assume a field of connections $\Gamma^i_{kl}(x^m)$, that is, a coefficient set attached to each point of the finite part of space considered and which obeys the connection transformation law (2.5). Then we set up the differential-equation system for a vector field in the parameter p:

$$\frac{d\xi^i}{dp} = \Gamma^i_{kl} \frac{dx^k}{dp} \xi^l(p) \tag{2.6'}$$

which allows us to compute the vectors $\xi^i(p)$ along the given curve $x^i(p)$ from the known value $\xi^i(0)$ at the initial point $p = 0$. But we still have to prove that the n-tuples $\xi^i(p)$ thus calculated transform indeed like vectors under a change of coordinates.

To do this we consider the barred coordinates $\bar{x}^i(p)$ and the quantities $\bar{\xi}^i(p)$ defined by the corresponding system (2.6′) in barred terms. We suppose that the equations

$$\bar{\xi}^i(p) - \frac{\partial \bar{x}^i}{\partial x^m} \xi^m(p) = 0$$

hold for $p = 0$ since we assumed at least that $\xi^i(0)$ is a vector. By use of (2.6′) and the analogous differential equation for $\bar{\xi}^i(p)$, we can calculate for all values of p the derivative

$$\frac{d}{dp}\left[\bar{\xi}^i(p) - \frac{\partial \bar{x}^i}{\partial x^m} \xi^m(p)\right] = \bar{\Gamma}^i_{kl} \frac{d\bar{x}^k}{dp} \bar{\xi}^l - \frac{\partial^2 \bar{x}^i}{\partial x^r \, \partial x^s} \frac{dx^r}{dp} \xi^s - \Gamma^m_{rs} \frac{\partial \bar{x}^i}{\partial x^m} \frac{dx^r}{dp} \xi^s$$

If we now use the transformation law (2.5) which we take as hypothesis, we obtain after easy rearrangement

$$\frac{d}{dp}\left[\bar{\xi}^i(p) - \frac{\partial \bar{x}^i}{\partial x^m} \xi^m(p)\right] = \bar{\Gamma}^i_{kl} \frac{d\bar{x}^k}{dp}\left(\bar{\xi}^l - \frac{\partial \bar{x}^l}{\partial x^\alpha} \xi^\alpha\right)$$

The validity of the vector transformation law of $\xi^i(p)$ along the entire curve follows from the uniqueness theorem for this linear homogeneous differential system which admits the solution $\bar{\xi}^i(p) - \dfrac{\partial \bar{x}^i}{\partial x^m} \xi^m(p) = 0$ as initial value for $p = 0$.

There are several interesting consequences of the transformation law (2.5):

1. If we restrict ourselves to linear transformations of coordinates, the term $\partial^2 \bar{x}^j / \partial x^\alpha \, \partial x^\beta$ vanishes and the set of connections Γ^i_{mj} transforms like a tensor. If now some transplanted vector field ξ^i has constant components in a particular coordinate system, it is clear that the terms $\Gamma^i_{mj}\xi^j$ will be zero in that system, as one can see from (2.6). Furthermore, since the Γ transform like a tensor under linear transformations and ξ^j is a vector, these expressions are zero in all coordinate systems related to the original by a linear transformation; it follows that ξ^i has constant components in all such coordinate systems. In this special case constancy of components is therefore an acceptable criterion for a constant vector field.

2. If two fields of connections are given, say, Γ^i_{kl} and $\tilde{\Gamma}^i_{kl}$, their difference is a tensor. Indeed, since the inhomogeneous terms in the transformation law (2.5) are independent of the individual connections, they cancel under subtraction, and $\Gamma^i_{kl} - \tilde{\Gamma}^i_{kl}$ transforms like a tensor. This observation is of particular interest when one varies the connections over a given manifold. The variation of the connection is then a tensor.

3. We can, furthermore, make certain axiomatic deductions from the general form of (2.5). Historically, the connections were introduced in classical differential geometry; their role in transplanting vector fields was first clearly brought out by Levi-Civita. In all these formulations they appeared to be symmetric in their lower indices as in (2.2). But in our more general development we have built up a consistent transplantation law of a vector without the assumption of any particular symmetry of the Γ coefficients. This was pointed out in 1950 by Einstein (Einstein, 1955) and considered independently by Schrödinger (Schrödinger, 1950). Suppose, therefore, that we are given a set of Γ^i_{ms} coefficients such that $\Gamma^i_{ms} \neq \Gamma^i_{sm}$ in a particular coordinate system; we then say that:

a. The Γ coefficients remain unsymmetric under any change of coordinates.

b. It is impossible to find a coordinate system in which all Γ coefficients are 0 at a point.

The proof of these statements is quite simple. For the first one, suppose that we could find a coordinate transformation in which the Γ^i_{mj} coefficients were symmetric; then (2.5) would give $\bar{\Gamma}^i_{\alpha\beta} = \bar{\Gamma}^i_{\beta\alpha}$ in *any* coordinate system, which contradicts our hypothesis. For the second we proceed as above: if such a coordinate system existed, then, by (2.5), the Γ^i_{ms} coefficients would be symmetric in their lower indices in any coordinate system since the inhomogeneous term in (2.5) is symmetric in α and β; this contradicts the hypothesis and completes the proof.

The use of unsymmetric Γ^i_{ms} coefficients has been considered only in later developments of the theory of general relativity, in the attempt to unite electromagnetic theory and gravitation theory. We shall deal in this book only with the original "classical" form of general relativity theory, which uses symmetric Γ^i_{ms} coefficients. It is indeed only in this case that one can connect the law of transplantation (2.3) with the intuitive idea of transplantation in a Euclidean space. This will be proved in the next theorem, which is sometimes referred to as the axiomatic definition of transplantation of Weyl (Weyl, 1950). In a Euclidean space described by a rectilinear coordinate system, we are used to transplanting vectors by simply keeping their components constant and attaching them to different points; in a Riemann space we expect to be able to do the same thing *locally* (in the neighborhood of a point) if we choose the right type of coordinate system at that point. In such a coordinate system we should have $d\xi^i = 0$ and therefore $\bar{\Gamma}^j_{ms} = 0$. Considering the inhomogeneous character of the transformation (2.5), we should often expect to be able to find such a coordinate system.

In the theory of surfaces in three-space, the original object of tensor analysis, the manifold is imbedded in a Euclidean space which determines the metric in the surface. Here transplantation of a vector might be defined by constancy of components in the extrinsic Cartesian coordinate system of the three-space. Locally, we can always find a coordinate system in the surface which coincides with two of the space coordinates. This particular situation provides motivation and illustration for Weyl's conception of transplantation. Weyl considered only transplantations that belong to connections which are symmetric in their lower indices. He showed that such transplantations could be completely characterized by the statement: At every point of the manifold there exists a local coordinate system in which the fields of the transplanted vectors possess constant components under infinitesimal displacement from that point.

From our more general point of view, Weyl's statement is contained in the following theorem.

Theorem. The necessary and sufficient condition for the existence of a particular local coordinate system in which the components of a vector are not altered by an infinitesimal transplantation according to the law (2.3) is that the coefficients of affine connection be symmetric in their lower indices.

Proof of the Necessary Condition. Suppose that, in a particular coordinate system $\bar{x}^i$, the components of an arbitrary vector ξ^s are unaltered under an infinitesimal transplantation from a given point P. This means that $d\xi^i = 0$ in that coordinate system, and therefore that $\bar{\Gamma}^j_{ms}\, d\bar{x}^m \xi^s = 0$. The product $d\bar{x}^m\, \xi^s$ is arbitrary; therefore $\bar{\Gamma}^j_{ms} = 0$ must be satisfied. Then, from (2.5), the coefficients $\Gamma^i_{\alpha\beta}$ will be symmetric in their lower indices in *any* coordinate system, for the inhomogeneous term and the coefficient of $\Gamma^i_{\alpha\beta}$ in (2.5) are symmetric in α and β.

Proof of the Sufficient Condition. Let us choose the point P as the origin of the coordinate system $x^i = 0$, and let us look for a particular coordinate system $\bar{x}^i$ (as mentioned in the theorem) by setting up the transformation

$$\bar{x}^i = x^i + \tfrac{1}{2}A^i_{jk}x^jx^k \qquad \left(\frac{\partial \bar{x}^i}{\partial x^j}\right)_{x=0} = \delta^i_j$$

where the A^i_{jk} coefficients have yet to be specified. From (2.5) the Γ^i_{ms} coefficients transform from unbarred to the above defined barred coordinates according to the equation

$$\bar{\Gamma}^j_{ms} = \Gamma^i_{\alpha\beta}\,\delta^j_i\,\delta^\alpha_m\,\delta^\beta_s + \tfrac{1}{2}(A^j_{\alpha\beta} + A^j_{\beta\alpha})\,\delta^\alpha_m\,\delta^\beta_s$$

which becomes

$$\bar{\Gamma}^j_{ms} = \Gamma^j_{ms} + \tfrac{1}{2}(A^j_{ms} + A^j_{sm})$$

The quantity $(A^j_{ms} + A^j_{sm})$ is symmetric in s and m. With the hypothesis that Γ^j_{ms} is also symmetric, we can choose the A^j_{ms} coefficients so that $\tfrac{1}{2}(A^j_{ms} + A^j_{sm}) = -\Gamma^j_{ms}$ and obtain $\bar{\Gamma}^j_{ms} = 0$; therefore $d\xi^i = 0$. Thus determined, the A^j_{ms} coefficients define a coordinate system in which the intuitive notion of transplantation by constancy of components is *locally* applicable.

The coordinate system obtained in the preceding paragraph is referred to as a *geodesic coordinate system* with respect to the connection Γ; it is clearly defined only locally. Furthermore, it is defined only up to a linear transformation, as we shall now show. From (2.5) we see that two coordinate systems x^i and $\bar{x}^j$ in which $\Gamma^i_{\alpha\beta}$ and $\bar{\Gamma}^j_{ms}$ are both zero must be related by

$$\frac{\partial^2 \bar{x}^j}{\partial x^\alpha\,\partial x^\beta}\,\frac{\partial x^\alpha}{\partial \bar{x}^m}\,\frac{\partial x^\beta}{\partial \bar{x}^s} = 0$$

Consider these equations as a system of four homogeneous equations labeled by s:

$$\left(\frac{\partial^2 \bar{x}^j}{\partial x^\alpha\,\partial x^\beta}\,\frac{\partial x^\alpha}{\partial \bar{x}^m}\right)\frac{\partial x^\beta}{\partial \bar{x}^s} = 0$$

with $\dfrac{\partial^2 \bar{x}^j}{\partial x^\alpha\,\partial x^\beta}\,\dfrac{\partial x^\alpha}{\partial \bar{x}^m}$ as unknowns; their determinant is the Jacobian of the coordinate transformation, which we always assume is different from zero.

Therefore the only solution of the system is

$$\frac{\partial^2 \bar{x}^j}{\partial x^\alpha\,\partial x^\beta}\frac{\partial x^\alpha}{\partial \bar{x}^m} = 0$$

But this again is a system of four homogeneous equations labeled by m, whose only solution is

$$\frac{\partial^2 \bar{x}^j}{\partial x^\alpha\,\partial x^\beta} = 0$$

This shows that x^i and $\bar{x}^j$ are indeed related locally by a linear transformation as stated above. A linear one-to-one coordinate transformation is called an affine transformation. The connections $\Gamma^i_{\alpha\beta}$ are called affine connections because of their linear transformation law (2.5).

As mentioned in our remark in Sec. 1.4, the existence of a geodesic coordinate system at each point in space provides us with a powerful tool in tensor calculus because of the very simple form which the $\Gamma^i_{\alpha\beta}$ coefficients have in such a coordinate system; that is, they are zero at the selected point. In the following chapters we shall always assume that the $\Gamma^i_{\alpha\beta}$ coefficients are symmetric in $\alpha\beta$ and that a geodesic coordinate system therefore exists. It should be kept in mind that the $\Gamma^i_{\alpha\beta}$ coefficients do not transform like tensors, but according to (2.5).

We end this section by mentioning some definitions often used in the literature in connection with the law of vector transplantation. A manifold in which a law of vector transplantation is defined is called "affinely related manifold," or an "affine space," and the Γ coefficients are called the "affine connections." We shall use the name *law of vector transplantation* for the general case treated in this section. Some authors use the name "law of parallel displacement" interchangeably with this, but we shall reserve the latter term for the more specific case of a metric space, which will be treated in the next section.

2.2 Parallel Displacement—Christoffel Symbols

In the previous section we introduced a law of *vector transplantation* on an affinely related manifold with the help of the coefficients of affine connection Γ. No metric properties were required to carry out such a program. In this section we particularize our study to Riemann spaces; we shall impose on the transplantation law previously defined the metric requirement that the scalar product of two vectors be invariant under the transplantation. In particular, the length of a vector will then remain

unchanged under the transplantation, as is the case with rectilinear coordinates in Euclidean space. This will give a unique determination of the Γ^i_{jk} coefficients as functions of the components of the metric tensor g_{ik} and their first derivatives.

Let us consider an infinitesimal displacement along a curve and express the fact that the scalar product of two vectors ξ^i and η^k remains constant as they are transplanted along the curve. The scalar product is $g_{ik}\xi^i\eta^k$, so our condition is

$$\frac{d}{ds}(g_{ik}\xi^i\eta^k) = 0$$

where ds is the element of arc along the curve. Expanded, this becomes

$$\frac{\partial g_{ik}}{\partial x^l}\frac{dx^l}{ds}\xi^i\eta^k + g_{ik}\frac{d\xi^i}{ds}\eta^k + g_{ik}\xi^i\frac{d\eta^k}{ds} = 0$$

Taking into account the transplantation law (2.1) for $d\xi^i/ds$ and $d\eta^k/ds$ and relabeling the dummy indices, we obtain for the coefficients of the arbitrary multinomial $\xi^i\eta^k\,dx^l$ the equation

$$\frac{\partial g_{ik}}{\partial x^l} + g_{rk}\Gamma^r_{il} + g_{ir}\Gamma^r_{lk} = 0 \tag{2.7a}$$

From this equation we can obtain a unique determination of Γ^r_{il}. Let us cyclically permute the indices ikl in (2.7a) to obtain the following two additional equations:

$$\frac{\partial g_{kl}}{\partial x^i} + g_{rl}\Gamma^r_{ki} + g_{kr}\Gamma^r_{il} = 0 \tag{2.7b}$$

$$\frac{\partial g_{li}}{\partial x^k} + g_{ri}\Gamma^r_{lk} + g_{lr}\Gamma^r_{ki} = 0 \tag{2.7c}$$

Using the symmetry of g_{ik} and of Γ^l_{ik}, we obtain, by adding (2.7c) and (2.7b) and subtracting (2.7a),

$$\frac{\partial g_{kl}}{\partial x^i} + \frac{\partial g_{li}}{\partial x^k} - \frac{\partial g_{ik}}{\partial x^l} + 2\Gamma^r_{ki}g_{rl} = 0$$

Thus

$$\Gamma^r_{ki} = -\tfrac{1}{2}g^{lr}\left(\frac{\partial g_{kl}}{\partial x^i} + \frac{\partial g_{li}}{\partial x^k} - \frac{\partial g_{ik}}{\partial x^l}\right)$$

For convenience in notation let us define

$$[ik,l] = \frac{1}{2}\left(\frac{\partial g_{il}}{\partial x^k} + \frac{\partial g_{kl}}{\partial x^i} - \frac{\partial g_{ik}}{\partial x^l}\right) \tag{2.8}$$

which we call the Christoffel symbol of the first kind (Christoffel, 1869), and

$$\left\{ {j \atop i\ k} \right\} = g^{jl}[ik,l] \tag{2.9}$$

which we call the *Christoffel symbol of the second kind.* With this notation the coefficients of connection can be simply written as

$$\Gamma^r_{ik} = -\left\{ {r \atop i\ k} \right\} \tag{2.10}$$

We now possess the unique determination of the coefficients of connection for which the scalar product of two vectors remains constant under the law of vector transplantation. We thus arrive at the *law of parallel displacement* in a metric space:

$$d\xi^i = -\left\{ {i \atop \alpha\ \beta} \right\} dx^\alpha\, \xi^\beta \tag{2.11}$$

The behavior of $\left\{ {i \atop \alpha\ \beta} \right\}$ under a coordinate transformation is the same as that of $-\Gamma^i_{\alpha\beta}$, which is given by (2.5).

At this point it might be appropriate to inject a word of comfort to the physicist. The introduction of complicated new symbols in the midst of tensor formalism which is already loaded with index conventions may seem unpleasant to a physicist, and the whole subject may appear to be hidden behind symbolism. In fact, we shall show here that Christoffel symbols actually occur in mechanics, but rarely appear in explicit form because of simplifications due to the very simple mechanical systems usually considered.

Let us consider the evolution in time of a mechanical system described by generalized coordinates $x^i(t)$, generalized velocities $\dot{x}^i = dx^i/dt$, a kinetic-energy quadratic form $T = \frac{1}{2}g_{ik}\dot{x}^i\dot{x}^k$, and a potential energy $V(x^i)$ which gives rise to a generalized force $F_i = -\partial V/\partial x^i$. As usual in analytical dynamics, we take $T\,dt^2 = ds^2$ to define a metric on the space of the generalized coordinates, which is called *configuration space.* In terms of the Lagrangian, $L = T - V$, we can write down the Lagrange equations of motion

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = \frac{\partial L}{\partial x^i} \tag{2.12}$$

In explicit form these become

$$g_{ik}\ddot{x}^k + \frac{\partial g_{ik}}{\partial x^l}\,\dot{x}^l\dot{x}^k = \frac{1}{2}\frac{\partial g_{lk}}{\partial x^i}\,\dot{x}^l\dot{x}^k + F_i$$

which we can rewrite as

$$g_{ik}\ddot{x}^k + \frac{1}{2}\left[\frac{\partial g_{ik}}{\partial x^l} + \frac{\partial g_{il}}{\partial x^k} - \frac{\partial g_{lk}}{\partial x^i}\right]\dot{x}^l\dot{x}^k = F_i$$

or multiplying by g^{ik},

$$\ddot{x}^i + \left\{ {i \atop l\ k} \right\} \dot{x}^l\dot{x}^k = F^i \tag{2.13}$$

This shows that the Lagrange equations in the general case are second-order differential equations in which Christoffel symbols occur explicitly. We see that, in the case of force-free motion, the generalized velocity vector $\dot{x}^i$ is displaced parallel to itself along the trajectory $x^i(t)$. The physico-geometric significance of such vector transplantation becomes evident.

2.3 Geodesics in Affine and Riemann Space

Let us look for a definition of a straight line in an affine space. In Euclidean space we can characterize such a line by the property that an arbitrary tangent vector along it remains parallel to itself when displaced along the curve. With the help of the law of vector transplantation introduced in Sec. 2.1, we can use the same characteristic property to define a *generalized straight line* in an affine space; we shall call such a curve a *geodesic.* Consider a curve $x^i(q)$ parametrized by q; if $\xi^i(q)$ is an arbitrary tangent vector to the curve, the equation

$$\frac{d\xi^i}{dq} - \Gamma^i_{\alpha\beta}\frac{dx^\alpha}{dq}\,\xi^\beta = 0$$

expresses the fact that it is displaced "unchanged" along the curve. A particular tangent vector is dx^i/dq; a more general tangent vector is therefore $\lambda(q)(dx^i/dq)$, where $\lambda(q)$ is an arbitrary function of q. The equations which define a geodesic in an affine space become

$$\frac{d}{dq}\left(\lambda(q)\,\frac{dx^i}{dq}\right) = \Gamma^i_{\alpha\beta}\,\frac{dx^\alpha}{dq}\,\lambda(q)\,\frac{dx^\beta}{dq} \tag{2.14}$$

This contains the apparently arbitrary function $\lambda(q)$. However, these equations can be brought into a standard form by the following simple transformation: we multiply both sides by $\lambda(q)$ and introduce a new parameter $p(q)$ defined by $dp = dq/[\lambda(q)]$, in terms of which the above equations take the so-called "normal form"

$$\frac{d^2x^i}{dp^2} - \Gamma^i_{\alpha\beta}\,\frac{dx^\alpha}{dp}\,\frac{dx^\beta}{dp} = 0 \tag{2.15}$$

From the way in which we obtained Eqs. (2.15) we see that, starting from an arbitrary parameter q, we can always find a parameter $p(q)$ in terms of which the equations defining geodesic lines in affine space take the "normal form" (2.15). We therefore take these equations to be a *definition of geodesic lines in an affine space.*

One should note that the "normal form" is not parameter-independent and is only valid when one uses a particular class of parameters p; indeed, if we change p into another parameter $\pi(p)$, Eqs. (2.15) become, using $\dfrac{dx^i}{dp} = \dfrac{dx^i}{d\pi}\,\dfrac{d\pi}{dp}$,

$$\frac{dx^i}{d\pi}\,\frac{d^2\pi}{dp^2} + \left[\frac{d^2x^i}{d\pi^2} - \Gamma^i_{\alpha\beta}\,\frac{dx^\alpha}{d\pi}\,\frac{dx^\beta}{d\pi}\right]\left(\frac{d\pi}{dp}\right)^2 = 0$$

This is still in the "normal form" only if

$$\frac{dx^i}{d\pi}\,\frac{d^2\pi}{dp^2} = 0$$

Since the tangent vector $dx^i/d\pi$ can always be taken to be nonzero, the above condition becomes $d^2\pi/dp^2 = 0$; the parameters π and p must therefore be *proportional*. This proportionality shows that the parameter p of the normal form is determined up to a constant factor; thus we have, even in a nonmetric affine space, a sort of pseudo-length.

Equations (2.15) are a system of ordinary second-order differential equations. Consequently, the solution of the initial-value problem for the geodesics defined by (2.15) [$x^\alpha(0)$ and $dx^\alpha(0)/dp$ are given at the origin, $p = 0$] is unique. Geometrically, this means that, through a given point and with a given tangent at that point, one and only one geodesic line can be drawn.

In the more particular case of a Riemann space we shall use the law of parallel displacement to define geodesics instead of the general law of vector transplantation; the tangent vector which we displace along the curve must in this case be of constant length along the curve. The only such vector is the unit tangent vector dx^i/ds (or a vector proportional to it), where s is the arc length of the geodesic; for this particular tangent vector

$$g_{ik}\,\frac{dx^i}{ds}\,\frac{dx^k}{ds} = \left(\frac{ds}{ds}\right)^2 = 1$$

Furthermore, the affine connections can be expressed explicitly in terms of Christoffel symbols, and the defining equations for a geodesic in a Riemann space become

$$\frac{d^2x^i}{ds^2} + \left\{ {i \atop \alpha\;\beta} \right\} \frac{dx^\alpha}{ds}\,\frac{dx^\beta}{ds} = 0 \tag{2.16}$$

Other definitions of geodesics in a Riemann space. A geodesic can also be defined by the requirement that it be the shortest curve between a point P_0 and another point P_1. The arc length of a curve between a point P_0 (characterized by the value $p = p_0$ of the curve parameter) and a point P_1 (at which $p = p_1$) is

$$s = \int_{P_0}^{P_1} \left(g_{ik}\,\frac{dx^i}{dp}\,\frac{dx^k}{dp}\right)^{1/2} dp \tag{2.17}$$

As an alternative definition of a geodesic, we require that the variation of this integral be zero. [Note that we must still show consistency with (2.15).] Thus we characterize a geodesic by the requirement

$$\delta \int_{P_0}^{P_1} \left(g_{ik}\,\frac{dx^i}{dp}\,\frac{dx^k}{dp}\right)^{1/2} dp = 0 \tag{2.18}$$

The condition (2.18) expresses a necessary condition for the shortest path between two given points in space (in fact, the mathematical requirement for a stationary path). We might expect this definition of a geodesic to be parameter-independent; in fact, p may be any parameter describing the curve. The change of p into q merely introduces the derivative dq/dp which cancels out of (2.18) because of the homogeneity in dp; therefore the form of Eq. (2.18) remains unaltered.

The definition (2.18) of a geodesic involves the square root of the purely geometric quantity $T = \frac{1}{2}g_{ik}\frac{dx^i}{dp}\frac{dx^k}{dp}$. It has the great theoretical advantage of yielding a parameter-independent integral, but the presence of the square root leads to cumbersome calculations in most applications. It is therefore of practical value to achieve greater flexibility and to show that a more general variational problem leads to precisely the same extremal curves. Let $F(T)$ be an *arbitrary* monotonic and differentiable function of its argument T, and consider the variational problem

$$\delta \int_{P_0}^{P_1} F(T)\, dp = 0 \tag{2.19}$$

This problem depends, of course, upon the particular choice of the parameter p. We shall assume that p is the arc-length parameter s of the extremal curve. More precisely, we deal with the following variational problem: A curve $x^i(s)$ is parametrized by its arc-length parameter s and is compared with all nearby curves $\tilde{x}^i(s)$ which coincide with it at the endpoints $P_0, x^i(0)$ and $P_1, x^i(l)$. What is the condition that, with this parameter $p = s$, the original curve is stationary in this family of competing curves?

Since $g_{ik} = g_{ik}(x^i)$ and $T = \frac{1}{2}g_{ik}\dot{x}^i\dot{x}^k$ (where $\dot{x}^i \equiv dx^i/dp$), Eq. (2.19) can be written with clearly displayed arguments:

$$\delta \int_{P_0}^{P_1} F[T(x^i, \dot{x}^i)]\, dp = 0$$

This is a typical problem of the calculus of variation, which leads to the Euler-Lagrange equations

$$\frac{d}{dp}\left(\frac{\partial F}{\partial \dot{x}^i}\right) = \frac{\partial F}{\partial x^i} \tag{2.20}$$

These may be written, using the functional form of T, as

$$\frac{d}{dp}\left[F'(T)g_{ik}\frac{dx^k}{dp}\right] = F'(T)\frac{1}{2}\frac{\partial g_{lk}}{\partial x^i}\frac{dx^l}{dp}\frac{dx^k}{dp}$$

If we make the particular choice of parameter, p = arc length s of the extremal line, we always have $T = \frac{1}{2}g_{ik}\dot{x}^i\dot{x}^k = \frac{1}{2}(ds/ds)^2 = \frac{1}{2}$ along the solution curve of the variational problem. Therefore $F(T)$ and $F'(T)$ are constant along the extremal curve. Thus we can take $F'(T)$ out of the left-hand-side bracket in these equations and still obtain equations defining the extremal curves of the problem:

$$\frac{d}{ds}\left(g_{ik}\frac{dx^k}{ds}\right) = \frac{1}{2}\frac{\partial g_{lk}}{\partial x^i}\frac{dx^l}{ds}\frac{dx^k}{ds}$$

By definition of the Christoffel symbols, and by the same calculations which led from (2.12) to (2.13), we obtain

$$\frac{d^2x^i}{ds^2} + \left\{\begin{matrix} i \\ l\ k \end{matrix}\right\}\frac{dx^l}{ds}\frac{dx^k}{ds} = 0 \tag{2.21}$$

This is identical with the original definition of a geodesic line in a Riemann space (2.16).

We have shown that the variational problem (2.19) and the differential equation (2.16) are consistent alternative definitions of a geodesic line when the curve parameter is the extremal arc length s. The special case $F(T) = \sqrt{T}$ in (2.18) is, however, parameter-independent, as we noted. So (2.18) is consistent with the differential equations (2.16), with no restriction on the choice of curve parameter.

In conclusion, one should notice that, among the three definitions of geodesics which we introduced, the two based on a variational principle, (2.18) and (2.19), are particularly well adapted to the spirit of general relativity because they do not require any particular specification of a coordinate system; therefore they fit immediately into the coordinate-invariant formulation of physics at which one aims in the theory of general relativity.

Connection with dynamics

1. Equations (2.13),

$$\ddot{x}^i + \left\{\begin{matrix} i \\ k\ l \end{matrix}\right\}\dot{x}^k\dot{x}^l = F^i$$

which we obtained as the normal form of the Lagrange equations, are the classical equations of motion of a particle in configuration space. If there are no external forces, $F^i = 0$, they become identical with (2.16) and define *geodesic lines* in configuration space. But in (2.13) we know that the parameter with respect to which the differentiation is performed must be the time parameter t along the trajectory. On the other hand, in (2.16) the parameter must be the arc length s along the geodesic in configuration space. Therefore, since both are normal parameters, the arc length s must be proportional to the time parameter t: $ds \propto dt$. One should note that this proportionality holds only on the geodesic lines;

along other curves the notions of time of transit and of arc length are different.

The fact that the vector $\dot{x}^i$ does not change its length under parallel displacement can be expressed by the equation

$$T = \tfrac{1}{2} g_{ik} \dot{x}^i \dot{x}^k = \text{const}$$

and is nothing but the principle of conservation of energy. Since

$$s = \int \sqrt{g_{ik} \dot{x}^i \dot{x}^k}\, dt$$

we see that

$$s = \sqrt{2T}\, t$$

which determines the above factor of proportionality.

2. With no external forces present the Lagrangian L and the kinetic energy T are equal, $L = T$. Taking the function $F(T)$ (defined above) as T itself, the trajectories of free particles, which are geodesics in configuration space, satisfy the variational problem

$$\delta \int_{P_0}^{P_1} T\, ds = 0 \tag{2.22}$$

Along a geodesic this is the same as

$$\delta \int_{t_0}^{t_1} T\, dt = 0$$

In this last form we recognize Hamilton's principle. The advantage of Hamilton's principle over Eq. (2.18) in defining geodesic lines lies in the fact that the integrand in Hamilton's principle is always defined whereas the presence of the square root in Eq. (2.18) does not allow us to define geodesics when $T \leq 0$. Since the two definitions are equivalent whenever (2.18) has a meaning, we shall adopt Hamilton's principle as a more general definition of geodesics; this allows in particular the definition of *null* geodesic lines, which will be of importance in Chap. 6, when we consider the trajectories of light rays.

It was noticed by Gauss that, for every given force-free dynamical trajectory, one can find a particular coordinate system in which the equations of motion (2.13) reduce to the Newtonian form $\ddot{x}^i = 0$. The proof is immediate, as follows.

We can choose a coordinate system in which the geodesic line described is the x^1 line, i.e., the trajectory is the curve $x^i = 0$ for $i \neq 1$. Furthermore, we may measure x^1 as the arc length along the trajectory. Thus $\dot{x}^1 = \text{const}$ since we know that arc length and time are proportional along the trajectory. Inserting these values into (2.13) we see that $\left\{ {i \atop 1\ 1} \right\} = 0$ along the entire trajectory and that $\ddot{x}^i = 0$ for all components. It is particularly interesting that the Christoffel symbols $\left\{ {i \atop 1\ 1} \right\}$ can be reduced to zero in the large along the entire trajectory. We shall utilize this fact later to introduce a so-called Gaussian coordinate system in which a normal form of the metric tensor is obtained in the large.

Lagrange's equations in the light of general relativity theory. Suppose one forgets about the physical origin of the generalized coordinates x^i and sees the equations of motion written in the form

$$\ddot{x}^i + \left\{ {i \atop k\ l} \right\} \dot{x}^k \dot{x}^l = F^i \tag{2.23}$$

When thinking in terms of Newtonian mechanics, one would like to see these equations take the form $\ddot{x}^i = \tilde{F}^i$, where $\tilde{F}^i$ represents external forces according to Newton's law. To be able to make this identification one considers the quantities

$$-\left\{ {i \atop k\ l} \right\} \dot{x}^k \dot{x}^l$$

as representing fictitious forces (such as centrifugal and Coriolis forces); these visibly depend on the coordinate system used (through the Christoffel symbols) and are often said to appear because one uses the "wrong kind" of coordinate system, for instance, a system attached to a rotating disk. A "right kind" of coordinate system is of course one in which these fictitious forces simply do not appear. However, one can use the alternative approach of treating all forces equally, be they external, fictitious, or due to a constraint, and accordingly write

$$\ddot{x}^i = \tilde{F}^i \qquad \tilde{F}^i = F^i - \left\{ {i \atop k\ l} \right\} \dot{x}^k \dot{x}^l \tag{2.24}$$

Clearly, the combined force depends very much on the coordinate system.

Obviously, such a viewpoint is not in the spirit of general relativity theory where all kinds of coordinate systems are considered equivalent. From the viewpoint of general relativity, one would instead like to reduce the equations of motion as much as possible to the *geometry* of the configuration space. That is, instead of explaining away wrong geometries by fictitious forces, we should like to explain away forces by proper choices of geometry. This will be possible at least in the case of gravitational forces. The easiest way to do this is to postulate that the gravitational forces F^i can be made to disappear from the above equations of motion by incorporating them into the geometric term $\left\{ {i \atop k\ l} \right\} \dot{x}^k\dot{x}^l$ just like a fictitious force. This approach is motivated by the fact that gravitational and fictitious forces both act on material bodies in the same way; they communicate an acceleration $\ddot{x}^i$ which is *independent of the body's mass.* (This is not the case for other types of forces; for instance, the acceleration communicated to a body by a spring is inversely proportional to the mass of the body.) The above property is the basis of the *principle of equivalence,* which states that the effect of a gravitational field can be "reproduced" by describing physics in an appropriately accelerated frame of reference without interior gravitational forces present. In such a frame of reference the generalized coordinates will be some y^j (which can be considered functions of x^i and t), and the kinetic energy of the system will be described by a new function $\bar{T}$. In order to bring in the principle of equivalence and incorporate all gravitational forces in the geometric term, one would like the Lagrange equations in the moving frame (with coordinates y^j and kinetic energy function $\bar{T}$) to take the form

$$\ddot{y}^j = - \left\{ {j \atop k\ l} \right\} \dot{y}^k\dot{y}^l \tag{2.25}$$

which are the equations of configuration-space geodesics in the moving frame.

Unfortunately, we can easily show that such an attempt to incorporate the principle of equivalence cannot succeed within the framework of classical mechanics: consider the concrete case of a particle moving under the influence of gravity along a three-dimensional trajectory described by $y^j(t)$ ($j = 1, 2, 3$) in a moving frame of reference. If Eqs. (2.25) were valid, the acceleration $\ddot{y}^j$ of the particle in that frame of reference would depend quadratically on the velocity $\dot{y}^j$ of the particle; doubling the velocity of a particle submitted to a gravitational field would therefore quadruple its acceleration. We know from experience that the movement of a particle in a gravitational field does not obey such a law in any frame of reference. Equations (2.25) are therefore unacceptable to describe the effects of gravitational forces, and *it is impossible to have gravitational forces take the same mathematical form as fictitious forces within the framework of classical analytical mechanics.*

We shall be able to formulate the principle of equivalence in mathematical terms only when we consider *Euler-Lagrange equations in a four-dimensional space* which includes time as an ordinary coordinate; in order to consider the solution of gravitational problems by a purely geometrical treatment, it will be necessary to make use of the concepts of special relativity theory and the Lorentz metric. In fact, in a four-dimensional space for which the zeroth coordinate is ct (time multiplied by the speed of light), Eqs. (2.25) are acceptable. When the velocities involved in the problem are small compared with the speed of light, we have

$$\dot{y}^0 = \frac{d}{dt}(ct) = c$$

$$\dot{y}^k \ll c \qquad k = 1, 2, 3$$

Thus Eqs. (2.25) reduce in lowest order to

$$\ddot{y}^j = - \left\{ {j \atop 0\ 0} \right\} c^2 \qquad j = 1, 2, 3 \tag{2.26}$$

The terms quadratic in velocity, which prevented us from making any progress in a three-dimensional framework, do not appear in these lower-order equations in a four-dimensional framework. The only term which survives is the constant c^2. Furthermore, we see that the Christoffel symbols (which are geometric entities) here play the role of forces. These considerations will be taken up in greater detail in Chap. 4.

2.4 Gaussian Coordinates

By letting a family of geodesics play a particular role among the coordinate lines, Gauss has shown the existence of a very useful coordinate system, which we shall now describe. Let us restrict ourselves to the case of a four-dimensional space with a hyperbolic metric, which we defined earlier as a metric of signature $(1,-1,-1,-1)$. Consider a three-dimensional hypersurface S imbedded in this four-dimensional space; we suppose that any vector n normal to S satisfies the inequality

$$n^0n_0 + (n^1n_1 + n^2n_2 + n^3n_3) > 0 \tag{2.27}$$

which, in the familiar language of special relativity theory, implies that the surface is "oriented in space" (whereas a vector normal to S is "oriented in time").

We introduce in the surface S three coordinates x^{*1}, x^{*2}, x^{*3} which serve to characterize the variable point $P^* \varepsilon S$. Through each point

P^* of the three-dimensional surface S we draw the geodesic which is orthogonal to S at P^*. These geodesics will form a field of nonintersecting curves in some finite neighborhood M of S such that, through each point P of M, there will pass exactly one of the geodesics constructed. We introduce now, in the entire four-dimensional domain M, coordinates as follows: Given P, we consider the geodesic of the field passing through P and its original point $P^* \varepsilon S$. We define the coordinates x^i of P in

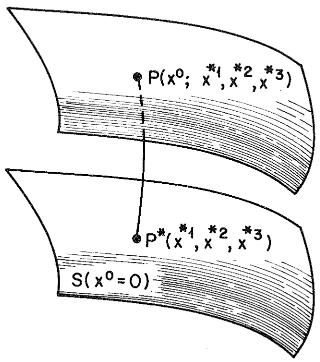

Fig. 2.1

terms of the arc length P^*P of the geodesic and of the coordinates x^{*i} of P^*:

$$x^0 = \text{arc length } P^*P \text{ along the geodesic}$$

$$x^1 = x^{*1}$$

$$x^2 = x^{*2}$$

$$x^3 = x^{*3}$$

In this manner, the three coordinates x^1, x^2, x^3 remain constant along any geodesic perpendicular to S; it follows that, along such a geodesic,

$$ds^2 = (dx^0)^2 \qquad g_{00} = 1$$

Next let us express the conditions resulting from the orthogonality of the x^0 (geodesic) lines to the hypersurface S; any vector $(0,a,b,c)$ in S must be orthogonal to the vector $(1,0,0,0)$ tangent to the x^0 line at the

same point, which requires that

$$g_{01} = g_{02} = g_{03} = 0$$

on the hypersurface S. Thus we see that, on S, the line element has the form

$$(dx^0)^2 + g_{ik}\, dx^i\, dx^k$$

Let us now attempt to show that the above form of the line element also holds outside the hypersurface S in the coordinate system we have constructed above. From the equations of a geodesic line (2.16),

$$\frac{d^2x^\alpha}{ds^2} + \begin{Bmatrix} \alpha \\ \mu \ \nu \end{Bmatrix} \frac{dx^\mu}{ds}\frac{ds^\nu}{ds} = 0$$

applied to the present case, for which x^0 alone varies along a geodesic and $ds = dx^0$, we deduce

$$\begin{Bmatrix} i \\ 0 \ 0 \end{Bmatrix} = 0 \qquad i = 1, 2, 3$$

and therefore

$$[00,i] = 0$$

Explicitly, writing out this last condition, we obtain

$$2\frac{\partial g_{0i}}{\partial x^0} - \frac{\partial g_{00}}{\partial x^i} = 0$$

Since $g_{00} = 1$ on *any* geodesic, the second term in the above is identically zero, so we have

$$\frac{\partial g_{0i}}{dx^0} = 0$$

along the x^0 lines. Therefore the elements g_{01}, g_{02}, g_{03} of the metric tensor, which are zero on S, remain zero in the entire domain M determined by the field of geodesics. In such a domain, the metric of the four-dimensional space expressed in Gaussian coordinates takes everywhere the form

$$ds^2 = (dx^0)^2 + g_{ij}(x^0,x^1,x^2.x^3)\, dx^i\, dx^j \qquad i, j = 1, 2, 3 \tag{2.28}$$

Any hypersurface (x_0 = const) is therefore orthogonal to the geodesic x^0 coordinate lines.

Furthermore, the lengths of the segments of any two geodesic lines between hypersurfaces S_1 and S_2 are clearly equal; the lengths of these segments correspond to the unique time interval Δx^0 between S_1 and S_2. Thus the surfaces $x^0 = \text{const}$ are equidistant level surfaces in terms of our geodesics; they correspond to parallel planes in Euclidean geometry.

The Gaussian coordinates introduced here will be very useful in later chapters because of the separation which they perform between a time coordinate x^0 valid everywhere in three-space and the metrical description of three-space itself. They allow us to connect the abstract four-dimensional framework of general relativity theory with the classical intuitive point of view which regards events as occurring in three-space and describes the development of these events in terms of a universal time parameter. (In Sec. 8.3 there is further discussion of the relation of Gaussian coordinates to the Einstein equations.)

Exercises

2.1 How many algebraically independent Christoffel symbols are there in two, three, and four dimensions? In n dimensions?

2.2 What are the Christoffel symbols in two-dimensional Euclidean space with orthogonal axes? What if the axes are canted, as mentioned in Sec. 1.9?

2.3 What are the Christoffel symbols for the surface of a unit sphere? (Use the results of Exercise 1.6.) Show that the equator and longitude lines on the surface of the earth are geodesics but that latitude lines are not.

2.4 Consider, in spherical coordinates r, θ, and φ, a diagonal metric with diagonal components $f(r)$, r^2, and $r^2 \sin^2 \theta$. Calculate the Christoffel symbols.

2.5 What are the geodesic equations for the metric of Exercise 2.4? Is the ray $\theta = \text{const}$, $\varphi = \text{const}$ a geodesic?

2.6 Consider a vector of small coordinate displacements $(\Delta r, \Delta\theta)$ attached to the point (r,θ) in polar coordinates. Parallel-displace it to a nearby point using the law (2.6). Check diagramatically that the displaced vector and the original vector are indeed parallel in the usual Euclidean sense.

2.7 Prove the theorem $g_{il\|k} = [ki,l] + [kl,i]$.

Problems

2.1 Use matrix theory to show that by a real coordinate transformation any metric can be brought into diagonal form at a given point. Moreover show that the diagonal elements can be made equal to $+1$, -1, or 0 and that the number of $+1$, -1, and 0 elements is independent of the manner in which the transformation is achieved. This diagonal matrix is called the *Cayley-Sylvester canonical form* of the matrix, and the set of diagonal elements is an alternative definition of the signature [see Eq. (1.2)]. It is always assumed to be $(+1,-1,-1,-1)$ in relativity theory, and the coordinate system defined by the transformation is termed a *tangent Lorentz space;* see Sec. 5.6 for more details.

2.2 The theorem in Sec. 2.1 guarantees that there exists a coordinate system in which the coefficients of affine connection vanish if they are symmetric in any system. Extend that theorem to show that the coordinate system may simultaneously be chosen to be a tangent Lorentz space by considering a transformation $\bar{x}^i = B^i_j x^j + \frac{1}{2} A^i_{jk} x^j x^k$, where the A^i_{jk} and B^i_j are suitably chosen constants.

Bibliography

Cartan, E. (1946): "Leçons sur la géométrie des espaces de Riemann," 2d ed., Paris
Christoffel, E. B. (1869): Über die Transformation der homogenen Differentialausdrücke zweiten Grades, "Journal für die reine und angewandte Mathematik (Crelle)," Vol. 70, pp. 46–70.
Einstein, A. (1955): "Meaning of Relativity," 5th ed., Princeton, N.J.
Eisenhart, L. P. (1949): "Riemannian Geometry," 2d ed., Princeton, N.J.
Kreyszig, E. (1959): "Differential Geometry," Toronto.
Lichnerowicz, A. (1955): "Théories relativistes de la gravitation et de l'électro-magnétisme," Paris.
Rainich, G. Y. (1950): "Mathematics of Relativity," New York–London.
Raschewski, P. K. (1959): "Riemannsche Geometrie und Tensoranalysis," Berlin. (Translation from the Russian, Moscow, 1953).
Schouten, J. A., and D. J. Struik (1935–1938): "Einführung in die neueren Methoden der Differentialgeometrie," 2 vols., Groningen.
Schrödinger, E. (1950): "Space-Time Structure," London.
Weatherburn, C. E. (1938): "Riemannian Geometry and the Tensor Calculus," London.
Weyl, H. (1950): "Space, Time, Matter," New York.
Whittaker, E. T. (1937): "Analytical Dynamics," 4th ed., London.
Willmore, T. J. (1959): "An Introduction to Differential Geometry," Oxford.

CHAPTER

3

Tensor Analysis

We return in this chapter to the problem of comparing tensors attached to neighboring points of space. To do this we shall first define intrinsically a differentiation process for a vector field.

3.1 Covariant Differentiation

Let us study how a contravariant vector field $\xi^i(x^j)$ varies when one goes from a point x^j to a neighboring point $x^j + dx^j$ in an affine space. To do this we shall compare at the point $x^j + dx^j$ the value $\xi^i(x^j + dx^j)$ of the vector field with the vector $\xi^{i*}(x^j + dx^j)$ obtained from $\xi^i(x^j)$ by the law of vector transplantation along the infinitesimal vector dx^j (Fig. 3.1). We thus form, at the point $x^j + dx^j$, the vector difference

$$(3.1) \qquad \xi^i(x^j + dx^j) - \xi^{i*}(x^j + dx^j)$$

By the use of a Taylor expansion, $\xi^i(x^j + dx^j)$ can be written as

$$\xi^i(x^j + dx^j) = \xi^i(x^j) + \frac{\partial \xi^i}{\partial x^k} dx^k + O(dx^k)^2$$

From the law of vector transplantation, the displaced vector $\xi^{i*}(x^j + dx^j)$ is equal to

$$\xi^{i*}(x^j + dx^j) = \xi^i(x^j) + \Gamma^i_{kl}\xi^l \, dx^k$$

Therefore the vector difference (3.1) becomes

$$(3.2) \quad \xi^i(x^j + dx^j) - \xi^{i*}(x^j + dx^j) = \left[\frac{\partial \xi^i}{\partial x^k} - \Gamma^i_{kl}\xi^l\right] dx^k + O(dx^k)^2$$

This expression has the form of the beginning of a series expansion and suggests interpreting the quantity

$$\left[\frac{\partial \xi^i}{\partial x^k} - \Gamma^i_{kl}\,\xi^l\right] \tag{3.3}$$

which is independent of the displacement dx^j, as a sort of "first derivative" of the field $\xi^i(x^j)$. Furthermore, formula (3.2) is valid for an *arbitrary* infinitesimal displacement dx. The left side of (3.2) is a vector (the difference of two vectors), and on the right side dx^k is an arbitrary vector. Thus, neglecting terms of higher order than the first in dx^k,

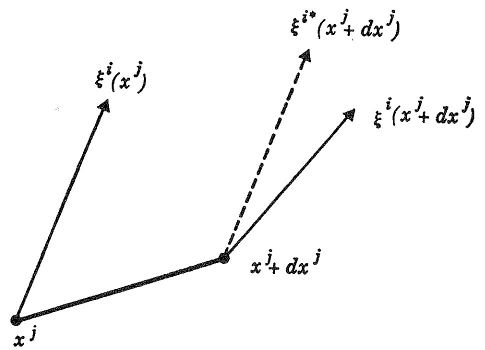

Fig. 3.1

we see by the quotient theorem that the quantity (3.3) is a tensor. We shall call it the *covariant derivative* of the contravariant vector field ξ^i.

One should notice that the derivatives $\partial\xi^i/\partial x^k$ taken with respect to arbitrary coordinates do not themselves form a tensor, as is evident from their transformation properties; neither do the coefficients of connection, as we already know. But the combination (3.3) is, as we have proved, a tensor. At this point it is advantageous for simplicity to introduce certain notation conventions. To indicate *ordinary differentiation* of a tensor component $T^\alpha{}_{\beta\gamma}$ with respect to x^k, we introduce the notation

$$T^\alpha{}_{\beta\gamma|k} = \frac{\partial T^\alpha{}_{\beta\gamma}}{\partial x^k}$$

The bar in front of the index k indicates that this particular index is an index of differentiation, while the others retain their usual meaning. To represent the *covariant derivative* of a vector, we introduce the notation

$$\xi^i_{\|k} = \frac{\partial \xi^i}{\partial x^k} - \Gamma^i_{kl}\xi^l$$

which can also be written in terms of the previous convention as

$$\xi^i_{\|k} = \xi^i_{|k} - \Gamma^i_{kl}\xi^l \tag{3.4}$$

The name covariant derivative given to this tensor is justified for the following reasons:

1. It appears already as playing the role of a first derivative in the Taylor series (3.2), where it first occurs.
2. It reduces to the usual derivative $\xi^i_{|k}$ in a geodesic coordinate system in which the connections Γ^i_{kl} vanish.

If $\xi^i_{\|k}$ were zero over a finite region of space, one could then build up a vector field by vector transplantation, starting with one given vector at one particular point (x^i). Then we should have

$$\xi^i(x^j + dx^j) = \xi^{i*}(x^j + dx^j)$$

over a finite region of space, and ξ^i would obey the law of vector transplantation. Such a vector field could then be called a generalized *constant vector field* since it would have a zero covariant derivative. However, one must be careful to note that the condition

$$\xi^i_{\|k} = \xi^i_{|k} - \Gamma^i_{kl}\xi^l = 0$$

can hold in a finite region of space *only if certain integrability conditions on the connections are satisfied.* These conditions will put restrictions on the type of spaces in which constant vector fields can be defined. Indeed, we shall show in Chap. 5 that a constant vector field can be defined *only* in a pseudo-Euclidean space. It is for this reason that we never attempted to compare vectors which are a finite distance apart by using the vector-transplantation law to carry one vector to the point of attachment of the other vector. We shall show in Chap. 5 that such a procedure depends in general on the path along which the transplantation is performed.

The concept of covariant derivative can be specialized in an obvious way to the case of Riemann spaces by replacing the concept of vector transplantation by that of parallel displacement and by expressing the affine connections in terms of Christoffel symbols. The covariant derivative of a vector field in a Riemann space is then defined by analogy with (3.4) to be

$$\xi^i_{\|k} = \xi^i_{|k} + \left\{ {i \atop k\ l} \right\} \xi^l \tag{3.5}$$

In the rest of this book we shall always deal with a Riemann space because one uses such a space in the theory of general relativity, but most concepts which will be introduced in this chapter can easily be generalized to an affine space.

We shall now give a direct proof that $\xi^i_{\|k}$ is a tensor, without recourse to series expansions. Given a vector field $\xi^i(x^\alpha)$ in a Riemann space, let us consider an arbitrary curve $x^\alpha(s)$ and define a vector field $\eta^i(s)$ at each point of the curve by the equations $\dfrac{d\eta^i}{ds} + \begin{Bmatrix} i \\ k\ l \end{Bmatrix} \dfrac{dx^l}{ds}\eta^k = 0$. These are ordinary first-order differential equations, and the $\eta^i(s)$ are defined along the curve once initial values have been chosen. The quantity

$$P(s) = g_{ik}\xi^i\eta^k$$

defined for each point of the curve is clearly a scalar; therefore

$$\frac{dP(s)}{ds} = P'(s)$$

is also a scalar. Let us write it out:

$$P'(s) = g_{ik|l}\frac{dx^l}{ds}\xi^i\eta^k + g_{ik}\xi^i_{|l}\frac{dx^l}{ds}\eta^k + g_{ik}\,\xi^i\eta^k_{|l}\frac{dx^l}{ds}$$

Using the defining equation for η^i and relabeling the dummy indices, we obtain

$$P'(s) = g_{ir|l}\,\xi^i\frac{dx^l}{ds}\eta^r + g_{ir}\xi^i_{|l}\frac{dx^l}{ds}\eta^r - g_{ik}\,\xi^i\begin{Bmatrix} k \\ r\ l \end{Bmatrix}\frac{dx^l}{ds}\eta^r$$

or

$$P'(s) = \left[\xi^i\left(g_{ir|l} - g_{ik}\begin{Bmatrix} k \\ r\ l \end{Bmatrix}\right) + g_{ir}\xi^i_{|l}\right]\frac{dx^l}{ds}\eta^r = T_{rl}\frac{dx^l}{ds}\eta^r$$

where T_{rl} represents the bracket. The left-hand side of this equation is a scalar, and dx^l/ds and η^r are arbitrary vectors at x^i; therefore, by the quotient theorem, the quantity T_{rl} *is a tensor*. The form of T_{rl} can be simplified using only the definitions of the Christoffel symbols of the first and second kind, (2.8) and (2.9). One obtains immediately

$$g_{ir|l} - g_{ik}\begin{Bmatrix} k \\ r\ l \end{Bmatrix} = g_{ir|l} - [rl,i] = [il,r] \tag{3.6}$$

Thus we may write T_{rl} as

$$T_{rl} = \xi^i[il,r] + g_{ir}\xi^i_{|l}$$

Multiplication and contraction with g^{rs} then gives the result

$$T^s{}_l = \xi^s_{|l} + \begin{Bmatrix} s \\ i\ l \end{Bmatrix}\xi^i$$

Since T_{rl} is a tensor, $T^s{}_l$ is also a tensor. Thus we have proved in a formal way that $\xi^s_{\|l} = T^s{}_l$ is indeed a tensor. However, our first derivation displays more clearly the significance of the term as a vector derivative under parallel displacement.

Covariant derivative of a covariant vector field. In the foregoing paragraphs we defined the covariant derivative of a contravariant vector field. Let us define an analogous operation for a covariant vector field $\eta_i(x^\alpha)$. We shall impose on the covariant derivative the defining conditions that it be a tensor and reduce to the ordinary derivative in a geodesic coordinate system, which we defined in Chap. 2. Given the covariant vector field $\eta_i(x^\alpha)$, consider an arbitrary contravariant vector field $\xi^i(x^\alpha)$ and form the scalar $\phi(x^\alpha) = \xi^i(x^\alpha)\eta_i(x^\alpha)$. We can write the gradient of the scalar ϕ, which we know is a covariant vector, as

$$v_l = \phi_{|l} = (\xi^i\eta_i)_{|l} = \xi^i_{|l}\eta_i + \xi^i\eta_{i|l}$$

Note that v_l is a *vector field*, which has been created from two other vector fields without the use of Christoffel symbols. By definition, the covariant derivative of ξ^i is

$$\xi^i_{\|l} = \xi^i_{|l} + \begin{Bmatrix} i \\ m\ l \end{Bmatrix}\xi^m$$

Using this tensor, let us form the covariant vector $\eta_i\xi^i_{\|l} = w_l$ and consider the vector difference $s_l = v_l - w_l$. We now have the vector

$$s_l = v_l - w_l = \eta_{i|l}\xi^i - \begin{Bmatrix} i \\ m\ l \end{Bmatrix}\xi^m\eta_i$$

By changing the first dummy index i into m, we may write this as

$$s_l = \xi^m\left(\eta_{m|l} - \begin{Bmatrix} r \\ m\ l \end{Bmatrix}\eta_r\right)$$

Since ξ^m is an arbitrary contravariant vector and s_l is (as we have constructed it) a covariant vector, we may use the quotient theorem to deduce that $\eta_{m|l} - \begin{Bmatrix} r \\ m\ l \end{Bmatrix} \eta_r$ is a tensor. Furthermore, since it clearly reduces to the ordinary derivative of η_m in a geodesic coordinate system, we may call it the covariant derivative of the covariant vector η_i. In analogy to the covariant derivative notation of (3.5) for a contravariant vector, we shall denote it by

$$\eta_{m\|l} = \eta_{m|l} - \begin{Bmatrix} r \\ m\ l \end{Bmatrix} \eta_r \tag{3.7}$$

Differentiation of a vector product and differentiation of tensors. Consider two arbitrary vector fields ξ^i and η^k and the tensor $\xi^i\eta^k$. We shall define the covariant derivative of this tensor $(\xi^i\eta^k)_{\|l}$ to meet the following requirement: In a geodesic coordinate system this quantity is to be identical with $(\xi^i\eta^k)_{|l} = \xi^i\eta^k_{|l} + \xi^i_{|l}\eta^k$. We know that, in a geodesic coordinate system, $\xi^i_{|l} = \xi^i_{\|l}$, and therefore

$$(\xi^i\eta^k)_{\|l} = \xi^i\eta^k_{\|l} + \xi^i_{\|l}\eta^k$$

But in this last form the right side of the expression is written as a tensor. We can therefore take it as a consistent definition of $(\xi^i\eta^k)_{\|l}$ in an arbitrary coordinate system.

We can evidently extend the previous reasoning to any tensor of the form $\xi^i\eta^k\zeta_m$ and obtain

$$(\xi^i\eta^k\zeta_m)_{\|l} = \xi^i\eta^k\zeta_{m\|l} + \xi^i\eta^k_{\|l}\zeta_m + \xi^i_{\|l}\eta^k\zeta_m \tag{3.8}$$

Replacing the covariant derivatives of the vectors by their values

$$\xi^i_{\|l} = \xi^i_{|l} + \begin{Bmatrix} i \\ l\ r \end{Bmatrix} \xi^r$$

etc., we obtain

$$(\xi^i\eta^k\zeta_m)_{\|l} = (\xi^i\eta^k\zeta_m)_{|l} + \begin{Bmatrix} i \\ l\ r \end{Bmatrix} \xi^r\eta^k\zeta_m + \begin{Bmatrix} k \\ l\ r \end{Bmatrix} \xi^i\eta^r\zeta_m - \begin{Bmatrix} r \\ l\ m \end{Bmatrix} \xi^i\eta^k\zeta_r \tag{3.9}$$

Note that, whenever a summation occurs over an index of a covariant vector, the summation index appears in the upper position in the Christoffel symbol; summations involving a contravariant vector have the summation index in the lower position in the Christoffel symbol.

Since we know from Sec. 1.5 that any tensor $T^{\alpha\beta}{}_\gamma$ can be written in the form of a sum of multinomials $T^{\alpha\beta}{}_\gamma = \Sigma\xi^\alpha\eta^\beta\zeta_\gamma$, we have also found the rule of differentiation of any tensor. A typical case is

$$T^{\alpha\beta}{}_{\gamma\|l} = T^{\alpha\beta}{}_{\gamma|l} + \begin{Bmatrix} \alpha \\ \tau\ l \end{Bmatrix} T^{\tau\beta}{}_\gamma + \begin{Bmatrix} \beta \\ \tau\ l \end{Bmatrix} T^{\alpha\tau}{}_\gamma - \begin{Bmatrix} \tau \\ \gamma\ l \end{Bmatrix} T^{\alpha\beta}{}_\tau$$

The generalization to any number of indices is evident; one need only be careful to balance indices and remember to use a plus sign for each contravariant index and a minus sign for each covariant index in the general case.

In the remainder of this book, we shall often follow the usage of physicists and speak of the covariant derivative of a tensor. It would be more precise but more verbose to speak of a tensor field.

Other approach to tensor covariant differentiation formula. Instead of deducing the law of differentiation for tensors from the law known for vectors, we could proceed directly as follows: If $T^\alpha{}_{\beta\gamma}$ is a tensor and ξ_α, η^β, ζ^γ are three arbitrary vectors, consider the scalar $T^\alpha{}_{\beta\gamma}\xi_\alpha\eta^\beta\zeta^\gamma$. Its gradient $(T^\alpha{}_{\beta\gamma}\xi_\alpha\eta^\beta\zeta^\gamma)_{|l}$ is a covariant vector which we shall call w_l:

$$w_l = T^\alpha{}_{\beta\gamma|l}\xi_\alpha\eta^\beta\zeta^\gamma + T^\alpha{}_{\beta\gamma}\xi_{\alpha|l}\eta^\beta\zeta^\gamma + T^\alpha{}_{\beta\gamma}\xi_\alpha\eta^\beta_{|l}\zeta^\gamma + T^\alpha{}_{\beta\gamma}\xi_\alpha\eta^\beta\zeta^\gamma_{|l}$$

Next consider the vector v_l:

$$v_l = T^\alpha{}_{\beta\gamma}(\xi_\alpha\eta^\beta\zeta^\gamma)_{\|l}$$

which can be written by (3.8)

$$v_l = T^\alpha{}_{\beta\gamma}\xi_{\alpha\|l}\eta^\beta\zeta^\gamma + T^\alpha{}_{\beta\gamma}\xi_\alpha\eta^\beta_{\|l}\zeta^\gamma + T^\alpha{}_{\beta\gamma}\xi_\alpha\eta^\beta\zeta^\gamma_{\|l}$$

Let us form the vector difference

$$w_l - v_l = T^\alpha{}_{\beta\gamma|l}\xi_\alpha\eta^\beta\zeta^\gamma + T^\alpha{}_{\beta\gamma}(\xi_{\alpha|l} - \xi_{\alpha\|l})\eta^\beta\zeta^\gamma + T^\alpha{}_{\beta\gamma}\xi_\alpha(\eta^\beta_{|l} - \eta^\beta_{\|l})\zeta^\gamma + T^\alpha{}_{\beta\gamma}\xi_\alpha\eta^\beta(\zeta^\gamma_{|l} - \zeta^\gamma_{\|l})$$

We know from (3.7) that

$$\xi_{\alpha|l} - \xi_{\alpha\|l} = \begin{Bmatrix} r \\ \alpha\ l \end{Bmatrix} \xi_r$$

and similarly for the differences $\eta^\beta{}_{|l} - \eta^\beta{}_{\|l}$ and $\zeta^\gamma{}_{|l} - \zeta^\gamma{}_{\|l}$. Substitution of these in the above expression for $w_l - v_l$ gives

$$w_l - v_l = T^\alpha{}_{\beta\gamma|l}\xi_\alpha\eta^\beta\zeta^\gamma + T^\alpha{}_{\beta\gamma}\begin{Bmatrix} r \\ \alpha\ l \end{Bmatrix}\eta^\beta\zeta^\gamma\xi_r - T^\alpha{}_{\beta\gamma}\begin{Bmatrix} \beta \\ l\ r \end{Bmatrix}\xi_\alpha\eta^r\zeta^\gamma$$
$$- T^\alpha{}_{\beta\gamma}\begin{Bmatrix} \gamma \\ l\ r \end{Bmatrix}\xi_\alpha\eta^\beta\zeta^r$$

Relabeling the dummy indices, we have

$$w_l - v_l = \left[T^\alpha{}_{\beta\gamma|l} + \begin{Bmatrix} \alpha \\ l\ s \end{Bmatrix} T^s{}_{\beta\gamma} - \begin{Bmatrix} s \\ l\ \beta \end{Bmatrix} T^\alpha{}_{s\gamma} - \begin{Bmatrix} s \\ l\ \gamma \end{Bmatrix} T^\alpha{}_{\beta s}\right]\xi_\alpha\eta^\beta\zeta^\gamma$$

Since ξ_α, η^β, ζ^γ are arbitrary vectors and $w_l - v_l$ is a vector, we may use the quotient theorem to infer that the quantity in brackets is a tensor. It reduces to $T^\alpha{}_{\beta\gamma|l}$ in a geodesic coordinate system. We therefore call it the covariant derivative of the tensor; as expected, this definition coincides with (3.9) obtained by our previous method.

$$T^\alpha{}_{\beta\gamma\|l} = T^\alpha{}_{\beta\gamma|l} + \begin{Bmatrix} \alpha \\ l\ s \end{Bmatrix} T^s{}_{\beta\gamma} - \begin{Bmatrix} s \\ l\ \beta \end{Bmatrix} T^\alpha{}_{s\gamma} - \begin{Bmatrix} s \\ l\ \gamma \end{Bmatrix} T^\alpha{}_{\beta s} \tag{3.10}$$

Properties of covariant differentiation. Let us now consider the derivative of the product of two tensors,

$$(T^\alpha{}_{\beta\gamma}S^\epsilon{}_{\sigma\nu})_{\|l}$$

We may ask whether the usual formula for differentiation of a product remains valid:

$$(T^\alpha{}_{\beta\gamma}S^\epsilon{}_{\sigma\nu})_{\|l} = T^\alpha{}_{\beta\gamma\|l}S^\epsilon{}_{\sigma\nu} + T^\alpha{}_{\beta\gamma}S^\epsilon{}_{\sigma\nu\|l}$$

The validity of this formula could be inferred directly from formula (3.8). However, it can be shown to be correct more directly by noting that the two sides of the equation are tensors. We know that they are equal in one particular coordinate system—the geodesic system in which covariant differentiation is equivalent to ordinary differentiation. Thus the equation is a tensor equality and is true in all systems.

We can now prove an important theorem due to Ricci: *The covariant derivative of the g_{ik} tensor is identically zero.* To prove this we first make use of (3.6) to obtain the ordinary derivative

$$g_{ir|l} = [il,r] + g_{ik}\begin{Bmatrix} k \\ r\ l \end{Bmatrix}$$

Thus the covariant derivative is

$$g_{ir\|l} = [il,r] + g_{ik}\begin{Bmatrix} k \\ r\ l \end{Bmatrix} - \begin{Bmatrix} k \\ r\ l \end{Bmatrix} g_{ik} - \begin{Bmatrix} k \\ i\ l \end{Bmatrix} g_{kr}$$
$$= [il,r] - g_{kr}\begin{Bmatrix} k \\ i\ l \end{Bmatrix}$$

Using the definition of the Christoffel symbol $\begin{Bmatrix} k \\ i\ l \end{Bmatrix}$ in (2.9), we see that this is identically zero. Thus the metric tensor which characterizes the Riemann space is a constant in the absolute sense; i.e., it has a zero covariant derivative. From this property it is obvious that the operations of raising or lowering indices commute with covariant differentiation. For instance,

$$(\xi_i)_{\|l} = (g_{ik}\xi^k)_{\|l} = g_{ik\|l}\xi^k + g_{ik}\xi^k{}_{\|l} = g_{ik}(\xi^k{}_{\|l})$$

since $g_{ik\|l} = 0$.

We conclude this section with the important remark that the operation of covariant differentiation does not possess in general the commutative property of ordinary differentiation.

$$\xi^\alpha{}_{\|\mu\|\nu} \neq \xi^\alpha{}_{\|\nu\|\mu} \qquad \text{but} \qquad \xi^\alpha{}_{|\mu|\nu} = \xi^\alpha{}_{|\nu|\mu}$$

We can verify this by direct computation or by noticing that, if we try to use the method of the geodesic coordinate system, we shall come across derivatives of Christoffel symbols which will neither vanish nor cancel out. Only in special spaces will the commutative property hold; this subject will be examined further in Chap. 5.

3.2 Applications of Tensor Analysis

Divergence of a vector. The quantity $\xi^i{}_{\|l}$ is a tensor, so $\xi^i{}_{\|i}$ is a scalar, being the contraction of a tensor; we call it the *divergence* of the vector ξ^i.

$$\xi^i{}_{\|i} = \text{div}\ \xi = \xi^i{}_{|i} + \begin{Bmatrix} i \\ i\ s \end{Bmatrix}\xi^s$$

We shall show that this expression can be put into a simpler form which allows very easy computation and no longer contains Christoffel symbols.

Indeed, the Ricci theorem proved above gives

$$g_{ij|h} - \left\{ {k \atop h\ i} \right\} g_{kj} - \left\{ {k \atop h\ j} \right\} g_{ik} = 0$$

Multiplication and contraction with g^{ij} leads to

$$g^{ij}g_{ij|h} - \left\{ {i \atop h\ i} \right\} - \left\{ {j \atop h\ j} \right\} = 0$$

and therefore, since j is a dummy index,

$$\left\{ {i \atop i\ h} \right\} = \tfrac{1}{2}g^{ij}g_{ij|h}$$

We know that g^{ik} are the elements of the inverse matrix of g_{ik}, and therefore $g^{ik} = \Delta^{ik}/g$, where

$$g = \det\,((g_{ik}))$$

and Δ^{ik} is the cofactor of g_{ik} as defined in the theory of determinants. But we could expand the determinant g along one of its rows, say the third row, into $g = g_{3k}\,\Delta^{3k}$; from this we see that $\partial g/\partial g_{ik} = \Delta^{ik}$, since Δ^{ik} is a subdeterminant of g which does not contain the variable g_{ik}. Replacing Δ^{ik} by this expression, we have

$$g^{ik} = \frac{1}{g}\frac{\partial g}{\partial g_{ik}}$$

and

(3.11) $$\left\{ {i \atop i\ h} \right\} = \frac{1}{2}\frac{1}{g}\frac{\partial g}{\partial g_{ik}}\frac{\partial g_{ik}}{\partial x^h} = \frac{1}{2g}\frac{\partial g}{\partial x^h} = \frac{1}{2}\frac{\partial}{\partial x^h}\log|g|$$
$$= \frac{\partial}{\partial x^h}\log\sqrt{-g}$$

a result which is important in its own right.

This last way of writing a contracted Christoffel symbol of the second kind is legitimate only if the quantity $-g$ under the square-root sign is positive. We shall assume usually that the metric considered has a g_{ik} tensor of signature identical with the one of special relativity $(+---)$. This assumption is usually made in general relativity in the form that one can find, at every point considered, a local coordinate

system in which the g_{ik} tensor reduces to the Lorentz metric tensor of special relativity.

Putting the above results into the divergence formula, we obtain

$$\xi^i{}_{\|i} = \operatorname{div}\xi = \frac{1}{\sqrt{-g}}\left[\xi^i{}_{|i}\sqrt{-g} + \xi^s\frac{\partial\sqrt{-g}}{\partial x^s}\right]$$

that is,

(3.12) $$\xi^i{}_{\|i} = \operatorname{div}\xi = \frac{1}{\sqrt{-g}}(\xi^i\sqrt{-g})_{|i}$$

We shall see in the next section that the divergence formula is specially suited to formulate Gauss's integral theorem in a Riemann space.

Note that the divergence operation has been defined only on the contravariant form of a vector; this is no restriction, however, since one can always operate with the contravariant components of a given vector.

Geometric interpretation of the divergence formula. Let us consider a change of coordinates from x^i to $\bar{x}^i$. The metric tensor g_{ik} transforms according to $\bar{g}_{ik} = \dfrac{\partial x^\alpha}{\partial \bar{x}^i}\dfrac{\partial x^\beta}{\partial \bar{x}^k}g_{\alpha\beta}$. Therefore, using the rules of matrix multiplication, we see that the determinant g transforms according to

$$\bar{g} = g\left(\det\frac{\partial x^\gamma}{\partial \bar{x}^i}\right)^2$$

and hence, since we always choose the positive value of the square root,

$$\sqrt{-\bar{g}} = \sqrt{-g}\left|\frac{\partial(x^1, x^2, \ldots)}{\partial(\bar{x}^1, \bar{x}^2, \ldots)}\right|$$

where the last factor is the inverse of the absolute value of the Jacobian of the coordinate transformation. On the other hand, one knows that the differential volume element, say, in four dimensions,

$$d\tau = dx^0\,dx^1\,dx^2\,dx^3$$

transforms according to

$$d\bar{\tau} = \left|\frac{\partial(\bar{x}^1, \bar{x}^2, \ldots)}{\partial(x^1, x^2, \ldots)}\right| d\tau$$

Thus

$$\sqrt{-\bar{g}}\, d\bar{\tau} = \sqrt{-g}\, d\tau$$

and therefore $\sqrt{-g}\, d\tau$ is a scalar. We shall call it the invariant four-dimensional volume element. As mentioned before, we assume that we can find always at each point of the Riemann space a coordinate system such that the g_{ik} tensor takes the form of the usual Lorentz metric tensor of special relativity

$$\begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

In such a coordinate system the invariant four-dimensional volume element $\sqrt{-g}\, d\tau$ becomes $d\tau = dx_0\, dv$, the "natural" volume element in that local system of reference in which a physicist measures lengths with rods and time with clocks.

Let us now consider the integral

$$\int_D (\mathrm{div}\, \xi)(\sqrt{-g}\, d\tau) = \int_D \xi^i{}_{\|i} \sqrt{-g}\, d\tau$$

over an n-dimensional domain D, where we have explicitly displayed the invariant volume element introduced above. This integral is obviously an invariant scalar quantity. With the help of the divergence formula (3.12) we can write it in the form

$$\int_D (\xi^i \sqrt{-g})_{|i}\, d\tau$$

which becomes, by direct application of Green's formula (integration by parts),

$$(3.13)\qquad \int_D (\xi^i \sqrt{-g})_{|i}\, d\tau = \sum_i \int_{\text{boundary}} \xi^i \sqrt{-g}\, dx^1 \cdots dx^{i-1}\, dx^{i+1} \cdots dx^n$$

We interpret expression (3.13) as the flux of the vector ξ^i across the boundary of the domain D. The identity (3.13) is called *Gauss's integral theorem* for a Riemann space. We see that the integral of a divergence of a vector in a volume depends only on the values of the vector on the

boundary, just as in the Euclidean space of ordinary integral calculus. This formula will be very useful later to express conservation laws in the theory of general relativity.

Laplacian of a scalar field. Let us consider the divergence of the gradient of a scalar function $W(x^\alpha)$. We know that this operation leads to the Laplacian in ordinary calculus. The gradient of $W(x^\alpha)$ is a covariant vector $W(x^\alpha)_{|k}$. We form its contravariant components by multiplying by g^{ik} and take its divergence:

$$(3.14)\qquad \frac{1}{\sqrt{-g}} (\sqrt{-g}\, g^{ik} W_{|k})_{|i} = \nabla^2 W = (g^{ik} W_{|k})_{\|i}$$

In analogy with the Euclidean case, we call this scalar the Laplacian of W.

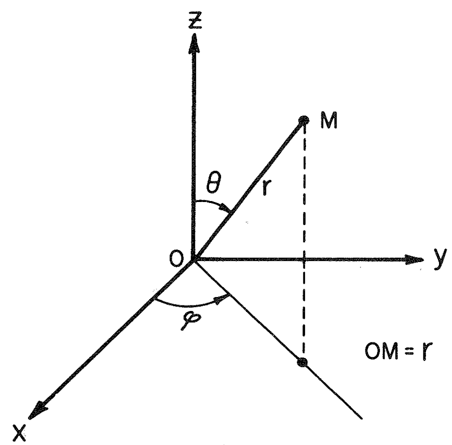

Fig. 3.2

The present formalism can be used to obtain very easily the expression for the Laplacian in any curvilinear coordinate system in ordinary analysis. Let us take as example the case of polar coordinates in *three-dimensional* Euclidean space: r, θ, φ, as indicated in Fig. 3.2. The Euclidean metric in these coordinates is

$$ds^2 = dr^2 + r^2\, d\theta^2 + r^2 \sin^2\theta\, d\varphi^2$$

Thus g_{ik} is diagonal, and

$$g_{11} = 1 \qquad g_{22} = r^2 \qquad g_{33} = r^2 \sin^2\theta$$

Therefore $g = r^4 \sin^2\theta$ and

$$g^{11} = 1 \qquad g^{22} = \frac{1}{r^2} \qquad g^{33} = \frac{1}{r^2\sin^2\theta}$$

Here g is positive, and therefore we simply replace $\sqrt{-g}$ by $\sqrt{g}$ in (3.14) to obtain the expression for the Laplacian in the present case. The Laplacian of a function $W(r,\theta,\varphi)$ is therefore

$$\nabla^2 W = \frac{1}{\sqrt{g}}\left[\sqrt{g}\left(g^{i1}\frac{\partial W}{\partial r} + g^{i2}\frac{\partial W}{\partial \theta} + g^{i3}\frac{\partial W}{\partial \varphi}\right)\right]_{|i}$$

$$\nabla^2 W = \frac{1}{r^2\sin\theta}\left[\frac{\partial}{\partial r}\left(r^2\sin\theta\frac{\partial W}{\partial r}\right) + \frac{\partial}{\partial\theta}\left(r^2\sin\theta\frac{1}{r^2}\frac{\partial W}{\partial\theta}\right) + \frac{\partial}{\partial\varphi}\left(r^2\sin\theta\frac{1}{r^2\sin^2\theta}\frac{\partial W}{\partial\varphi}\right)\right]$$

which becomes, after simplifications,

$$\nabla^2 W = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial W}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial W}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 W}{\partial\varphi^2}$$

We thus obtain this well-known formula in a straightforward way by using the powerful coordinate invariant formalism of tensor analysis.

3.3 Symmetric and Antisymmetric Tensors

We introduce two classes of tensors which play an important role in physical theories. They are defined in the following way:

1. A totally symmetric tensor† is a tensor whose components remain unchanged under interchange of any two indices:

$$T_{ikl} = T_{kil} = T_{lik}$$

2. A totally antisymmetric tensor† is a tensor whose components change sign under any odd permutation of its indices. Clearly, it will therefore remain unchanged under an even permutation of indices:

$$T_{ikl} = -T_{kil} = T_{kli}$$

† Totally symmetric and totally antisymmetric tensors will be simply called symmetric and antisymmetric in further developments.

The simplest example of a symmetric tensor is the metric tensor g_{ik}. The most important antisymmetric tensor in physics is probably the well-known Maxwell tensor, which will be studied in detail in Chap. 4. A justification for singling out the symmetric or antisymmetric character of tensors rests on the following theorem.

Theorem. The symmetric and antisymmetric character of a tensor is an intrinsic property; that is, the symmetry characteristics do not depend upon the coordinate system used.

Proof. Given the tensor $T_{\alpha\beta}$ in one coordinate system x^i, consider another coordinate system $\bar{x}^i$ in which the tensor has components

$$\bar{T}_{ik} = \frac{\partial x^\alpha}{\partial \bar{x}^i}\frac{\partial x^\beta}{\partial \bar{x}^k} T_{\alpha\beta}$$

and

$$\bar{T}_{ki} = \frac{\partial x^\beta}{\partial \bar{x}^k}\frac{\partial x^\alpha}{\partial \bar{x}^i} T_{\beta\alpha}$$

We see that, if $T_{\alpha\beta} = T_{\beta\alpha}$, then $\bar{T}_{ik} = \bar{T}_{ki}$, and if $T_{\alpha\beta} = -T_{\beta\alpha}$, then $\bar{T}_{ik} = -\bar{T}_{ki}$, which proves the theorem for tensors of rank 2. The generalization to tensors of arbitrary rank is immediate.

Notice that, in the three-dimensional space, an antisymmetric tensor of rank 2 has three independent components with which a vector can be associated. Indeed, let ξ^i and η^i be two arbitrary vectors and form the tensor $\xi^i\eta^k - \xi^k\eta^i$. Its three independent components coincide with the components of the exterior or vector product of ξ^i and η^i. Thus, in many applications in three-dimensional space, antisymmetric tensors of this form are identified with vectors. There is, however, one simple way of distinguishing between genuine vectors and such apparent vectors. If all coordinate axes are reversed in direction, each vector ξ^i goes over into $-\xi^i$, while the antisymmetric tensor $\xi^i\eta^k - \xi^k\eta^i$ remains unchanged. This difference in transformation behavior led to the distinction in classical vector theory between polar and axial vectors.

In a four-dimensional space a symmetric tensor of the second rank has 10 independent components, whereas an antisymmetric tensor of the second rank has 6 independent components and is sometimes called a six-vector. For example, the independent components of the Maxwell field tensor are the components of the electric and magnetic fields **E** and **H**. An antisymmetric tensor of the fourth rank has only one independent component; so has any antisymmetric tensor of rank equal to the dimension of the space considered. Antisymmetric tensors of rank higher than the space dimension are identically zero.

Curl of a vector. An antisymmetric tensor field can be obtained from a vector field in the following manner. Consider a covariant vector η_i and form the expressions

$$\eta_{i\|k} = \eta_{i|k} - \begin{Bmatrix} r \\ i\ k \end{Bmatrix} \eta_r$$

$$\eta_{k\|i} = \eta_{k|i} - \begin{Bmatrix} r \\ k\ i \end{Bmatrix} \eta_r$$

The Christoffel symbols are symmetric in their lower indices. Therefore the difference of the two tensors written above gives the tensor

$$\eta_{i\|k} - \eta_{k\|i} = \eta_{i|k} - \eta_{k|i}$$

which does *not* involve Christoffel symbols. This tensor is obviously antisymmetric; we call it the curl or rotation of the vector η_i. Note that the curl operation can be performed only on the covariant components of a vector, since, in the above, we made use of the symmetry of the Christoffel symbols in its two lower indices.

Homogeneous forms and symmetry character of tensors. The study of antisymmetric tensors is linked directly with the theory of exterior differential forms, which we shall introduce now. From the point of view of integration in an n-dimensional Riemann space, a differential volume element $dx^1\,dx^2 \cdots dx^n$ is not independent of the order of the factors. Indeed, $dx^1\,dx^2 \cdots dx^n$ is not equivalent to $dx^2\,dx^1 \cdots dx^n$, because going from one integration scheme to the other involves the change of coordinates

$$\bar{x}^1 = x^2$$

$$\bar{x}^2 = x^1$$

from which we get, by computing the Jacobian of the transformation,

$$dx^2\,dx^1 = d\bar{x}^1\,d\bar{x}^2 = \frac{\partial(\bar{x}^1,\bar{x}^2)}{\partial(x^1,x^2)}\,dx^1\,dx^2 = -dx^1\,dx^2$$

We see that the differential elements appear anticommutative from the point of view of integration; this simply reflects the change of orientation of the surface element $dx^1 \wedge dx^2$ when one changes the order of integration. We introduce in the space of differential elements the concept of exterior product, which preserves the laws of algebra, except that commutativity is replaced by anticommutativity. The symbol for exterior multiplication is $\wedge$. We thus have

$$dx^i \wedge dx^k = -dx^k \wedge dx^i$$

which in particular leads to

$$dx^i \wedge dx^i = 0$$

This notion goes back to Grassmann (Grassmann, 1878) and is most familiar in the vector product of elementary vector algebra. As we saw above, the exterior product will play a major role in integration theory. However, at this point we shall utilize only the fact that a totally antisymmetric tensor can be best defined as the coefficient system of an exterior differential form. We define an exterior differential form as a form of differentials in which multiplication is understood as exterior multiplication; for example,

$$G = A_{ik}\,dx^i \wedge dx^k$$

is an exterior differential form of order 2. One knows that any usual form $L = B_{ikl}x^ix^kx^l$ can be written with totally symmetric coefficients, since if B_{ikl} is not totally symmetric in i, k, l, we replace it by

$$C_{ikl} = \tfrac{1}{6}(B_{ikl} + B_{kil} + B_{lki} + B_{ilk} + B_{lik} + B_{kli})$$

which gives the same value to L and is symmetric in i, k, l. The study of such forms L reduces, therefore, to the study of forms with symmetric coefficients. The analogous property for exterior differential forms is that one can always utilize coefficients which are antisymmetric in all indices.

Let us show this in the case of two indices. Consider a form

$$G = A_{ik}\,dx^i \wedge dx^k$$

where A_{ik} is arbitrary. Because of our freedom in the choice of dummy indices, we can write it as $G = \frac{1}{2}(A_{ik}\,dx^i \wedge dx^k + A_{ki}\,dx^k \wedge dx^i)$ by simply relabeling summation indices. But the above expression may be written

$$G = \tfrac{1}{2}(A_{ik} - A_{ki})\,dx^i \wedge dx^k$$

by virtue of the anticommutativity of dx^i and dx^k. Since $(A_{ik} - A_{ki})$ is obviously antisymmetric, this proves our original proposition.

By considering an exterior differential form built up from an arbitrary tensor A_{ikl}, we can construct a new tensor which is antisymmetric. Consider, for example, $F = A_{ikl}\,dx^i \wedge dx^k \wedge dx^l$. It can be written

$$F = \frac{1}{3!}(A_{ikl} - A_{kil} + A_{kli} - A_{ilk} + A_{lik} - A_{lki})\,dx^i \wedge dx^k \wedge dx^l$$

That is, we sum over all permutations of the indices, giving a plus sign to an even permutation and a minus sign to an odd permutation. We thus define

$$F = \{A_{ikl}\}\,dx^i \wedge dx^k \wedge dx^l$$

in which the notation $\{A_{ikl}\}$ designates the antisymmetrized sum in the above sense over all permutations of the indices i, k, l:

$$\{A_{ikl}\} = \frac{1}{3!}(A_{ikl} - A_{kil} + A_{kli} - A_{ilk} + A_{lik} - A_{lki}) \tag{3.15}$$

The extension of this definition to an arbitrary number of indices is obvious.

Remarks

1. The antisymmetrization operation $\{\ \}$ is linear: one verifies immediately from the above definition that

$$\{A_{ikl} + B_{ikl}\} = \{A_{ikl}\} + \{B_{ikl}\}$$

2. Any tensor of rank 2 can be uniquely decomposed into the sum of a symmetric and an antisymmetric tensor. For tensors of rank higher than 2, this is not possible; in these cases the sum of a symmetric and an antisymmetric tensor does not give enough independent components for the representation of an arbitrary tensor. However, an arbitrary tensor of rank larger than 2 can be decomposed into the sum of more than two tensors whose components have distinguished symmetry properties upon particular permutation of the indices.

3. As a particular case we include tensors of rank 1 as antisymmetric tensors. The antisymmetrization operation is then the identity by definition.

Creation of a new antisymmetric tensor of rank $n + 1$ from an antisymmetric tensor of rank n. In Sec. 3.3 we created an antisymmetric tensor from a vector; we shall here generalize this process. The remarkable feature of this procedure is that the new tensor does not contain Christoffel symbols. Let us first consider the case $q = 3$ for convenience of notation. Given an antisymmetric tensor A_{ikl}, consider its covariant derivative

$$A_{ikl\|m} = A_{ikl|m} - \left\{ {r \atop i\ m} \right\} A_{rkl} - \left\{ {r \atop k\ m} \right\} A_{irl} - \left\{ {r \atop l\ m} \right\} A_{ikr}$$

This tensor is in general not antisymmetric because of the distinguished role of the index m. We antisymmetrize it with the method sketched above, forming

$$A_{ikl\|m}\,dx^i \wedge dx^k \wedge dx^l \wedge dx^m = \{A_{ikl\|m}\}\,dx^i \wedge dx^k \wedge dx^l \wedge dx^m$$

The new antisymmetric tensor $\{A_{ikl\|m}\}$ does not depend upon Christoffel symbols. This results from the following theorem.

Theorem. For every tensor field one has

$$\{A_{ikl\|m}\} = \{A_{ikl|m}\}$$

Proof. By definition, we can write

$$\{A_{ikl\|m}\}\,dx^i \wedge dx^k \wedge dx^l \wedge dx^m = \left[A_{ikl|m} - \left\{ {r \atop i\ m} \right\} A_{rkl} - \left\{ {r \atop k\ m} \right\} A_{irl} - \left\{ {r \atop l\ m} \right\} A_{ikr}\right] dx^i \wedge dx^k \wedge dx^l \wedge dx^m$$

Since each Christoffel symbol is symmetric in its lower indices whereas the differential elements anticommute, this last expression is equal to $A_{ikl|m}\,dx^i \wedge dx^k \wedge dx^l \wedge dx^m$. But, by definition,

$$A_{ikl|m}\,dx^i \wedge dx^k \wedge dx^l \wedge dx^m = \{A_{ikl|m}\}\,dx^i \wedge dx^k \wedge dx^l \wedge dx^m$$

Therefore

$$\{A_{ikl\|m}\} = \{A_{ikl|m}\} \tag{3.16}$$

Note that this process of creation of tensors by differentiation and antisymmetrization always has to start with completely covariant tensors.

By repeating the above procedure, one might think that one can keep creating new antisymmetric tensors of higher ranks. This is not the case, and the above procedure leads to a tensor identically zero at the second step. We illustrate this property in the following examples.

1. Let us begin with a tensor of rank zero, that is, a scalar W. Take its covariant derivative, and do not antisymmetrize since there is only one index present; one obtains the gradient of W, $W_{|k}$. Now repeat the previous operation: Take the covariant derivative of the gradient, antisymmetrize to get the tensor $\{W_{|k|l}\} = \frac{1}{2}(W_{|k|l} - W_{|l|k})$, which is visibly identically zero. We have found the well-known fact that the curl of a gradient vanishes. Thus observe the resultant sequence

$$W; \quad W_{|k}; \quad \{W_{|k|l}\} = 0$$

2. Consider the case of a tensor of rank $q = 1$, that is, the covariant vector ξ_i. Taking its covariant derivative and antisymmetrizing, we obtain the tensor

$$\{\xi_{i\|k}\} = \{\xi_{i|k}\} = \tfrac{1}{2}(\xi_{i|k} - \xi_{k|i})$$

which is proportional to the curl of the vector ξ_i. We repeat the previous operation. Taking the covariant derivative of the curl and antisymmetrizing, we get, in view of the linearity of the antisymmetrization operation,

$$\{\{\xi_{i|k}\}_{|l}\} = \tfrac{1}{2}\{(\xi_{i|k} - \xi_{k|i})_{|l}\} = \tfrac{1}{2}[\{\xi_{i|k|l}\} - \{\xi_{k|i|l}\}]$$

But we note that $\{\xi_{i|k|l}\} = -\{\xi_{k|i|l}\}$, since this antisymmetrized tensor is antisymmetric in i and k. Therefore

$$\{\{\xi_{i|k}\}_{|l}\} = \{\xi_{i|k|l}\}$$

On the other hand, since the order of differentiation is immaterial, we have $\xi_{i|k|l} = \xi_{i|l|k}$, and thus $\{\xi_{i|k|l}\} = \{\xi_{i|l|k}\}$. Since a cyclic permutation leaves the antisymmetrized expression invariant, we obtain also

$$\{\xi_{i|k|l}\} = \{\xi_{k|l|i}\}$$

On the other hand, the antisymmetry in i and k of $\{\xi_{i|k|l}\}$ implies

$$\{\xi_{i|k|l}\} = -\{\xi_{k|i|l}\}$$

Comparing the last two equations, we find

$$\{\xi_{i|k|l}\} = \{\{\xi_{i|k}\}_{|l}\} = 0 \tag{3.17}$$

Again the second antisymmetric differentiation leads to a null tensor.

3. In the case of a second-rank antisymmetric tensor t_{ik}, we consider $t_{ik|l}$ and antisymmetrize. As one can verify immediately, we obtain

$$\{t_{ik|l}\} = \tfrac{1}{3}(t_{ik|l} + t_{kl|i} + t_{li|k}) \tag{3.18}$$

This is an antisymmetric tensor of rank 3 which visibly does not contain any Christoffel symbols. The operation of antisymmetric differentiation on an antisymmetric tensor of rank 2 is analogous to the curl operation on a vector. (In the next chapter we shall see that the Maxwell field tensor satisfies the equations $\{t_{ik|l}\} = 0$, which are equivalent to some of Maxwell's equations involving a curl operation.)

If we now try to repeat the previous operation, we obtain $\{\{t_{ik|l}\}_{|m}\} = 0$. This result is general. In fact, our demonstration for the second example can be extended as it stands to any antisymmetric tensor $A_{jm\cdots npi}$ by forming $\{A_{jm\cdots npi|k}\}$ and proving that

$$\{\{A_{jm\cdots npi|k}\}_{|l}\} = 0 \tag{3.19}$$

since one sees that the indices $jm \cdots np$ are not relevant to the proof. This shows that the number of antisymmetric tensors which can be created by the above process is limited. We can, in fact, create only one such tensor from a given antisymmetric tensor of any rank; repeating the operation gives a null tensor.

3.4 Closed and Exact Tensors

It is a well-known property in ordinary vector analysis that a vector field whose curl is zero is derivable from a potential, and vice versa. We shall introduce here a similar notion into tensor analysis in a Riemann space.

We make the following *definitions:*

1. An antisymmetric tensor t_{ikl} is called a *closed* antisymmetric tensor if

$$\{t_{ikl|n}\} = 0$$

2. An antisymmetric tensor t_{ikl} is called an *exact* antisymmetric tensor if there exists a tensor T_{ik} such that

$$t_{ikl} = \{T_{ik|l}\}$$

T_{ik} is called the tensor potential of t_{ikl}.

Analogously, closed and exact tensors can be defined for every rank. In view of the foregoing definition, the result of the previous section can be formulated as a theorem.

Theorem. *Every exact tensor is closed.*

Indeed, take an antisymmetric tensor T_{ik} as potential and consider $\{T_{ik|l}\}$, which is by definition exact; we know that $\{\{T_{ik|l}\}_{|m}\} = 0$, which proves that $\{T_{ik|l}\}$ is closed. We shall show in the following theorem that the converse proposition is also true.

Theorem. *Every closed tensor is exact,* that is, admits a tensor potential.

We first illustrate this theorem with the familiar case of a vector in Euclidean geometry. If a vector t_i has zero curl, $t_{i|k} - t_{k|i} = 0$, we know that it is derivable from a potential ϕ:

$$t_i = \frac{\partial \phi}{\partial x^i} \tag{3.20}$$

The curl condition represents simply the integrability conditions of (3.20) and allows the quantity

$$d\phi = \frac{\partial \phi}{\partial x^i}\, dx^i = t_i\, dx^i$$

to be an exact differential.

We now prove the above theorem for tensors of rank 2 in an n-dimensional space. The proof will be by induction on the dimension n. We start with the case $n = 2$ and consider an antisymmetric tensor

$$t_{ik} = \begin{pmatrix} 0 & t_{12} \\ -t_{12} & 0 \end{pmatrix}$$

Let us form $\{t_{ik|l}\}$; it is identically zero because i, k, l can take only the values 1 and 2, and therefore there is always one index, 1 or 2, which is present twice. It follows that the result of the antisymmetrization is zero. This is an example of an antisymmetric tensor of rank higher than the space dimension and which must therefore be identically zero, as mentioned earlier. In two dimensions a second-rank antisymmetric tensor is always closed. Let us show that such a closed tensor is exact. We must find out if two functions t_1 and t_2 exist such that

$$t_{12} = \frac{\partial t_1}{\partial x^2} - \frac{\partial t_2}{\partial x^1}$$

This is obviously so, for one can take $t_2 = 0$ and $t_1 = \int t_{12}\, dx^2$. Therefore every second-rank antisymmetric tensor in two dimensions is exact. Before extending this proof to higher dimensions, let us first make the following remark: If we suppose that t_{12} depends on a parameter p and is k times differentiable in p, then, from the known laws of differentiation under the integral sign, the tensor potential which we defined above depends on p with the same differentiability properties.

To complete the induction, we go from dimension n to $n + 1$: Suppose that the indices i, k take the values $1, 2, \ldots, n + 1$ and that t_{ik} is a closed tensor in an $(n + 1)$-dimensional space:

$$\{t_{ik|l}\} = 0 \qquad \text{for} \qquad i, k, l = 1, 2, \ldots, n + 1$$

By our induction hypothesis we can now assert that the matrix t_{ik} considered as a function of x^1 to x^n, and for every fixed x^{n+1}, has a potential t_i such that

$$t_{ik} = t_{i|k} - t_{k|i} \qquad \text{for} \qquad i, k = 1, 2, \ldots, n$$

By the induction hypothesis again, these t_i depend on the parameter x^{n+1} in a differentiable manner. Therefore there exist functions $t_i(x^1, x^2, \ldots, x^n; x^{n+1})$ such that

$$t_{ik} = t_{i|k} - t_{k|i} \qquad \text{for} \qquad i, k = 1, 2, \ldots, n \tag{3.21}$$

To complete the induction we need only prove that we can write

$$t_{ik} = t_{i|k} - t_{k|i}$$

for i or k equal to $n + 1$.

The closure condition on t_{ik} for $l = n + 1$ and $i, k = 1, \ldots, n$ is $\{t_{ik|n+1}\} = 0$. It can be written out as

$$\frac{\partial t_{ik}}{\partial x^{n+1}} + \frac{\partial t_{k,n+1}}{\partial x^i} + \frac{\partial t_{n+1,i}}{\partial x^k} = 0$$

Using the potential representation (3.21) of t_{ik} for $i, k \leq n$, we obtain

$$\frac{\partial t_{i|k}}{\partial x^{n+1}} - \frac{\partial t_{k|i}}{\partial x^{n+1}} + \frac{\partial t_{k,n+1}}{\partial x^i} + \frac{\partial t_{n+1,i}}{\partial x^k} = 0$$

Taking into account the definition of $t_{i|k} = \partial t_i/\partial x^k$ and the fact that t_{ik} is antisymmetric, we arrive at

$$\frac{\partial t_{i|n+1}}{\partial x^k} - \frac{\partial t_{k|n+1}}{\partial x^i} - \frac{\partial t_{n+1,k}}{\partial x^i} + \frac{\partial t_{n+1,i}}{\partial x^k} = 0$$

which can be written

$$\frac{\partial}{\partial x^k}(t_{i|n+1} + t_{n+1,i}) = \frac{\partial}{\partial x^i}(t_{k|n+1} + t_{n+1,k})$$

If we define $W_i = t_{i|n+1} + t_{n+1,i}$, we have $(\partial/\partial x^k)W_i - (\partial/\partial x^i)W_k = 0$, which means that the curl of W_i is 0. In the present case of a vector we know that this condition implies that W_i admits a potential which we call t_{n+1}. We therefore write

$$W_i = \frac{\partial t_{n+1}}{\partial x^i} = t_{n+1|i} \qquad i = 1, 2, \ldots, n$$

which with the definition of W_i gives

$$t_{n+1,i} = t_{n+1|i} - t_{i|n+1}$$

This is the expression we were looking for, analogous to (3.21), but for $k = n + 1$.

We have used a fixed coordinate system x^i and found in it $n + 1$ functions $t_k(x_1, \ldots, x_{n+1})$ such that

$$t_{ik} = t_{i|k} - t_{k|i} \qquad i = 1, \ldots, n + 1$$

We allow the t_k to transform as covariant vectors under a change of coordinates, so that this equation becomes a tensor identity and is therefore valid independently of the coordinate system.

We have thus completed the proof that every closed tensor of rank 2 is exact. This property and its converse are true for tensors of arbitrary rank, but we shall make use of it only in the case of a tensor of rank 2 in four-dimensional space when writing Maxwell's equations in tensor form in the next chapter.

3.5 Tensor Densities—Dual Tensors

When studying the properties of the divergence of a vector under integration, we introduced the expression $\sqrt{-g}$ in order to be able to form

invariant integrals. This expression is a special case of an entity which plays an important role in general relativity, the so-called tensor density. A tensor density (which is usually denoted by a script letter such as $\mathfrak{I}_{\alpha\beta}{}^{\gamma}$) depends on the coordinate system in such a way that, under change of variables $x^\alpha \to \bar{x}^\alpha$, it obeys the transformation law

$$\bar{\mathfrak{I}}_{\alpha\beta}{}^{\gamma} = \frac{\partial x^k}{\partial \bar{x}^\alpha}\frac{\partial x^l}{\partial \bar{x}^\beta}\frac{\partial \bar{x}^\gamma}{\partial x^m}\mathfrak{I}_{kl}{}^m \frac{\partial(x^1, x^2, \ldots)}{\partial(\bar{x}^1, \bar{x}^2, \ldots)} \tag{3.22}$$

In order to avoid sign difficulties we shall restrict ourselves to transformations with a positive Jacobian. Then, in the sense of this definition, $\sqrt{-g}$ may be called a scalar density. This name comes from the fact that, if one integrates such a quantity, forming, for instance, the integral $I = \int \sqrt{-g}\, d\tau$, the result is an invariant as we saw earlier; thus $\sqrt{-g}$ behaves like the physical density in space of the quantity I.

If $T_{\alpha\beta}{}^{\gamma}$ is a tensor, clearly

$$\mathfrak{I}_{\alpha\beta}{}^{\gamma} = T_{\alpha\beta}{}^{\gamma}\sqrt{-g} \tag{3.23}$$

is a tensor density. The transition from tensor to tensor density can always be performed by a correspondence (3.23), and hence the knowledge of $\sqrt{-g}$ leads to a complete understanding of all tensor densities.

As illustration, let us consider in four-dimensional space four arbitrary contravariant vectors $\xi^\alpha_{(j)}$, $j = 1, 2, 3, 4$, and form the determinant

$$D = \det(\xi^\alpha_{(j)}) = \epsilon_{\alpha\beta\gamma\delta}\xi^\alpha_{(1)}\xi^\beta_{(2)}\xi^\gamma_{(3)}\xi^\delta_{(4)} \tag{3.24}$$

D appears here as a multilinear form of the four vectors with the coefficient system

$$\epsilon_{\alpha\beta\gamma\delta} = \begin{cases} 0 & \text{if any two indices are equal} \\ \pm 1 & \text{according to } (\alpha\beta\gamma\delta)\text{, being an even or odd permutation of the numbers } (1,2,3,4) \end{cases}$$

The ϵ-system is well known from elementary determinant theory. To keep the symmetry between upper and lower indices in tensor calculus, we also introduce the symbol $\epsilon^{\alpha\beta\gamma\delta}$, which is exactly the same numerical array as $\epsilon_{\alpha\beta\gamma\delta}$ but allows us to keep Einstein's summation convention when dealing with covariant vector components.

In our geometry the ϵ-system provides a multilinear form of vectors with a very simple transformation behavior under change of coordinates.

Indeed, from

$$\bar{\xi}^{\alpha}_{(j)} = \frac{\partial \bar{x}^{\alpha}}{\partial x^{l}} \xi^{l}_{(j)}$$

and the multiplication rule for determinants, we find that

$$\bar{D} = \frac{\partial(\bar{x}^1, \bar{x}^2, \ldots)}{\partial(x^1, x^2, \ldots)} D$$

Thus D^{-1} is a scalar density and $D\sqrt{-g}$ is a proper scalar. Multiplying both sides of (3.24) by $\sqrt{-g}$ and applying the quotient theorem to $D\sqrt{-g}$ and the four arbitrary vectors $\xi_{(j)}$, we find that

$$\epsilon_{\alpha\beta\gamma\delta}\sqrt{-g} = e_{\alpha\beta\gamma\delta} \tag{3.25a}$$

is an antisymmetric covariant tensor of rank 4, the Levi-Civita tensor. As remarked earlier, in Sec. 3.3, in a four-dimensional space, **e** is the only such antisymmetric tensor (within a multiplicative factor). The covariant components of **e** can be obtained from similar reasoning applied to covariant vector components; they are, as an easy calculation shows,

$$e^{\alpha\beta\gamma\delta} = \frac{-1}{\sqrt{-g}} \epsilon^{\alpha\beta\gamma\delta} \tag{3.25b}$$

We see that $\epsilon^{\alpha\beta\gamma\delta}$ is a tensor density in the sense of (3.23).

While the tensor density $\epsilon^{\alpha\beta\gamma\delta}$ has components which are independent of the coordinate system (similar to the Kronecker tensor $g^{i}{}_{k} = \delta^{i}{}_{k}$), it should be observed that the tensors $e^{\alpha\beta\gamma\delta}$ and $e_{\alpha\beta\gamma\delta}$ both have zero covariant derivatives. Indeed,

$$\begin{aligned} e_{\alpha\beta\gamma\delta\|l} = \epsilon_{\alpha\beta\gamma\delta}(\sqrt{-g})_{|l} &- \begin{Bmatrix} r \\ \alpha\ l \end{Bmatrix} e_{r\beta\gamma\delta} \\ &- \begin{Bmatrix} r \\ \beta\ l \end{Bmatrix} e_{\alpha r\gamma\delta} - \begin{Bmatrix} r \\ \gamma\ l \end{Bmatrix} e_{\alpha\beta r\delta} - \begin{Bmatrix} r \\ \delta\ l \end{Bmatrix} e_{\alpha\beta\gamma r} \end{aligned}$$

However, the four indices in $e_{\alpha\beta\gamma\delta}$ must be different to give nonzero components. Hence, in each Christoffel symbol, r must be identical with the index which it replaces in $e_{\alpha\beta\gamma\delta}$. We find, therefore,

$$e_{\alpha\beta\gamma\delta\|l} = \epsilon_{\alpha\beta\gamma\delta}\left[(\sqrt{-g})_{|l} - \begin{Bmatrix} r \\ r\ l \end{Bmatrix}\sqrt{-g}\right]$$

which is identically zero by virtue of (3.11).

By means of the coordinate independent coefficient system $\epsilon_{\alpha\beta\gamma\delta}$, we can establish a very simple correspondence between antisymmetric tensor densities of rank 2 and antisymmetric tensors of rank 2. Let $\mathfrak{T}^{\alpha\beta}$ be an antisymmetric tensor density,

$$\mathfrak{T}^{\alpha\beta} = T^{\alpha\beta}\sqrt{-g}$$

We define

$$(^*T)_{\alpha\beta} = \tfrac{1}{2}\epsilon_{\alpha\beta\gamma\delta}\mathfrak{T}^{\gamma\delta} = \tfrac{1}{2}e_{\alpha\beta\gamma\delta}T^{\gamma\delta} \tag{3.26}$$

which is an antisymmetric covariant tensor of rank 2. We call $^*T_{\alpha\beta}$ the *dual tensor* of the tensor $T^{\gamma\beta}$. Clearly, the tensor component $^*T_{\alpha\beta}$ coincides with the tensor density component $\mathfrak{T}^{\gamma\delta}$, with complementary indices, where $(\alpha,\beta,\gamma,\delta)$ form an even permutation of (1,2,3,4). For example,

$$^*T_{12} = \mathfrak{T}^{34} \qquad ^*T_{14} = \mathfrak{T}^{23} \qquad ^*T_{13} = \mathfrak{T}^{42} = -\mathfrak{T}^{24} \tag{3.27}$$

This fact shows the validity of the inverse formula

$$\mathfrak{T}^{\alpha\beta} = \tfrac{1}{2}\epsilon^{\alpha\beta\gamma\delta}(^*T)_{\gamma\delta} \qquad T^{\alpha\beta} = -\tfrac{1}{2}e^{\alpha\beta\gamma\delta}(^*T)_{\gamma\delta} \tag{3.26'}$$

which leads back from the dual tensor to the original one.

The notions of tensor density and dual tensor will be used in the next chapter to write Maxwell's equations in a very condensed form in a four-dimensional space. Furthermore, we shall make use of a property of second-rank antisymmetric tensors, which we now prove.

Theorem. *The two following properties of a second-rank antisymmetric tensor in four-dimensional space are equivalent: being closed and having a zero-divergence dual tensor;* that is,

$$\{T_{\alpha\beta|\lambda}\} = \{T_{\alpha\beta\|\lambda}\} = 0 \quad \text{implies} \quad (^*T^{\mu\nu})_{\|\nu} = 0 \quad \text{and vice versa.}$$

To prove this theorem note that from the definition (3.26) and the fact that the tensor $e_{\alpha\beta\gamma\delta}$ has a zero covariant derivative, we may write

$$(^*T^{\mu\nu})_{\|\nu} = \tfrac{1}{2}e^{\mu\lambda\alpha\beta}T_{\alpha\beta\|\lambda} \tag{3.28}$$

The summation indices λ, α, β, are dummy indices, and we can permute them arbitrarily in the above identity. If we write out all permutations and add the resulting six equations, the symmetry properties of $e^{\mu\lambda\alpha\beta}$ allow us to write the result as

$$(^*T^{\mu\nu})_{\|\nu} = e^{\mu\lambda\alpha\beta}\{T_{\alpha\beta|\lambda}\} \tag{3.29}$$

where the covariant derivative has been replaced by the ordinary derivative as allowed by (3.16). The two expressions in (3.29) are zero or nonzero together, which proves the theorem. Our reasoning also shows that the covariant divergence of an antisymmetric tensor can be expressed without the use of Christoffel symbols, as is evident from (3.29).

Finally, it should be remembered that to form invariant scalar quantities by integration, one must integrate over scalar densities. The expression

$$\int \varphi \sqrt{-g}\, d\tau \tag{3.30}$$

is a scalar if φ is a scalar function. On the other hand, the integral of a tensor density,

$$\int \mathfrak{T}^{\alpha\beta}\, d\tau = \int T^{\alpha\beta} \sqrt{-g}\, d\tau \tag{3.31}$$

has no well-defined transformation properties since it is not attached to a single point in space, and therefore no transformation coefficients $\partial \bar{x}^\alpha/\partial x^\beta$ can be defined. Only in the limiting case, when the volume τ shrinks down to an *infinitesimal neighborhood* of a given point P, can we say that (3.31) has meaning, for then the transformation coefficients $\partial \bar{x}^\alpha/\partial x^\beta$ are definable at P, and

$$\int \bar{\mathfrak{T}}^{\alpha\beta}\, d\bar{\tau} = \left(\frac{\partial \bar{x}^\alpha}{\partial x^\gamma}\right)_P \left(\frac{\partial \bar{x}^\beta}{\partial x^\delta}\right)_P \int \mathfrak{T}^{\gamma\delta}\, d\tau$$

is the transformation law for the integral.

3.6 Vector Fields on Curves

We introduced in Sec. 3.1 the concept of covariant derivative $\xi^i_{\|k}$ for a vector field $\xi^i(x^j)$ and used it to create from a given vector field new tensor fields by covariant differentiation. The basic idea in this operation is the comparison of the local change of the component ξ^i due to the form of the function $\xi^i(x^j)$ with the corresponding change according to the law of vector transplantation. The difference of these changes measures the absolute variation of the vector field and gives rise to covariant expressions.

It is natural to apply the same method in the case of a vector field which is defined only on a curve. Let $x^i(s)$ be the parametric representation of a curve Λ, and consider a vector field $\xi^i(s)$ given as a function of the curve parameter. If we move from $x^i(s)$ to $x^i(s + \Delta s)$, the vector components will change by

$$\xi^i(s + \Delta s) = \xi^i(s) + \frac{d\xi^i}{ds}\Delta s + 0(\Delta s^2) \tag{3.32}$$

while the vector transplantation (3.2) of $\xi^i(s)$ along the curve would have led to the vector

$$\xi^{i*}(s + \Delta s) = \xi^i(s) + \Gamma^i_{kl}\xi^l(s)\frac{dx^k}{ds}\Delta s + 0(\Delta s^2) \tag{3.33}$$

Hence we define the absolute derivative of the vector field $\xi^i(s)$ along the curve Λ by the formula

$$\frac{D\xi^i(s)}{Ds} = \frac{d\xi^i}{ds} - \Gamma^i_{kl}\frac{dx^k}{ds}\xi^l \tag{3.34}$$

The generalization of this operation to tensor fields given only along a curve is obvious. If $\xi^i(s)$ happens to be the restriction to the curve $x^j(s)$ of a general vector field $\xi^i(x^j)$, we have

$$\frac{D\xi^i}{Ds} = \left(\frac{\partial \xi^i}{\partial x^k} - \Gamma^i_{kl}\xi^l\right)\frac{dx^k}{ds} = \xi^i_{\|k}\frac{dx^k}{ds} \tag{3.35}$$

and more generally for corresponding tensor fields,

$$\frac{D}{Ds}T^{ik}(x^l) = T^{ik}_{\|l}\frac{dx^l}{ds} \tag{3.35$'$}$$

The operation of absolute differentiation creates from vector fields along Λ new vector fields:

$$\dot{\xi}^i = \frac{D\xi^i}{Ds} \tag{3.36}$$

That this is indeed a vector can be verified from the transformation law

(2.5) for connections and from the definition (3.34). It is now easily seen that

$$\dot{\xi}_i = g_{ik}\dot{\xi}^k = \frac{D}{Ds}(g_{ik}\xi^k) = \frac{D}{Ds}\,\xi_i \tag{3.36'}$$

since the tensor g_{ik} is constant under absolute differentiation. We also verify easily the laws of product differentiation; for example,

$$\frac{D}{Ds}(U_iV^j) = \dot{U}_iV^j + U_i\dot{V}^j \tag{3.37}$$

The concept of absolute differentiation allows us to develop a differential geometry for curves in a Riemann space, which is quite analogous to the corresponding curve theory in Euclidean space. The most important vector field along a curve is its tangent vector field $t^i(s)$; it is by definition proportional to dx^i/ds. To be more specific, let us suppose that Λ is a timelike curve and that s is its arc-length parameter. We may then take t^i to be a unit vector

$$t^i = \frac{dx^i}{ds} \qquad t^it_i = 1 \tag{3.38}$$

Next we let

$$m^i = \dot{t}^i = \frac{Dt^i}{Ds} \tag{3.39}$$

From the second equation (3.38) we conclude by absolute differentiation

$$m^it_i = m_it^i = 0 \tag{3.40}$$

The vector m^i is orthogonal to the timelike tangent vector t^i and is therefore spacelike. We define a unit vector in the direction of m^i by

$$n^i = \frac{1}{\kappa}\,m^i \qquad n_in^i = -1$$

and can then write (3.39) in the form

$$\dot{t}^i = \kappa n^i \tag{3.41}$$

The unit vector n^i is the principal normal of Λ, and formula (3.41) is the generalization of Frenet's formula in classical curve theory. It relates the derivative of the tangent vector to the normal vector n^i by means of the principal curvature κ.

We shall not pursue the theory of curves any further, but shall use the two important vector fields $t^i(s)$ and $n^i(s)$ along a curve to define a law of vector transport along a curve which is very similar to the law of vector transplantation treated before. We define the tensor field along Λ by use of the quantities t^i, n^i, and κ defined in (3.41):

$$T^{ik}(s) = \kappa(s)\left[n^i(s)t^k(s) - t^i(s)n^k(s)\right] \tag{3.42}$$

and consider the differential equation defining the vector field $V^i(s)$,

$$\frac{DV^i}{Ds} = T^i{}_k(s)V^k \tag{3.43}$$

By use of (3.34) we can bring (3.43) into the form

$$\frac{dV^i}{ds} = \left[\Gamma^i_{kl}\frac{dx^k}{ds} + T^i{}_l(s)\right]V^l(s) \tag{3.44}$$

which shows that (3.43) is a first-order linear homogeneous differential equation for the unknown vector $V^i(s)$. If we prescribe V^i at one point of Λ, say, for $s = 0$, we can determine the vector field $V^i(s)$ along Λ uniquely by means of (3.44). We may thus conceive (3.43) as a law of vector transport along the curve Λ.

Observe now that the vector $t^i(s)$ satisfies the differential equation (3.43) identically. By substitution, using (3.42) and (3.38), (3.40), and (3.41), we obtain

$$\frac{Dt^i}{Ds} = \kappa\left[n^it_kt^k - t^in_kt^k\right] = \kappa n^i = \dot{t}^i \tag{3.45}$$

which verifies our assertion; that is, the tangent vector $t^i(s)$ is carried along Λ by the special transport law (3.43).

Next, let $V^i(s)$ and $W^i(s)$ be two vector fields on Λ which are transported by the same law (3.43). We easily find

$$\frac{D(V^iW_i)}{Ds} = T^i{}_kV^kW_i + T^i{}_kW^kV_i = T^{ik}(V_iW_k + W_iV_k) = 0 \tag{3.46}$$

because of the antisymmetry of the tensor $T^{ik}(s)$. Thus the transport law (3.43) preserves the scalar product, and hence the length and angles of all vectors so displaced. It is thus closely analogous to the displacement law for general vector fields discussed above.

However, in general, our new transport law (3.43) will be different from the law of parallel displacement valid for the entire Riemann space.

Indeed, the only curves whose tangent vectors are obtained by parallel displacement are the geodesics, while our new transport formula is so constructed that it carries the tangent vector of Λ into itself. On the other hand, the transport law (3.43) is much more specialized than the parallel displacement law and depends strongly on Λ. However, in many physical applications the transport formula (3.43) is of significant value. It may happen that a particular curve in space-time plays a distinguished role without being a geodesic; for example, the timelike world-line of an observer will usually not be a geodesic. It will then be a great convenience to introduce a coordinate system which moves with the observer, preserves all geometrical relations, and has the world-line of the observer as one coordinate axis. The observer will refer his observations to an orthogonal triad of axes in his laboratory which are all three orthogonal in four-space to his four-velocity, thus forming an orthogonal tetrad of reference; he can transport along with him this orthogonal tetrad using the transport law (3.43). It is called the Fermi-Walker transport law. For details and applications, we refer the reader to the bibliography (Fermi, 1922; Pirani, 1957; Synge, 1960; and Walker, 1932).

3.7 Intrinsic Symmetries and Killing Vectors

In general relativity we try to rid ourselves of accidental properties of coordinate frames and are led naturally to the study of tensors. However, many problems possess an intrinsic symmetry which may be difficult to recognize in an arbitrary coordinate system. For example, the problem of describing the gravitational field of a stationary spherical body, which we shall discuss in Chap. 6, possesses intrinsic spherical symmetry; this is evident when the problem is expressed in spherical coordinates but may become hidden if other coordinates are used. We wish to consider how such hidden symmetries can be discovered in an invariant manner when working in an arbitrary coordinate system.

We consider a metric tensor $g_{\alpha\beta}(x^\gamma)$ which admits a one-parameter group of continuous transformations

$$\bar{x}^\mu = \varphi^\mu(x^\gamma;\lambda) \tag{3.47}$$

where λ is the parameter of the group. Under all these changes of markers we assume the metric to be the same. In this sense we can say that the metric has a hidden symmetry. The transformation (3.47) may, for example, represent a rotation group (around a fixed axis since we restricted ourselves to a one-parameter group) in properly chosen coordinates; but in an arbitrary marker system x^γ its appearance must be as general as written down. Under any change of markers we have the transformation formula for the $g_{\mu\nu}$ tensor:

$$\bar{g}_{\mu\nu}(\bar{x}^\kappa)\,\frac{\partial\bar{x}^\mu}{\partial x^\alpha}\frac{\partial\bar{x}^\nu}{\partial x^\beta} = g_{\alpha\beta}(x^\gamma) \tag{3.48}$$

But since we demanded that the metric tensor remain unchanged under the transformations (3.47), we have the symmetry requirement

$$\bar{g}_{\mu\nu}(\bar{x}^\kappa) = g_{\mu\nu}(\bar{x}^\kappa) \tag{3.49}$$

That is, the functional dependence of the tensor $g_{\mu\nu}$ on the markers must be the same before and after the transformation. Thus (3.48) yields

$$g_{\mu\nu}(\varphi^\kappa(x^\gamma;\lambda))\varphi^\mu{}_{|\alpha}\varphi^\nu{}_{|\beta} = g_{\alpha\beta}(x^\gamma) \tag{3.50}$$

We differentiate this identity with respect to the parameter λ. This presupposes, of course, that the continuous group of transformations is differentiable, and we make this assumption. We denote differentiation with respect to λ by a dot over the function considered and define the vector field

$$\Psi^\mu(x^\gamma) = \dot{\varphi}^\mu(x^\gamma;0) \tag{3.51}$$

This is indeed a vector field, since it describes the infinitesimal shift of the point x^γ under an infinitesimal increase of the parameter λ and has thus an intrinsic geometric meaning. The differentiation of (3.50) with respect to λ yields, for $\lambda = 0$,

$$g_{\mu\nu|\kappa}(\varphi^\rho)\Psi^\kappa\varphi^\mu{}_{|\alpha}\varphi^\nu{}_{|\beta} + g_{\mu\nu}(\varphi^\rho)\Psi^\mu{}_{|\alpha}\varphi^\nu{}_{|\beta} + g_{\mu\nu}(\varphi^\rho)\varphi^\mu{}_{|\alpha}\Psi^\nu{}_{|\beta} = 0 \tag{3.52}$$

If we assume that the parameter λ is chosen such that $\lambda = 0$ corresponds to the identity transformation

$$x^\mu = \varphi^\mu(x^\gamma;0) \tag{3.53}$$

we derive by differentiation

$$\delta^\mu{}_\gamma = \varphi^\mu{}_{|\gamma} \qquad \text{for } \lambda = 0 \tag{3.54}$$

Hence (3.52) reduces to the identity

$$g_{\alpha\beta|\kappa}\Psi^{\kappa} + g_{\mu\beta}\Psi^{\mu}{}_{|\alpha} + g_{\alpha\nu}\Psi^{\nu}{}_{|\beta} = 0 \tag{3.55}$$

The existence of a hidden symmetry of the metric tensor thus leads us to postulate the existence of a vector field Ψ^μ which satisfies the differential system (3.55). The integrability condition for this system is a differential relation for the metric tensor which is a covariant formulation of the symmetry. Indeed, from the theory of continuous groups it follows that a field of infinitesimal generators Ψ^μ guarantees the existence of an integral group $\varphi^\mu(x^\gamma;\lambda)$ as desired.

We do not enter into the mathematical theory of integration of the system (3.55), but wish only to bring this system into a very elegant and suggestive form. We introduce the covariant vector field

$$\Psi_\sigma = g_{\sigma\mu}\Psi^\mu \tag{3.56}$$

and calculate its covariant derivative:

$$\begin{aligned} \Psi_{\sigma\|\rho} &= \Psi_{\sigma|\rho} - \begin{Bmatrix} \mu \\ \sigma\ \rho \end{Bmatrix} \Psi_\mu = \Psi_{\sigma|\rho} - \tfrac{1}{2}(g_{\sigma\mu|\rho} + g_{\rho\mu|\sigma} - g_{\sigma\rho|\mu})\Psi^\mu \\ &= \tfrac{1}{2}(g_{\sigma\mu|\rho} - g_{\rho\mu|\sigma} + g_{\sigma\rho|\mu})\Psi^\mu + g_{\sigma\mu}\Psi^\mu{}_{|\rho} \end{aligned} \tag{3.57}$$

From this equation we obtain

$$\Psi_{\sigma\|\rho} + \Psi_{\rho\|\sigma} = g_{\sigma\rho|\mu}\Psi^\mu + g_{\sigma\mu}\Psi^\mu{}_{|\rho} + g_{\rho\mu}\Psi^\mu{}_{|\sigma} \tag{3.58}$$

Thus we can express the condition (3.55) in the form

$$\Psi_{\sigma\|\rho} + \Psi_{\rho\|\sigma} = 0 \tag{3.59}$$

A vector field which satisfies this equation is called a *Killing vector*, after its discoverer. Using the definition (3.59), we have proved that a necessary condition for *the metric $g_{\mu\nu}$ to have a hidden symmetry is that it admits a Killing vector field $\Psi_\sigma(x^\gamma)$*.

In this development, we have restricted ourselves to a one-parameter group of transformations for the sake of clarity in the exposition. The theory can be easily extended to an n-parameter group of transformations. By a straightforward generalization one verifies that the necessary condition for a metric to admit an n-parameter group of transformations is that it admits n Killing vector fields. Such a group of transformations is often called a *group of motions* of the space with the metric $g_{\mu\nu}$.

One of the most interesting physical symmetries that we shall encounter is time-independence. From the preceding discussion we see that an

invariant characterization of this property is that there must exist a Killing vector field that is timelike. This characterization has meaning even in a coordinate system where x^0 is not a convenient time label and where the metric may depend on x^0. For such a geometry a coordinate system can be found in which the metric is time-independent; it is called *stationary*. (In Chap. 6 we shall discuss also a special case of stationary metric that we shall term *static*. These two terms should not be confused.)

Exercises

3.1 Show that if a vector is parallel-displaced along a geodesic in a Riemann space, its angle with the tangent vector to the geodesic remains unchanged. Assume that the metric is positive-definite; see Exercise 1.5.

3.2 Show that lowering an index of $\xi^i{}_{\|k}$ leads to the expression (3.7) for the covariant derivative of a covariant vector field.

3.3 Let Γ^i_{kl} be a set of symmetric affine connections and *demand* that the metric tensor have a zero covariant derivative. Show that this implies $\Gamma^i_{kl} = -\begin{Bmatrix} i \\ k\ l \end{Bmatrix}$. This is an alternative motivation for working in a Riemann space.

3.4 Consider a vector field w_i on a two-dimensional plane. Show that

$$\int_A \{w_{i|k}\}\, dx^i \wedge dx^k = \oint_C w_j\, dx^j$$

for a closed curve C enclosing the area A. From this it follows that if $\{w_{i|k}\} = 0$, then the line integral is zero. Show that it then follows that the vector w_i has a vector potential ϕ, that is, $w_i = \phi_{|i}$. (This proves the theorem in Sec. 3.4 for rank 1 and two dimensions.)

3.5 Show that the following tensor identity holds by working in a tangent Lorentz space:

$$\begin{aligned} e_{\alpha\beta\gamma\delta}e^{\alpha}{}_{\sigma\omega\tau} = -[&g_{\beta\sigma}g_{\gamma\omega}g_{\delta\tau} - g_{\beta\omega}g_{\gamma\sigma}g_{\delta\tau} + g_{\beta\omega}g_{\delta\sigma}g_{\gamma\tau} - g_{\beta\sigma}g_{\delta\omega}g_{\gamma\tau} \\ &+ g_{\gamma\sigma}g_{\delta\omega}g_{\beta\tau} - g_{\delta\sigma}g_{\gamma\omega}g_{\beta\tau}] \end{aligned}$$

From this show by contraction that

$$e_{\alpha\beta\gamma\delta}e^{\alpha\beta}{}_{\omega\tau} = -2[g_{\gamma\omega}g_{\delta\tau} - g_{\delta\omega}g_{\gamma\tau}]$$

From this it is easy to show that the dual operation operating twice produces the negative of the original tensor:

$$*(*T^{\alpha\beta}) = -T^{\alpha\beta}$$

3.6 Consider a diagonal metric in four dimensions

$$ds^2 = g_{00}(dx^0)^2 + g_{11}(dx^1)^2 + g_{22}(dx^2)^2 + g_{33}(dx^3)^2$$

Show that the invariant four-volume element is

$$(\sqrt{|g_{00}|}\,dx^0)(\sqrt{|g_{11}|}\,dx^1)(\sqrt{|g_{22}|}\,dx^2)(\sqrt{|g_{33}|}\,dx^3)$$

This illustrates the identification of $\sqrt{|g_{00}|}\,dx^0$ with a physical time interval, $\sqrt{|g_{11}|}\,dx^1$ with the physical space interval in the 1 direction, etc. Write out these intervals explicitly in spherical coordinates and verify that this identification agrees with the geometric picture.

3.7 We call a second-rank tensor traceless if

$$g^{\mu\nu}T_{\mu\nu} = T^{\alpha}_{\alpha} = 0$$

Given an arbitrary second rank tensor $S_{\mu\nu}$ in four dimensions, show that

$$S_{\mu\nu} - \tfrac{1}{4}(S^{\alpha}_{\alpha})g_{\mu\nu}$$

is traceless. From this show that an arbitrary second-rank tensor may be written as the sum of an antisymmetric tensor, a symmetric traceless tensor, and a multiple of the metric tensor.

3.8 Consider the simplest case of a field of Killing vectors. In the space of special relativity the metric is independent of position, so that a translation by a constant four-vector ξ^μ is a symmetry. Write this translation as $x'^\mu = x^\mu + \lambda\xi^\mu$. What is the Killing vector corresponding to this symmetry?

3.9 The nature of the translational Killing vector is quite obvious in the above exercise. Now make a transformation from the Cartesian coordinates to cylindrical coordinates ρ, θ, z where it is not so obvious. What is the translational Killing vector in these coordinates? Verify explicitly that it satisfies the fundamental equation (3.59).

3.10 Consider a metric that is invariant under the translation in time, $x'^\mu = x^\mu + \lambda\delta^\mu{}_0$. Such a metric is clearly independent of x^0. What is the Killing vector? Verify Eq. (3.59) explicitly.

Problems

3.1 An alternative way to introduce the concept of the exterior multiplication symbol $\wedge$ of Sec. 3.3 is to consider a two-dimensional surface labeled by *intrinsic* markers (a,b) and imbedded in Euclidean three space (x^1,x^2,x^3). Any point on the surface may also be labeled with x^i, and we can form three independent Jacobians, $\partial(x^i,x^j)/\partial(a,b)$. Show that these form the components of an antisymmetric second-rank tensor. We can then define

$$dx^i \wedge dx^j = \frac{\partial(x^i,x^j)}{\partial(a,b)}\,da\,db$$

which is clearly a second-rank antisymmetric tensor. Illustrate that $ds^k = \frac{1}{2}\epsilon^{kij}\,dx^i \wedge dx^j$ represents an element of surface by considering some special cases.

3.2 Discuss how the above concepts can be generalized to surfaces of any number of dimensions in any Riemann space.

Bibliography

Compare Bibliography for Chap. 2. See also:

Fermi, E. (1922): Sopra i fenomeni che avvengono in vicinanza di una linea oraria, *Atti R. Accad. Lincei Rend. Cl. Sci. Fis. Mat. Nat.*, vol. 31_1, pp. 21–51.

Flanders, H. (1963): "Differential Forms," New York.

Grassmann, H. (1878): "Die Ausdehnungslehre von 1844," Leipzig.

Hammermesh, M. (1962): "Group Theory and Its Application to Physical Problems," New York.

Levi-Civita, T. (1927): "The Absolute Differential Calculus," London.

McConnell, A. J. (1936): "Applications of the Absolute Differential Calculus," London.

Murnaghan, F. D. (1922): "Vector Analysis and the Theory of Relativity," Baltimore.

Pirani, F. A. E. (1957): Tetrad formulation of general relativity theory, *Bull. Acad. Polón. Sci.*, vol. 5, pp. 143–147.

Schouten, J. A. (1951): "Tensor Analysis for Physicists," Oxford.

Schouten, J. A. (1954): "Ricci Calculus," 2d ed., Berlin-Göttingen-Heidelberg.

Synge, J. L. (1960): "Relativity: The General Theory," Amsterdam.

Walker, A. G. (1932): Relative coordinates, *Proc. Roy. Soc. Edinburgh*, vol. 52, p. 345.

Weinberg, S. (1972): "Gravitation and Cosmology," New York (sec. 13.1 on Killing vectors).