Preface

In the early 1900s, three events took place that dramatically changed the course of modern physics. In 1905 Albert Einstein formulated the Special Theory of Relativity. Then, in 1915, he developed the General Theory of Relativity, and around 1925 quantum mechanics took its present form. Since then, physics has progressed rapidly. Beginning in 1930, quantum mechanics and special relativity were united into what is known as the relativistic quantum field theory. This merger was very rewarding in that it provides, at the least, partial explanation of the laws and interactions governing elementary particle physics.

Among the four types of forces (strong, electromagnetic, weak, and gravitational) known today, gravity is perhaps the strangest. Weak though it is, gravity dominates the other three forces over cosmic distances. Any cosmology must be founded on a logically secure theory of gravitation.

The first three forces could be explained through particle interactions taking place in the flat space-time of special relativity. However, gravity defies such an explanation. In order to describe the mysterious force known as gravity, Einstein in 1915 was compelled to generalize the ideas of his special relativity, and he eventually connected gravity with the geometry of space-time. In other words, Einstein's General Theory of Relativity is a relativistic theory of gravitation.

For a long time, Einstein's Theory of General Relativity occupied an isolated position within the domain of general physics. This was attributable in part to the mathematical framework of the theory, which is based on Riemannian geometry, a kind of geometry not needed in most other physical applications. The extreme difficulty in devising suitable experiments that might verify the theory and the growth of more fertile fields of investigation, such as atomic and nuclear physics as well as the study of elementary particles, also contributed to the isolation of the theory.

However, Einstein's Theory of General Relativity is now enjoying renewed interest. This is due partly to the development of new technological capabilities that opened up previously inaccessible avenues for the experimental verification of general relativity and partly to the conjecture of some theoretical physicists that the fundamental difficulties confronting quantum field theory may find their resolution in a suitable combination of the two disciplines. The discovery of extremely compact celestial objects—neutron stars and black holes, for instance—provided the final turning point. The study of these objects demanded the application of Einstein's Theory of General Relativity. Today, physics and astronomy have joined forces to form the discipline called relativistic astrophysics. Einstein's Theory of General Relativity is also essential to modern cosmology, since the overall space-time structure is intimately related to the gravitational field. In the past decade interest in cosmology and general relativity has grown considerably.

Today, there is increased demand for undergraduate courses in relativity and cosmology. There are many advanced books on the Theory of General Relativity and cosmology for the specialist, and many elementary expositions for the lay reader. But there is a gap at the undergraduate level. This book is an attempt to fill the gap. We will try to make available to the student a working acquaintance with the concepts and fundamental ideas in general relativity and modern cosmology. For the modes of calculation we choose the old-fashioned tensor calculus for pedagogical reasons. Most undergraduates have not been exposed to the many new formalisms developed in general relativity. Hopefully after reading this book, the student can continue delving more deeply into particular aspects or topics in general relativity and cosmology that interest him or her.

This book evolved from a set of lecture notes for a course that I have taught over the past 10 years. I am making the assumption that the student has been exposed to a calculus-based course in general physics and a course in calculus (including the handling of differentiations of field equations). Some exposure to tensor analysis would be helpful but is not necessary; this subject is covered in the text.

The student will find that in the derivations of equations, a generous amount of detail has been given. However, to ensure that the student does not lose sight of the development underway, some of the more lengthy and tedious algebraic manipulations have been omitted.

Turlock, California *Tai L. Chow, Ph.D.*

Chapter 2
Curvilinear Coordinates and General Tensors

2.1 Curvilinear Coordinates

We devote this chapter to the development of four-dimensional geometry in arbitrary curvilinear coordinates. We shall deal with field quantities. A field quantity has the same nature at all points of space. Such a quantity will be disturbed by the curvature.

If we take a point quantity Q (or one of its components if it has several), we can differentiate with respect to any of the four coordinates. We write the result

$$\frac{\partial Q}{\partial x^{\mu}} = Q_{,\mu}.$$

A subscript preceded by a comma will always denote a derivative in this way. We put the index downstairs in order that we may maintain a balancing of the indexes in the general equations. We can see this balancing by noting that the change in Q, when we move from the point x^{μ} to a neighboring point $x^{\mu} + dx^{\mu}$, is

$$\delta Q = Q_{,\mu} \delta x^{\mu} \tag{2.1}$$

where summation over μ is understood. The repeated indexes appearing once in the lower and once in the upper position are automatically summed over; this is Einstein's summation convention.

Let us consider the transformation from one coordinate system x^0, x^1, x^2, x^3 to another x'^0, x'^1, x'^2, x'^3:

$$x^{\mu} = f^{\mu}(x'^0, x'^1, x'^2, x'^3) \tag{2.2}$$

where the f^{μ} are certain functions. When we transform the coordinates, their differentials transform according to the relation

$$dx^{\mu} = \frac{\partial x^{\mu}}{\partial x'^{\nu}} dx'^{\nu}. \tag{2.3}$$

Any set of four quantities $A^{\mu} (\mu = 0, 1, 2, 3)$ which, under coordinate change, transform like the coordinate differentials, is called a *contravariant vector:*

$$A^{\mu} = \frac{\partial x^{\mu}}{\partial x'^{\nu}} A'^{\nu}. \tag{2.4}$$

If ϕ is somewhat scalar, under a coordinate change, the four quantities $\partial\phi/\partial x^{\mu}$ transform according to the formula

$$\frac{\partial \phi}{\partial x^{\mu}} = \frac{\partial x'^{\nu}}{\partial x^{\mu}} \frac{\partial \phi}{\partial x'^{\nu}}. \tag{2.5}$$

Any set of four quantities $A_{\mu}(\mu = 0, 1, 2, 3)$ that, under a coordinate transformation, transform like the derivatives of a scalar is called a *covariant vector*:

$$A_{\mu} = \frac{\partial x'^{\nu}}{\partial x^{\mu}} A'_{\nu} \tag{2.6}$$

From the two contravariant vectors A^{μ} and B^{μ} we may form the 16 quantities $A^{\mu}B^{\nu}(\mu, \nu = 0, 1, 2, 3)$. These 16 quantities form the components of a contravariant tensor of the second rank: Any aggregate of 16 quantities $T^{\mu\nu}$ that, under a coordinate transformation, transform like the products of two contravariant vectors

$$T^{\mu\nu} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}} \frac{\partial x^{\nu}}{\partial x'^{\beta}} T'^{\alpha\beta} \tag{2.7}$$

is a contravariant tensor of rank two. We may also form a covariant tensor of rank two from two covariant vectors, which transform according to the formula

$$T_{\mu\nu} = \frac{\partial x'^{\alpha}}{\partial x^{\mu}} \frac{\partial x'^{\beta}}{\partial x^{\nu}} T'_{\alpha\beta}. \tag{2.8}$$

Similarly, we can form a mixed tensor T^{μ}_{ν} of order two that transforms as follows

$$T^{\mu}_{\nu} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}} \frac{\partial x'^{\beta}}{\partial x^{\nu}} T'^{\alpha}_{\beta}. \tag{2.9}$$

We may continue this process and multiply more than two vectors together, taking care that their indexes are all different. In this way we can construct tensors of higher rank. The total number of free indexes of a tensor is called its *rank* (or *order*).

We may set a subscript equal to a superscript and sum over all values of this index, which results in a tensor having two fewer free indexes than the original one. This process is called *contraction*. For example, if we start with a fourth-order tensor $T^{\mu}{}_{\nu\rho}{}^{\sigma}$, one way of contracting it is to put $\sigma = \rho$, which gives the second-rank tensor $T^{\mu}{}_{\nu\rho}^{\rho}$, having only 16 components, arising from the four values of μ and ν. We could contract again to get the scalar $T^{\mu}{}_{\mu\rho}^{\rho}$ with just one component.

It is easy to show that the inner product of contravariant and covariant vectors, $A_{\mu}B^{\mu}$, is an invariant, that is, independent of the coordinate system

$$A'_{\mu}B'^{\mu} = \frac{\partial x^{\alpha}}{\partial x'^{\mu}} \frac{\partial x'^{\mu}}{\partial x^{\beta}} A_{\alpha}B^{\beta} = \frac{\partial x^{\alpha}}{\partial x^{\beta}} A_{\alpha}B^{\beta} = \delta^{\alpha}_{\beta} A_{\alpha}B^{\beta} = A_{\alpha}B^{\alpha}.$$

The square of the line element in curvilinear coordinates is a quadratic form in the differentials dx^μ : $ds^2 = \mathrm{g}_{\mu\nu} dx^\mu dx^\nu$. Since the contracted product of $\mathrm{g}_{\mu\nu}$ and the contravariant tensor $dx^\mu dx^\nu$ is a scalar, the $\mathrm{g}_{\mu\nu}$ forms a covariant tensor:

$$ds^2 = g'_{\alpha\beta} dx'^\alpha dx'^\beta = g_{\alpha\beta} dx^\alpha dx^\beta$$

Now, $dx'^\alpha = (\partial x'^\alpha / (\partial x^\mu) dx^\mu$, so that

$$g'_{\alpha\beta} \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x'^\beta}{\partial x^\nu} dx^\mu dx^\nu = g_{\mu\nu} dx^\mu dx^\nu$$

or

$$\left(g'_{\alpha\beta} \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x'^\beta}{\partial x^\nu} - g_{\mu\nu} \right) dx^\mu dx^\nu = 0.$$

The above equation is identically zero for arbitrary dx^μ, so we have

$$g_{\mu\nu} = \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x'^\beta}{\partial x^\nu} g'_{\alpha\beta} \tag{2.10}$$

that is, $g_{\mu\nu}$ is a covariant tensor of rank two. It is called the *metric tensor* or the *fundamental tensor*. The metric tensor is locally Minkoskian.

So far, covariant and contravariant vectors have no direct connection with each other except that their inner product is an invariant. A space in which covariant and contravariant vectors exist separately is called *affine*. Physical quantities are independent of the particular choice of the mode of description, that is, independent of the possible choices of contravariance or covariance. In metric space, contravariant and covariant vectors can be converted into each other with the help of the fundamental tensor $g_{\mu\nu}$. For example, we can get the covariant vector A_μ from the contravariant vector A^ν

$$A_\mu = g_{\mu\nu} A^\nu. \tag{2.11}$$

Since the determinant $|g|$ does not vanish, these equations can be solved for A^ν in terms of the A_μ. Let the result be

$$A^\nu = g^{\mu\nu} A_\mu \tag{2.12}$$

By combining the two transformations (2.11) and (2.12), we have

$$A_\mu = g_{\mu\nu} g^{\nu\alpha} A_\alpha.$$

Since the equation must hold for any four quantities A_μ, we can infer

$$g_{\mu\nu} g^{\nu\alpha} = \delta_\mu{}^\alpha. \tag{2.13}$$

In other words, $g^{\mu\nu}$ is the inverse of $g_{\mu\nu}$ and vice versa, that is

$$g^{\mu\nu} = \frac{M^{\mu\nu}}{|g|} \tag{2.14}$$

where $M^{\mu\nu}$ is the minor of the element $g_{\mu\nu}$.

Equation (2.11) may be used to lower any upper index occurring in a tensor. Similarly, (2.12) can be used to raise any downstairs index. It is necessary to remember the position from which the index was lowered or raised, because when we bring the index back to its original site, we do not want to interchange the order of indexes, in general $T^{\mu\nu} \neq T^{\nu\mu}$.

Two tensors, $A_{\mu\nu}$ and $B^{\mu\nu}$, are said to be reciprocal to each other if

$$A_{\mu\nu} B^{\nu\alpha} = \delta^{\alpha}{}_{\mu}. \tag{2.15}$$

A tensor is called symmetric with respect to two contravariant or two covariant indexes if its components remain unaltered on interchange of the indexes. For example, if $A^{\mu\nu\alpha}_{\beta\gamma} = A^{\nu\mu\alpha}_{\beta\gamma}$, the tensor is symmetric in μ and ν. If a tensor is symmetric with respect to any two contravariant and any two covariant indexes, it is called symmetric.

A tensor is called skew-symmetric with respect to two contravariant or two covariant indexes if its components change sign upon interchange of the indexes. Thus, if $A^{\mu\nu\alpha}_{\beta\gamma} = -A^{\nu\mu\alpha}_{\beta\gamma}$, the tensor is skew-symmetric in μ and ν.

If a tensor is symmetric (or skew-symmetric) with respect to two indexes in one coordinate system, it remains symmetric (skew-symmetric) with respect to those two indexes in any other coordinate system. It is easy to prove this. For example, if $B^{\alpha\beta}$ is symmetric, $B^{\alpha\beta} = B^{\beta\alpha}$, then

$$B'^{\alpha\beta} = \frac{\partial x'^{\alpha}}{\partial x^{\gamma}} \frac{\partial x'^{\beta}}{\partial x^{\delta}} B^{\gamma\delta} = \frac{\partial x'^{\alpha}}{\partial x^{\gamma}} \frac{\partial x'^{\beta}}{\partial x^{\delta}} B^{\delta\gamma} = B'^{\beta\alpha} \tag{2.16}$$

i.e., the tensor remains symmetric in the primed coordinate system.

Every tensor can be expressed as the sum of two tensors, one of which is symmetric and the other skew-symmetric in a pair of covariant or contravariant indices. Consider, for example, the tensor $B^{\alpha\beta}$. We can write it as

$$B^{\alpha\beta} = 1/2(B^{\alpha\beta} + B^{\beta\alpha}) + 1/2(B^{\alpha\beta} - B^{\beta\alpha}) \tag{2.17}$$

with the first term on the right-hand side symmetric and the second term skew-symmetric. By similar reasoning the result is seen to be true for any tensor.

It is obvious that the sum or difference of two or more tensors of the same rank and type (i.e., same number of contravariant indices and same number of covariant indices) is also a tensor of the same rank and type. Thus if $A_{\lambda}{}^{\mu\nu}$ and $B_{\lambda}{}^{\mu\nu}$ are tensors, then $C_{\lambda}{}^{\mu\nu} = A_{\lambda}{}^{\mu\nu} + B_{\lambda}{}^{\mu\nu}$ and $D_{\lambda}{}^{\mu\nu} = A_{\lambda}{}^{\mu\nu} - B_{\lambda}{}^{\mu\nu}$ are also tensors.

A given quantity $N^{\mu\ldots}_{\nu\rho\ldots}$ with various up and down indexes may or may not be a tensor. We can test whether it is a tensor or not by using the quotient law, which can be stated as follows:

Suppose we have a quantity X and we do not know whether it is a tensor or not. If an inner product of X with an arbitrary tensor is a tensor, then X is also a tensor.

For example, let $X = P_{\lambda\mu\nu}$, and A^{λ} is an arbitrary contravariant vector; if $P_{\lambda\mu\nu} A^{\lambda} = Q_{\mu\nu}$ is a tensor, then $P_{\lambda\mu\nu}$ is a contravariant tensor of rank 3. We can prove this explicitly:

$$A^\lambda P_{\lambda\mu\nu} = \frac{\partial x'^\alpha}{\partial x^\mu}\frac{\partial x'^\beta}{\partial x^\nu} A'^\gamma\, P'_{\gamma\alpha\beta}.$$

but

$$A'^\gamma = \frac{\partial x'^\gamma}{\partial x^\lambda} A^\lambda.$$

Hence,

$$A^\lambda P_{\lambda\mu\nu} = \frac{\partial x'^\alpha}{\partial x^\mu}\frac{\partial x'^\beta}{\partial x^\nu}\frac{\partial x'^\gamma}{\partial x^\lambda} A^\lambda P'_{\gamma\alpha\beta}.$$

This equation must hold for all values of A^λ, so we have

$$P_{\lambda\mu\nu} = \frac{\partial x'^\alpha}{\partial x^\mu}\frac{\partial x'^\beta}{\partial x^\nu}\frac{\partial x'^\gamma}{\partial x^\lambda} P'_{\gamma\alpha\beta}, \tag{2.18}$$

showing that $P_{\lambda\mu\nu}$ is a contravariant tensor of rank 3.

For a nontensor $N^{\mu\cdots}_{\nu\rho\cdots}$ we can raise and lower indexes by the same rules as for a tensor. Thus, for example

$$g^{\alpha\nu} N^{\rho}_{\nu\eta} = N^{\alpha\mu}_{\eta}. \tag{2.19}$$

2.2 Parallel Displacement and Covariant Differentiation

In this section and the following three sections we will give the full of apparatus of differential geometry. The reader may be in danger of being overwhelmed by algebra, but to simplify mathematics is not to make it simple, either. Please do not despair; just relax and try to enjoy it.

We have seen that a covariant vector is transformed according to the formula

$$A_i = \frac{\partial x'^k}{\partial x^i} A'_k \tag{2.20}$$

where the coefficients are functions of the coordinates. So vectors at different points transform differently. Because of this fact, dA_i is not a vector, since it is the difference of two vectors located at two infinitesimally separated points of space-time. We can easily verify this directly from (2.20)

$$dA_i = \frac{\partial x'^k}{\partial x^i} dA'_k + A'_k \frac{\partial^2 x'^k}{\partial x^i \partial x^j} dx^j$$

which shows that dA_i does not transform at all like a vector. The same also applies to the differential of a contravariant vector.

When using curvilinear coordinates, a differential can be obtained only when the two vectors to be subtracted from each other are located at the same point in space-time. In order to do so, we must what we call *parallel displace* one of the vectors to the point where the other vector is located, after which we determine the difference of two vectors, which now refer to one and the same point in space-time.

The concept of parallel displacement of a vector is very clear in Cartesian coordinates: displace a vector parallel to itself so that both its length and orientation are unchanged. We can extend the idea of parallel displacement of a vector to curved spaces in a consistent way. This requires us to assume that there always exist Galilean coordinates in the immediate vicinity of a point in space-time; in such a coordinate system the idea of an infinitesimal parallel displacement of a vector works. In other words, a vector can be transported parallel to itself without changing its length and orientation. We illustrate, in Fig. 2.1, with the example of a curved two-dimensional surface in a three-dimensional Euclidean space. During the infinitesimal parallel displacement of two vectors, A^μ and B^μ, the angle between them clearly remains unchanged, and so the inner (scalar) product of two vectors, $A_\mu B^\mu$, does not change under parallel displacement. For arbitrary coordinates we define the operation of infinitesimal parallel displacement of a vector A^μ from Point P to a neighboring Point Q to be one that leaves the inner product with an arbitrary vector B^μ invariant.

Parallel displacements are independent of the paths taken on a Euclidean plane (a flat surface), as shown in Fig. 2.2a. On a curved surface, however, we will obtain a different final result on the path taken (Figure 2.2b).

We can transfer a vector continuously along a path by the process of parallel displacement. In curvilinear coordinates, the components of a vector would be expected

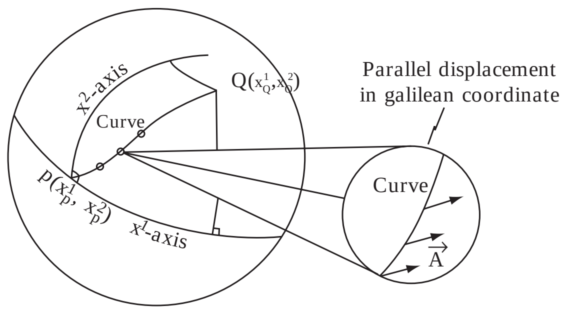

Fig. 2.1 Parallel displacement in curvilinear coordinates.

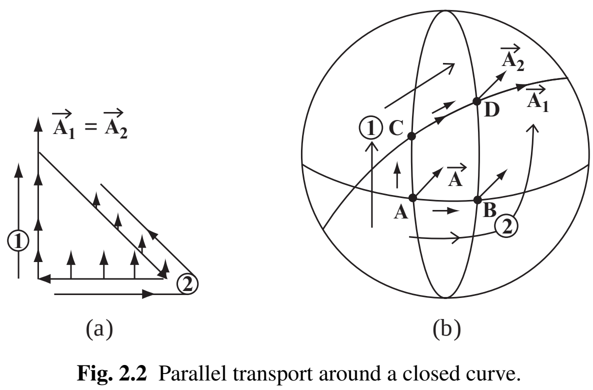

Fig. 2.2 Parallel transport around a closed curve.

to change under a parallel displacement, unlike the case of a Cartesian coordinate. Therefore, if A^μ are the components of a contravariant vector at the point $\mathrm{P}(x^\mu)$, and $A^\mu + dA^\mu$ the components at a neighboring point $\mathrm{Q}(x^\mu + dx^\mu)$, where

$$A^\mu + dA^\mu = A^\mu + \frac{\partial A^\mu}{\partial x^\sigma} dx^\sigma \tag{2.21}$$

an infinitesimal parallel displacement of A^μ from P to Q would produce a variation of its components, δA^μ. δA^μ should be a linear function of the coordinate differentials and the components A^μ. We write it in the form

$$\delta A^\nu = -\Gamma^\nu_{\alpha\beta} A^\alpha dx^\beta \tag{2.22}$$

where the $\Gamma^\nu_{\alpha\beta}$ are certain functions of the coordinates and are called Christoffel symbols of the second kind. Their form depends on the coordinate system. It will be proved in Section 4.4 that in a Galilean coordinate system $\Gamma^\nu_{\alpha\beta} = 0$. From this it is already clear that the quantities $\Gamma^\nu_{\alpha\beta}$ do not form a tensor, since a tensor that is equal to zero in one coordinate system is equal to zero in every other one. In a curved space it is impossible to make all the $\Gamma^\nu_{\alpha\beta}$ vanish over all of space.

The vector resulting from parallel displacement from Point P to Point Q is $A^\mu + \delta A^\mu$. Subtraction of these two quantities gives us

$$DA^\mu = dA^\mu - \delta A^\mu = \left(\frac{\partial A^\mu}{\partial x^\sigma} + \Gamma^\mu_{\sigma\alpha} A^\alpha \right) dx^\sigma . \tag{2.23}$$

We would expect the difference $dA^\mu - \delta A^\mu$ to be a vector since it is the difference of two vectors at the same point; the quantity

$$\frac{\partial A^\mu}{\partial x^\sigma} + \Gamma^\mu_{\sigma\alpha} A^\alpha$$

then is a mixed tensor called the covariant derivative of A^μ and written

$$A^\mu_{;\sigma} = \frac{\partial A^\mu}{\partial x^\sigma} + \Gamma^\mu_{\sigma\alpha} A^\alpha . \tag{2.24}$$

From $\delta(A_\mu A^\mu) = 0$ it follows, using (2.22), that

$$\delta A_\mu = \Gamma^\alpha_{\mu\beta} A_\alpha dx^\beta .$$

From this and a similar procedure that leads to (2.23) and (2.24), we obtain the covariant derivative of A_μ:

$$A_{\mu;\sigma} = \frac{\partial A_\mu}{\partial x^\sigma} - \Gamma^\alpha_{\mu\sigma} A_\alpha . \tag{2.25}$$

The tensor character of (2.24) and (2.25) can be established formally by showing that they obey the required transformation laws. This will require us first to establish the transformation laws for the $\Gamma^\alpha_{\mu\sigma}$. It is not difficult to do this, but very tedious, and so we shall not do it in this book.

To obtain the contravariant derivative, we raise the index that denotes differentiation,

$$A^{\mu;\sigma} = g^{\sigma\ \alpha} A^{\mu}_{;\alpha}. \tag{2.26}$$

In Galilean coordinates, $\Gamma^{\alpha}_{\mu\sigma} = 0$, and so covariant differentiation reduces to ordinary differentiation.

We may also obtain the covariant derivative of a tensor by determining the change in the tensor under an infinitesimal parallel displacement. For example, let us consider any arbitrary tensor $T^{\mu\nu}$ expressible as a product of two contravariant vectors $A^{\mu}B^{\nu}$. Under infinitesimal parallel displacement

$$\delta(A^{\mu}B^{\nu}) = A^{\mu}\delta B^{\nu} + B^{\nu}\delta A^{\mu} = -A^{\mu}\Gamma^{\nu}{}_{\alpha\beta}B^{\alpha}dx^{\beta} - B^{\nu}\Gamma^{\mu}{}_{\beta\sigma}A^{\sigma}dx^{\beta}.$$

By virtue of the linearity of this transformation we also have

$$\delta A^{\mu\nu} = -\left(A^{\mu\beta}\Gamma^{\nu}{}_{\beta\alpha} + A^{\beta\nu}\Gamma^{\mu}{}_{\beta\alpha}\right)dx^{\alpha}.$$

Substituting this in

$$DA^{\mu\nu} = dA^{\mu\nu} - \delta A^{\mu\nu} = A^{\mu\nu}{}_{;\alpha}dx^{\alpha}$$

we get the covariant derivative of the tensor $T^{\mu\nu}$ in the form

$$T^{\mu\nu}{}_{;\alpha} = \frac{\partial T^{\mu\nu}}{\partial x^{\alpha}} + \Gamma^{\mu}{}_{\beta\alpha}T^{\beta\nu} + \Gamma^{\nu}{}_{\beta\alpha}T^{\mu\beta}. \tag{2.27}$$

In similar fashion we obtain the covariant derivative of the mixed tensor $T^{\mu}{}_{\nu}$ and the covariant tensor $T_{\nu\mu}$ in the form

$$T^{\mu}{}_{\nu;\alpha} = \frac{\partial T^{\mu}{}_{\nu}}{\partial x^{\alpha}} - \Gamma^{\beta}{}_{\nu\alpha}T^{\mu}{}_{\beta} + \Gamma^{\mu}{}_{\beta\alpha}T^{\beta}{}_{\nu}, \tag{2.28}$$

$$T_{\mu\nu;\alpha} = \frac{\partial T_{\mu\nu}}{\partial x^{\alpha}} - \Gamma^{\beta}{}_{\mu\alpha}T_{\beta\nu} - \Gamma^{\beta}{}_{\nu\alpha}T_{\mu\beta}. \tag{2.29}$$

One can similarly determine the covariant derivative of a tensor of arbitrary rank. In doing this one finds the following rule of covariant differentiation:

To obtain the covariant derivative of the tensor $T^{\cdots\cdots}_{\cdots\cdots}$ *with respect to* x^{μ}, *you add to the ordinary derivative* $\partial T^{\cdots\cdots}_{\cdots\cdots}/\partial x^{\mu}$ *for each covariant index* $\nu(T^{\cdots\cdots}_{..\nu..})$ *a term* $-\Gamma^{\alpha}_{\mu\nu}T^{\cdots\cdots}_{..\alpha..}$, *and for each contravariant index* $\nu(T^{..\nu..}_{\cdots\cdots})$ *a term* $+\Gamma^{\nu}_{\alpha\mu}T^{..\alpha..}_{\cdots\cdots}$.

The covariant derivative of the metric tensor $g_{\mu\nu}$ is zero. To show this we note that the relation

$$DA_{\mu} = g_{\mu\nu}DA^{\nu}$$

is valid for the vector DA_{μ} as for any vector. On the other hand, we have $A_{\mu} = g_{\mu\nu}A^{\nu}$, so that

$$DA_{\mu} = D(g_{\mu\nu}A^{\nu}) = g_{\mu\nu}DA^{\nu} + A^{\nu}Dg_{\mu\nu}.$$

Comparing with $DA_{\mu} = g_{\mu\nu}DA^{\nu}$, we have $A^{\nu}Dg_{\mu\nu} = 0$. But the vector A^{ν} is arbitrary, so

$$Dg_{\mu\nu} = 0.$$

Therefore, the covariant derivative is

$$g_{\mu\nu;\alpha} = 0. \tag{2.30}$$

Thus, $g_{\mu\nu}$ may be considered as a constant during covariant differentiation.

The covariant derivative of a product can be found by the same rule as for ordinary differentiation of products. In doing this we must consider the covariant derivative of a scalar ϕ as an ordinary derivative, that is, as the covariant vector $\phi_k = \partial\phi/\partial x^k$, in accordance with the fact that for a scalar $\delta\phi = 0$, and hence $D\phi = d\phi$. For example

$$(A_\mu B_\nu)_{;\alpha} = A_{\mu;\alpha} B_\nu + A_\mu B_{\nu;\alpha}. \tag{2.31}$$

2.3 Symmetry Properties of the Christoffel Symbols

We now show that $\Gamma^\nu_{\alpha\beta}$ is symmetric in the subscripts. If δA^ν is a coordinate differential dx^ν, then (2.22) becomes

$$\delta(dx^\nu) = -\Gamma^\nu_{\alpha\beta} dx^\alpha dx^\beta. \tag{2.32}$$

Next, we return to the local Cartesian coordinate system by the transformations

$$\left.\begin{aligned} x^\alpha &= f^\alpha(x'^1, x'^2, \ldots.) \\ x'^\alpha &= \varphi^\alpha(x^1, x^2, \ldots\ldots) \end{aligned}\right\} \tag{2.33}$$

where the primed coordinates are local Cartesian coordinates. From (2.33) we obtain

$$dx^\alpha = \frac{\partial f^\alpha}{\partial x'^\beta} dx'^\beta. \tag{2.34}$$

Under a parallel displacement, $\delta(dx'^\beta) = 0$, so that from (2.34) we have

$$\delta(dx^\nu) = \frac{\partial^2 f^\nu}{\partial x'^\delta \partial x'^\gamma} dx'^\delta dx'^\gamma = \frac{\partial^2 f^\nu}{\partial x'^\delta \partial x'^\gamma} \frac{\partial \varphi^\delta}{\partial x^\alpha} \frac{\partial \varphi^\gamma}{\partial x^\beta} dx^\alpha dx^\beta.$$

Comparing this with (2.32) we obtain

$$\Gamma^\nu_{\alpha\beta} = -\frac{\partial^2 f^\nu}{\partial x'^\delta \partial x'^\gamma} \frac{\partial \varphi^\delta}{\partial x^\alpha} \frac{\partial \varphi^\gamma}{\partial x^\beta}. \tag{2.35}$$

The right-hand side is clearly symmetric in the indexes α and β, so that $\Gamma^\nu_{\alpha\beta}$ is also symmetric in α and β.

2.4 Christoffel Symbols and the Metric Tensor

It is very useful to express the Γ's in terms of the metric tensors. Let A^μ be any contravariant vector, $A_\mu = g_{\mu\nu} A^\nu$ a covariant vector. From the definition of parallel displacement $\delta(A_\mu A^\mu) = 0$, we have

$$\delta(A_\mu A^\mu) = g_{\mu\nu}(x^\mu + dx^\mu)[A^\nu + \delta A^\nu][A^\mu + \delta A^\mu] - g_{\mu\nu}(x^\mu) A^\nu A^\mu = 0.$$

Carrying out these operations gives us

$$\frac{\partial g_{\mu\nu}}{\partial x^\alpha} A^\mu A^\nu dx^\alpha + g_{\mu\nu} A^\mu \delta A^\nu + g_{\mu\nu} A^\nu \delta A^\mu = 0.$$

Making use of (2.22) to eliminate δA^μ and δA^ν gives us

$$\frac{\partial g_{\mu\nu}}{\partial x^\alpha} - g_{\nu\beta}\Gamma^\beta_{\nu\alpha} - g_{\nu\beta}\Gamma^\beta_{\mu\alpha} = 0. \tag{2.36}$$

Now, $\Gamma^\beta_{\nu\alpha}$ is symmetric in the lower indexes ν and α, and this symmetry allows permutation of ν and α to obtain

$$\frac{\partial g_{\mu\alpha}}{\partial x^\nu} - g_{\mu\beta}\Gamma^\beta_{\nu\alpha} - g_{\alpha\beta}\Gamma^\beta_{\mu\nu} = 0. \tag{2.37}$$

Similarly, we write

$$\frac{\partial g_{\nu\alpha}}{\partial x^\mu} - g_{\nu\beta}\Gamma^\beta_{\mu\alpha} - g_{\alpha\beta}\Gamma^\beta_{\mu\nu} = 0. \tag{2.38}$$

Solving (2.36), (2.37), (2.38) for $\Gamma^\gamma_{\mu\alpha}$, we obtain

$$\Gamma^\gamma_{\mu\alpha} = \frac{1}{2} g^{\gamma\nu}\left[\frac{\partial g_{\nu\mu}}{\partial x^\alpha} + \frac{\partial g_{\nu\alpha}}{\partial x^\mu} - \frac{\partial g_{\mu\alpha}}{\partial x^\nu}\right]. \tag{2.39}$$

The Christoffel symbol of the first kind is

$$\Gamma_{\nu,\mu\alpha} = \frac{1}{2}\left[\frac{\partial g_{\nu\mu}}{\partial x^\alpha} + \frac{\partial g_{\nu\alpha}}{\partial x^\mu} - \frac{\partial g_{\nu\alpha}}{\partial x^\nu}\right]. \tag{2.40}$$

It is often written as $[\mu\alpha, \nu]$. Clearly $\Gamma_{\nu,\mu\alpha} = \Gamma_{\nu,\alpha\mu}$. The Christoffel symbols are also known as the *affine connections*. The Christoffel symbols all vanish in Galilean coordinates, as the metric tensors are all constants in Galilean coordinates.

The equation $g_{\mu\nu;\sigma} = 0$ can be used to offer an alternative derivation of (2.39) and (2.40). We write in accordance with the general definition (2.29):

$$g_{\mu\nu;\alpha} = \frac{\partial g_{\mu\nu}}{\partial x^\alpha} - g_{\beta\nu}\Gamma^\beta_{\mu\alpha} - g_{\mu\beta}\Gamma^\beta_{\nu\alpha} = \frac{\partial g_{\mu\nu}}{\partial x^\alpha} - \Gamma_{\nu,\mu\alpha} - \Gamma_{\mu,\nu\alpha} = 0.$$

From this we have, permuting the indexes μ, ν, α:

$$\frac{\partial g_{\mu\nu}}{\partial x^{\alpha}} = \Gamma_{\nu,\mu\alpha} + \Gamma_{\mu,\nu\alpha}.$$
$$\frac{\partial g_{\alpha\mu}}{\partial x^{\nu}} = \Gamma_{\mu,\nu\alpha} + \Gamma_{\alpha,\mu\nu}.$$
$$\frac{\partial g_{\nu\alpha}}{\partial x^{\mu}} = \Gamma_{-\alpha,\nu\mu} - \Gamma_{\nu,\alpha\mu}.$$

Taking half the sum of these equations and remembering that $\Gamma_{\mu,\nu\alpha} = \Gamma_{\mu,\alpha\nu}$, we find

$$\Gamma_{\mu,\nu\alpha} = \frac{1}{2}\left(\frac{\partial g_{\mu\nu}}{\partial x^{\alpha}} + \frac{\partial g_{\mu\alpha}}{\partial x^{\nu}} - \frac{\partial g_{\nu\alpha}}{\partial x^{\mu}}\right). \tag{2.41}$$

From this we have for the symbols $\Gamma^{\mu}_{\nu\alpha} = g^{\nu\beta}\Gamma_{\mu,\nu\alpha}$

$$\Gamma^{\mu}_{\nu\alpha} = \frac{1}{2}g^{\mu\beta}\left(\frac{\partial g_{\beta\nu}}{\partial x^{\alpha}} + \frac{\partial g_{\beta\alpha}}{\partial x^{\nu}} - \frac{\partial g_{\nu\alpha}}{\partial x^{\beta}}\right). \tag{2.42}$$

A coordinate system in which the Christoffel symbols vanish at Point P is called a *geodesic coordinate system*, and Point P is said to be the *pole*.

2.5 Geodesics

As an application of the notion of parallel displacement and covariant differentiation, let's consider the geodesic equation. A geodesic is the curve defined by the requirement that each element of it is a parallel displacement of the preceding element. We shall see later that the world line of a point-like particle not acted upon by any forces, except gravitation, is a time-like geodesic.

If we take a point with coordinates x^{μ} and move it along a path, we then have x^{μ} as a function of some parameter s. There is a tangent vector $t^{\mu} = dx^{\mu}/ds$ at each point of the path. As we go along the path the vector t^{μ} gets shifted by parallel displacement: we shift the initial position from x^{μ} to $x^{\mu} + t^{\mu}ds$, and then shift the vector t^{μ} to this new position by parallel displacement, then shift the point again in the direction fixed by the new t^{μ}, and so on. If we are given the initial point and the initial value of the vector t^{μ}, not only can the path be determined but also the parameter s along it. A path produced in this way is called a geodesic. We get the geodesic equations by applying (2.22) with $A^{\mu} = t^{\mu}$

$$\frac{dt^{\nu}}{ds} + \Gamma^{\nu}_{\mu\sigma}t^{\mu}\frac{dx^{\sigma}}{ds} = 0, \tag{2.43}$$

or

$$\frac{d^2x^{\nu}}{ds^2} + \Gamma^{\nu}_{\mu\sigma}\frac{dx^{\sigma}}{ds}\frac{dx^{\mu}}{ds} = 0. \tag{2.44}$$

If the vector t^μ is initially a null vector, it always remains a null vector and the path is called a *null geodesic*. If the vector t^μ is initially time-like (i.e., $t^\mu t^\mu > 0$), it is always time-like and we have a time-like geodesic. If t^μ is initially space-like ($t^\mu\ t^\mu < 0$), it is always space-like and we have a space-like geodesic.

For a time-like geodesic we may multiply the initial t^μ by a factor so as to make its length unity. This requires only a change in the scale of s. The vector t^μ now always has a unit length. It is simply the velocity vector $u^\mu = dx^\mu/d\tau$, and the parameter s has becomes the proper time τ. (2.43) becomes

$$\frac{du^\mu}{d\tau} + \Gamma^\nu_{\mu\sigma} u^\mu u^\sigma = 0 \tag{2.43a}$$

and (2.44) becomes

$$\frac{d^2x^\nu}{d\tau^2} + \Gamma^\nu_{\mu\sigma} \frac{dx^\sigma}{d\tau} \frac{dx^\mu}{d\tau} = 0. \tag{2.44a}$$

We make the physical assumption that the world line of a particle not acted on by any forces, except gravitation, is a time-like geodesic. Note that the terms $-m\Gamma^\nu_{\mu\sigma} u^\mu u^\sigma$ in (2.43a) may be interpreted as the gravitational forces, and the components of the metric tensor $g_{\mu\nu}$ play the role of the classical gravitational potential (as the Christoffel symbol is proportional to the derivatives of the metric tensor; see [2.39] and [2.40]).

Choosing a local Galilean frame in which $g_{\mu\nu}$ = constants, the Christoffel symbols $\Gamma^\nu_{\mu\sigma} = 0$, and $du^\mu/d\tau = 0$. Therefore, the gravitational forces can be locally eliminated, and the geodesic equations can be locally reduced to the special relativistic equations of motion, in agreement with the equivalence.

The path of a light ray is a null geodesic. It is fixed by (2.44) referring to some parameter s along the path. The proper time τ cannot now be used because $d\tau$ vanishes.

2.6 The Stationary Property of Geodesics

We now examine the stationary property of geodesics. A geodesic that is not a null geodesic joining two points P and Q has a stationary value compared with the interval (line element) measured along another neighboring curve joining P and Q. This property holds good for a straight line in flat space and, in that case, it is also true that the straight line gives the shortest interval from one point to another. In curved space, the geodesic is no longer a straight line because space-time is no longer flat, and the particle motion is not rectilinear and uniform, in general. However, we can show that the geodesic is a path of extreme length. (We will not enquire whether or not the geodesic in a curved space gives the minimum or maximum value of the interval between any of its points). To show this, we demonstrate that the relations which must be satisfied to give a stationary value to the integral

$$s = \int \sqrt{g_{\lambda\mu} dx^\lambda dx^\mu} \tag{2.45}$$

are simply the equations of geodesics (2.44) of the previous section. Let us first introduce a parameter α and write (2.45) as

$$s = \int_{\alpha_P}^{\alpha_Q} \left[g_{\lambda\mu} \frac{dx^\lambda}{d\alpha} \frac{dx^\mu}{d\alpha} \right]^{1/2} d\alpha$$

where α varies from point to point of the geodesic curve described by the relations which we are seeking for, and we write it as

$$x^\mu = f^\mu(\alpha). \tag{2.46}$$

Any other neighboring curve joining P and Q has equations of the form

$$\bar{x}^\lambda = x^\lambda + \varepsilon y^\lambda = f^\lambda(\alpha) + \varepsilon y^\lambda(\alpha)$$

where $y^\lambda = 0$ at the end points P and Q, i.e., at $\alpha = \alpha_{\mathrm{P}}$ and $\alpha = \alpha_{\mathrm{Q}}$, and ε is a small quantity whose square and higher powers are negligibly small. If $\bar{s}$ is the line element along the neighboring curve joining $P$and Q, then

$$\bar{s} = \int_{\alpha_P}^{\alpha_Q} \left(g_{\lambda\mu}(\bar{x}) \frac{d\bar{x}^\lambda}{d\alpha} \frac{d\bar{x}^\mu}{d\alpha} \right)^{1/2} d\alpha$$

and therefore, neglecting all powers of ε higher than the first,

$$\bar{s} - s = \int_{\alpha_P}^{\alpha_Q} \left[g_{\lambda\mu} \frac{dx^\lambda}{d\alpha} \frac{dx^\mu}{d\alpha} + \varepsilon \left(\frac{\partial g_{\lambda\mu}}{dx^\sigma} \frac{dx^\lambda}{d\alpha} y^\sigma + 2 g_{\lambda\mu} \frac{dy^\lambda}{d\alpha} \right) \frac{dx^\mu}{d\alpha} \right]^{1/2} d\alpha$$
$$- \int_{\alpha_{\mathrm{P}}}^{\alpha_O} \left[g_{\lambda\mu} \frac{dx^\lambda}{d\alpha} \frac{dx^\mu}{d\alpha} \right]^{1/2} d\alpha,$$

which can be reduced to

$$\bar{s} - s = \frac{1}{2} \varepsilon \int_{\alpha_P}^{\alpha_Q} \left[\frac{\partial g_{\lambda\mu}}{dx^\sigma} \frac{dx^\lambda}{d\alpha} y^\sigma + 2 g_{\lambda\mu} \frac{dy^\lambda}{d\alpha} \right] \frac{dx^\mu}{d\alpha} \frac{d\alpha}{ds} ds.$$

Note that $ds = \sqrt{g_{\lambda\mu} dx^\lambda dx^\mu}$.

We can simplify the calculation, if $s \neq 0$, by assuming that the parameter α is identical with s itself measured along the geodesic. Then $d\alpha/ds = 1$ and

$$\bar{s} - s = \frac{1}{2} \varepsilon \int_{\alpha_P}^{\alpha_Q} \left[\frac{\partial g_{\lambda\mu}}{dx^\sigma} \frac{dx^\lambda}{d\alpha} y^\sigma + 2 g_{\lambda\mu} \frac{dy^\lambda}{d\alpha} \right] \frac{dx^\mu}{ds} ds,$$

with the x^λ, y^λ now regarded as functions of s. Integration of the second term in the last equation by parts yields

$$\bar{s} - s = \frac{1}{2} \varepsilon \int_{\alpha_P}^{\alpha_Q} \left[\frac{\partial g_{\lambda\mu}}{dx^\sigma} \frac{dx^\lambda}{ds} \frac{dx^\mu}{ds} - 2 \frac{d}{ds} \left(g_{\sigma\mu} \frac{dx^\mu}{ds} \right) \right] y^\sigma ds + \varepsilon \left(g_{\lambda\mu} \frac{dx^\mu}{ds} y^\lambda \right)_{s_P}^{s_Q}.$$

But the functions y^λ vanish at s_P and s_Q; hence the integrated term is zero. Therefore, if the interval is to have a stationary value for the geodesic curve compared with neighboring curves, $\bar{s} - s$ must be zero for any choice of the function y^λ. This is possible only if the coefficient of each y^σ in the integrand is separately zero, and therefore the differential equations of the geodesic are the n equations

$$\frac{d}{ds}\left(g_{\sigma\mu}\frac{dx^\mu}{ds}\right) - \frac{1}{2}\frac{\partial g_{\lambda\mu}}{\partial x^\sigma}\frac{dx^\lambda}{ds}\frac{dx^\mu}{ds} = 0 \quad (\sigma = 1, 2, \ldots, n). \tag{2.47}$$

Now,

$$\begin{aligned}\frac{d}{ds}\left(g_{\sigma\mu}\frac{dx^\mu}{ds}\right) &- g_{\sigma\mu}\frac{d^2x^\mu}{ds^2} + \frac{\partial g_{\mu\sigma}}{\partial x^\lambda}\frac{dx^\lambda}{ds}\frac{dx^\mu}{ds}\\ &= g_{\sigma\mu}\frac{d^2x^\mu}{ds^2} + \frac{1}{2}(g_{\mu\sigma,\lambda} + g_{\lambda\sigma,\mu})\frac{dx^\lambda}{ds}\frac{dx^\mu}{ds}.\end{aligned}$$

Thus, the equations of the geodesic may be written

$$\begin{aligned}g_{\sigma\mu}\frac{d^2x^\mu}{ds^2} &+ \frac{1}{2}(g_{\mu\sigma,\lambda} + g_{\lambda\sigma,\mu} - g_{\lambda\mu,\sigma})\frac{dx^\lambda}{ds}\frac{dx^\mu}{ds}\\ &= g_{\sigma\mu}\frac{d^2x^\mu}{ds^2} + \Gamma_{\sigma,\lambda\mu}\frac{dx^\lambda}{ds}\frac{dx^\mu}{ds} = 0.\end{aligned} \tag{2.48}$$

Multiplying this by $g^{\tau\sigma}$ and summing over σ, the result is

$$\frac{d^2x^\tau}{ds^2} + \Gamma^\mu_{\lambda\mu}\frac{dx^\lambda}{ds}\frac{dx^\mu}{ds} = 0, \quad (\tau = 1, 2, \ldots, n), \tag{2.49}$$

which is simply the standard form of (2.44) for geodesics.

The above work shows that we may use the stationary condition as the definition of a geodesic, except in dealing with the propagation of light. In that case, we have null geodesics, so the deduction as given above cannot be applied because *ds* vanish throughout.

2.7 The Curvature Tensor

In a flat space, if we perform two (ordinary) differentiations in succession their order does not matter. However, this does not, in general, hold for covariant differentiation in a curved space, except for a scalar ϕ. For the case of a scalar, we have

$$\phi_{;\mu;\nu} = (\phi_{;\mu})_{,\nu} - \Gamma^\alpha_{\mu\nu}\phi_{;\alpha} = \phi_{,\mu,\nu} - \Gamma^\alpha_{\mu\nu}\phi_{,\sigma}. \tag{2.50}$$

Since $\Gamma^\alpha_{\mu\nu}$ is symmetric in the lower indexes μ and ν, so the order of differentiation does not matter.

Now if we take a contravariant vector A^μ and apply covariant differentiations twice to it, we will find the order of differentiation is very important. First, the covariant differentiation of A^μ gives a mixed tensor

$$A^{\mu}_{;\nu} = \frac{\partial A^\mu}{\partial x^\nu} + \Gamma^{\mu}_{\alpha\nu} A^\alpha$$

Covariant differentiation of this mixed tensor gives

$$A^{\mu}_{;\nu;\beta} = \frac{\partial}{\partial x^\beta}\left(A^{\mu}_{;\nu}\right) + \Gamma^{\mu}_{\alpha\beta} A^{\alpha}_{;\nu} - \Gamma^{\alpha}_{\nu\beta} A^{\mu}_{;\alpha}$$

or

$$A^{\mu}_{;\nu;\beta} = \frac{\partial^2 A^\mu}{\partial x^\beta \partial x^\nu} + \Gamma^{\mu}_{\alpha\nu}\frac{\partial A^\alpha}{\partial x^\beta} + A^\alpha \frac{\partial \Gamma^{\mu}_{\alpha\nu}}{\partial x^\beta} + A^{\alpha}_{;\nu}\Gamma^{\mu}_{\alpha\beta} - A^{\mu}_{;\alpha}\Gamma^{\alpha}_{\beta\nu}. \tag{2.51}$$

Interchanging β and ν, we obtain

$$A^{\mu}_{;\beta;\nu} = \frac{\partial^2 A^\mu}{\partial x^\nu \partial x^\beta} + \Gamma^{\mu}_{\alpha\beta}\frac{\partial A^\alpha}{\partial x^\nu} + A^\alpha \frac{\partial \Gamma^{\mu}_{\alpha\beta}}{\partial x^\nu} + A^{\alpha}_{;\beta}\Gamma^{\mu}_{\alpha\nu} - A^{\mu}_{;\alpha}\Gamma^{\alpha}_{\beta\nu} \tag{2.52}$$

Subtracting (2.52) from (2.51), we get

$$A^{\mu}_{;\nu;\beta} - A^{\mu}_{;\beta;\nu} = A^\alpha\left[\frac{\partial \Gamma^{\mu}_{\alpha\nu}}{\partial x^\beta} - \frac{\partial \Gamma^{\mu}_{\alpha\beta}}{\partial x^\nu} + \Gamma^{\gamma}_{\alpha\nu}\Gamma^{\mu}_{\gamma\beta} - \Gamma^{\gamma}_{\alpha\beta}\Gamma^{\mu}_{\gamma\nu}\right] = A^\alpha R^{\mu}_{\alpha\nu\beta} \tag{2.53}$$

where

$$R^{\mu}_{\alpha\nu\beta} = \frac{\partial \Gamma^{\mu}_{\alpha\nu}}{\partial x^\beta} - \frac{\partial \Gamma^{\mu}_{\alpha\beta}}{\partial x^\nu} + \Gamma^{\gamma}_{\alpha\nu}\Gamma^{\mu}_{\gamma\beta} - \Gamma^{\gamma}_{\alpha\beta}\Gamma^{\mu}_{\gamma\nu}. \tag{2.54}$$

Since $A^{\mu}_{;\nu;\beta} - A^{\mu}_{;\beta;\nu}$ and A^α are tensors, $R^{\mu}_{\alpha\nu\beta}$ must be the component of a tensor, by the quotient law. It is called the *curvature tensor* or the *Riemann tensor*, and it depends solely on the Christoffel symbols and their derivatives. In flat space all $g_{\mu\nu}$ can be transformed into constants (rectangular or Galilean coordinates) and all Christoffel symbols vanish; hence $R^{\mu}_{\alpha\nu\beta} = 0$. Being a tensor equation, it holds in all coordinate systems (Cartesian, oblique, or curvilinear). In a curved space, $R^{\mu}_{\alpha\nu\beta}$ will not vanish. Therefore it is a measure of the curvature of space.

From (2.54) it follows that the curvature tensor is antisymmetric in the indices ν and β:

$$R^{\mu}_{\alpha\nu\beta} = -R^{\mu}_{\alpha\beta\nu}. \tag{2.55}$$

Furthermore, it is easy to verify that the following identity is valid

$$R^{\alpha}_{\beta\gamma\delta} + R^{\alpha}_{\delta\beta\gamma} + R^{\alpha}_{\gamma\delta\beta} = 0. \tag{2.56}$$

In addition to the mixed curvature tensor $R^{\mu}_{\alpha\nu\beta}$, we can also use the covariant curvature tensor

$$R_{\alpha\beta\gamma\delta} = g_{\alpha\eta} R^{\eta}_{\beta\gamma\delta}. \tag{2.57}$$

From the transformation law for $R_{\alpha\beta\gamma\delta}$

$$R_{\alpha\beta\gamma\delta} = \frac{1}{2}\left(\frac{\partial^2 g_{\alpha\delta}}{\partial x^\beta \partial x^\gamma} + \frac{\partial^2 g_{\beta\gamma}}{\partial x^\alpha \partial x^\delta} - \frac{\partial^2 g_{\alpha\gamma}}{\partial x^\beta \partial x^\delta} - \frac{\partial^2 g_{\beta\delta}}{\partial x^\alpha \partial x^\gamma}\right)$$
$$+ g_{\mu\nu}\left(\Gamma^\mu_{\beta\gamma}\Gamma^\nu_{\alpha\delta} - \Gamma^\mu_{\beta\delta}\Gamma^\nu_{\alpha\gamma}\right). \tag{2.58}$$

It is not difficult to derive this transformation law, but it is very tedious; hence we simply give it here without derivation. From (2.58) we see the following symmetry properties:

$$R_{\alpha\beta\gamma\delta} = -R_{\beta\alpha\gamma\delta} = -R_{\alpha\beta\delta\gamma}\,,\; R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta} \tag{2.59}$$

i.e., the tensor is antisymmetric in each of the index pairs α, β, and γ, δ and is symmetric under the interchange of the two pairs with each other. Thus, all components $R_{\alpha\beta\gamma\delta}$, in which $\alpha = \beta$ or $\gamma = \delta$ are zero. As for $R_{\alpha\beta\gamma\delta}$ we also have the identity

$$R_{\alpha\beta\gamma\delta} + R_{\alpha\delta\beta\gamma} + R_{\alpha\gamma\delta\beta} = 0. \tag{2.60}$$

A fourth-rank tensor has $4^4 = 256$ components. However, because of the above symmetries, the number of algebraically independent components of $R_{\alpha\beta\gamma\delta}$ is only 20.

By contracting the curvature tensor, we get the symmetric Ricci tensor:

$$R_{\sigma\mu} = R^\lambda_{\sigma\mu\lambda} = R_{\mu\sigma}. \tag{2.61}$$

According to (2.54), we have

$$R_{\sigma\mu} = \frac{\partial \Gamma^\lambda_{\sigma\mu}}{\partial x^\lambda} - \frac{\partial \Gamma^\lambda_{\sigma\lambda}}{\partial x^\mu} + \Gamma^\lambda_{\sigma\mu}\Gamma^\nu_{\lambda\nu} - \Gamma^\nu_{\sigma\lambda}\Gamma^\lambda_{\mu\nu}. \tag{2.62}$$

This tensor is clearly symmetric: $R_{\sigma\mu} = R_{\mu\sigma}$.

Finally, contracting $R_{\sigma\mu}$ we obtain

$$R = g^{\sigma\mu} R_{\sigma\mu} = g^{\sigma\lambda} g^{\mu\beta} R_{\sigma\mu\lambda\beta}, \tag{2.63}$$

which is called the *scalar curvature of the space*.

One of the most important tensors in the study of gravitation is the Einstein tensor, defined by

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = G_{\nu\mu}. \tag{2.64}$$

The Einstein tensor $G_{\mu\nu}$ is purely geometric in character, being built up from $g_{\mu\nu}$ and their first and second derivatives. And it is linear in the second derivatives of $g_{\mu\nu}$. It can be shown that the covariant divergence of $G_{\mu\nu}$ vanishes identically (we leave it as a problem). This property will be used later to formulate the gravitational field equations.

There is apparent obscurity surrounding the physical meaning of the Riemann tensor; a simple example of calculating the curvature of space doesn't elucidate the physical meaning of the Riemann tensor. This may be helpful to visualize the Riemann tensor. Let us consider the two-dimensional space on the surface of a sphere of radius a. The metric of this two-dimensional space is

$$ds^2 = a^2(d\theta^2 + \sin^2\theta d\phi^2).$$

The covariant and contravariant metric tensors are

$$g_{\mu\nu} = \begin{pmatrix} a^2 & 0 \\ 0 & a^2\sin^2\theta \end{pmatrix}, \quad g^{\mu\nu} = \begin{pmatrix} 1/a^2 & 0 \\ 0 & 1/(a^2\sin^2\theta) \end{pmatrix}$$

and $g = a^4\sin^2\theta$.

The Christoffel symbols are given by (2.42). By direction calculation, we find that the only non-zero symbols are

$$\Gamma^{\theta}_{\phi\phi} = \frac{1}{2}g^{\theta\alpha}\left(\frac{\partial g_{\phi\alpha}}{\partial\phi} + \frac{\partial g_{\alpha\phi}}{\partial\phi} - \frac{\partial g_{\phi\phi}}{\partial\alpha}\right), \text{ and } \Gamma^{\phi}_{\theta\phi} = \frac{1}{2}g^{\phi\alpha}\left(\frac{\partial g_{\theta\alpha}}{\partial\phi} + \frac{\partial g_{\phi\alpha}}{\partial\theta} - \frac{\partial g_{\theta\phi}}{\partial\alpha}\right)$$

where α takes the values θ and ϕ; these become

$$\Gamma^{\theta}_{\phi\phi} = -g^{\theta\theta}\frac{\partial g_{\phi\phi}}{\partial\theta} = -\sin\theta\cos\theta, \quad \Gamma^{\phi}_{\theta\phi} = \frac{1}{2}g^{\phi\phi}\frac{\partial g_{\phi\phi}}{\partial\theta} = \cot\theta.$$

Let us first calculate the Ricci tensor (the contracted Riemann tensor) that is given by (2.62). The non-zero components are

$$R_{\theta\theta} = \Gamma^{\phi}_{\theta\phi}\Gamma^{\phi}_{\theta\phi} + \frac{\partial\cot\theta}{\partial\theta} = \cot^2\theta - \frac{1}{\sin^2\theta} = -1$$

and

$$R_{\phi\phi} = -\sin^2\theta.$$

The Riemann scalar curvature R of the space is given by

$$R = g^{\theta\theta}R_{\theta\theta} + g^{\phi\phi}R_{\phi\phi} = -2/a^2.$$

If the reader is familiar with the analytical geometry of surfaces, then recall that R here is equivalent to

$$R = \frac{2}{\rho_1\rho_2}$$

where ρ_1 and ρ_2 are the two principal radii of curvature of the surfaces. On the sphere these radii coincide and are equal to the spherical radius. The Riemann scalar curvature thus bears a simple relationship to the radius curvature of the two-dimensional space on the spherical surface.

Proceeding similarly, we find, from (2.57), the only independent component of the Riemann tensor

$$R_{\theta\phi\theta\phi} = g_{\theta\theta}R^{\theta}_{\phi\theta\phi} = g_{\theta\theta}\left(\frac{\partial\Gamma^{\theta}_{\phi\phi}}{\partial\theta} - \frac{\partial\Gamma^{\theta}_{\theta\phi}}{\partial\phi} + \Gamma^{\theta}_{\theta\alpha}\Gamma^{\alpha}_{\phi\phi} - \Gamma^{\theta}_{\phi\alpha}\Gamma^{\alpha}_{\phi\theta}\right) = a^2\sin^2\theta,$$

and by (2.59) we have other non-zero components

$$R_{\phi\theta\phi\theta} = -R_{\theta\phi\phi\theta} = -R_{\phi\theta\theta\phi} = R_{\theta\phi\theta\phi} = a^2\sin^2\theta.$$

2.8 The Condition for Flat Space

If space is flat, we may choose a Cartesian coordinate system. Then the $g_{\mu\nu}$ are all constant and all Christoffel symbols vanish; hence, from (2.54),

$$R^{\mu}_{\alpha\nu\beta} = 0$$

in this coordinates system. But if $R^{\mu}_{\alpha\nu\beta} = 0$ in one coordinate system, the components are zero in all coordinate systems. Hence if a space is flat, the Riemann curvature tensor must vanish, which is a necessary condition for a space to be flat. The converse is a sufficient condition: if $R^{\mu}_{\alpha\nu\beta} = 0$, the space is flat. We now proceed to prove it.

If $R^{\mu}_{\alpha\nu\beta} = 0$, and if we can find a coordinate system for which its metric tensor is constant, then the space is flat. To this end, let us take vector A_μ located at Point x and parallel displace it to Point $x + dx$, and then parallel displace it to Point $x + dx + \delta x$. If the Riemann curvature tensor $R^{\mu}_{\alpha\nu\beta}$ is zero, the result would be the same if we had displace A_μ from x to $x + \delta x$ and then to $x + \delta x + dx$. That is, the displacement is independent of the path. Thus we can displace the vector to a distant point and the result we get is independent of the path to the distant point. If we displace the vector A_μ at x to all points by parallel displacement, we will get a vector field that satisfies $A_{\mu;\nu} = 0$:

$$A_{\mu;\nu} = \frac{\partial A_\mu}{\partial x^\nu} - \Gamma^{\sigma}_{\mu\nu} A_\sigma = 0, \quad \text{or} \quad A_{\mu,\nu} = \Gamma^{\sigma}_{\mu\nu} A_\sigma.$$

If A_μ is the gradient of a scalar S, $A_\mu = \partial S/\partial x^\mu = S_{,\mu}$, then the above equation becomes

$$S_{,\mu\nu} = \Gamma^{\sigma}{}_{\mu\nu} S_{,\sigma}.$$

Because $\Gamma^{\sigma}{}_{\mu\nu} = \Gamma^{\sigma}{}_{\nu\mu}$, $S_{,\mu\nu} = S_{,\nu\mu}$ the above equations can be integrated.

Now let us take four independent scalars satisfying the last equations and let them to be the coordinates y^α of a new coordinate system. Thus, we have

$$y^{\alpha}_{,\mu\nu} = \Gamma^{\sigma}{}_{\mu\nu} y_{,\sigma},$$

where $y^{\alpha}_{,\mu\nu} = \partial^2 y^\alpha / \partial x^\mu \partial x^\nu$.

Let us go back to Eq. (10) that now takes the form

$$g_{\mu\lambda}(x) = g_{\alpha\beta}(y) \frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial y^\beta}{\partial x^\lambda}.$$

Differentiation of this equation with respect to x^ν yields

$$\begin{aligned}
\frac{\partial g_{\mu\lambda}}{\partial x^\nu} &= \frac{\partial g_{\alpha\beta}(y)}{\partial x^\nu} \frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial y^\beta}{\partial x^\lambda} + g_{\alpha\beta}(y) \left(\frac{\partial^2 y^\alpha}{\partial x^\mu \partial x^\nu} \frac{\partial y^\beta}{\partial x^\lambda} + \frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial^2 y^\beta}{\partial x^\lambda \partial x^\nu} \right) \\
&= \frac{\partial g_{\alpha\beta}(y)}{\partial x^\nu} \frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial y^\beta}{\partial x^\lambda} + g_{\alpha\beta}(y) \left(\Gamma^{\sigma}_{\mu\nu} y^{\alpha}{}_{,\sigma} y^{\beta}{}_{,\lambda} + y^{\alpha}{}_{,\mu} \Gamma^{\sigma}_{\lambda\nu} y^{\beta}{}_{,\sigma} \right) \\
&= \frac{\partial g_{\alpha\beta}(y)}{\partial x^\nu} \frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial y^\beta}{\partial x^\lambda} + g_{\alpha\lambda}(x) \Gamma^{\sigma}_{\mu\nu} + g_{\mu\sigma}(x) \Gamma^{\sigma}_{\lambda\nu}.
\end{aligned}$$

The last two terms can be rewritten in terms of the Christoffel symbols of the first kind:

$$g_{\alpha\lambda}(x)\Gamma^{\sigma}_{\mu\nu} + g_{\mu\sigma}(x)\Gamma^{\sigma}_{\lambda\nu} = \Gamma_{\lambda,\mu\nu} + \Gamma_{\mu,\lambda\nu} = \partial g_{\mu\lambda}/\partial x^{\nu} = g_{\mu\lambda,\nu}$$

where the Christoffel symbols are given by (2.40). Combining this with the last equation, we obtain

$$\frac{\partial g_{\alpha\beta}(y)}{\partial x^{\nu}}\frac{\partial y^{\alpha}}{\partial x^{\mu}}\frac{\partial y^{\beta}}{\partial x^{\lambda}} = 0.$$

It follows that

$$\frac{\partial g_{\alpha\beta}(y)}{\partial x^{\nu}} = 0.$$

Thus, the metric tensor of the new system of coordinates is constant, and the Christroffel symbols all vanish identically. In other words, we have flat space.

2.9 Geodesic Deviation

We can get a good insight into the nature of a space just by examining the problem of geodesic deviation. For example, consider two nearby freely falling particles which travel on paths $x^{\mu}(\tau)$ and $x^{\mu}(\tau) + \delta x^{\mu}(\tau)$. The equations of motion are then given by

$$\frac{d^2x^{\mu}}{d\tau^2} + \Gamma^{\mu}_{\nu\lambda}(x)\frac{dx^{\nu}}{d\tau}\frac{dx^{\lambda}}{d\tau} = 0 \tag{2.65}$$

and

$$\frac{d^2}{d\tau^2}[x^{\mu} + \delta x^{\mu}] + \Gamma^{\mu}_{\nu\lambda}(x + \delta x)\frac{d}{d\tau}[x^{\nu} + \delta x^{\nu}]\frac{d}{d\tau}[x^{\lambda} + \delta x^{\lambda}] = 0. \tag{2.66}$$

Evaluating the difference between these equations to first order in δx^{μ} gives

$$\frac{d^2\delta x^{\mu}}{d\tau^2} + \frac{\partial\Gamma^{\mu}_{\nu\lambda}}{\partial x^{\rho}}\delta x^{\rho}\frac{dx^{\nu}}{d\tau}\frac{dx^{\lambda}}{d\tau} + 2\Gamma^{\mu}_{\nu\lambda}\frac{dx^{\nu}}{d\tau}\frac{dx^{\rho}}{d\tau} = 0 \tag{2.67}$$

or, in terms of covariant derivatives along the curves $x^{\mu}(\tau)$,

$$\frac{D^2}{D\tau^2}\delta x^{\lambda} = R^{\lambda}_{\nu\mu\rho}\delta x^{\mu}\frac{dx^{\nu}}{d\tau}\frac{dx^{\rho}}{d\tau}. \tag{2.68}$$

This (2.68) is called the *equation of geodesic deviation*.

Although a freely falling particle appears to be at rest in a coordinate system falling with the particle, a pair of nearby freely falling particles will exhibit a relative motion that can reveal the presence of a gravitational field to an observer who falls with the particles. The effect of the right-hand side of (2.68) becomes negligible when the separation between particles is much less than the characteristic dimensions of the field. This indicates clearly that the local inertial frames are only locally applicable; otherwise, the principle of equivalence will be violated.

2.10 Laws of Physics in Curved Spaces

The laws of physics must be valid in all coordinate systems. If they are expressed as tensor equations, whenever they involve the derivative of a field quantity, it must be a covariant derivative. Even if we are working with flat space (which means neglecting the gravitational field) and we are using curvilinear coordinates, we must write our equations in terms of covariant derivatives if we want them to hold in all coordinate systems.

As for the problem of generalizing a particular physics law from the flat Minkowski space to a general curved space, there is not a unique solution at all. In fact, the problem is highly complicated, since in general there will be an interaction between the space (the $g_{\mu\nu}$) and the physical phenomenon whose laws we are trying to formulate. But if the object under consideration does not appreciably influence the $g_{\mu\nu}$, that is, if the $g_{\mu\nu}$ are determined by objects much more massive than the object under consideration, we may then consider the $g_{\mu\nu}$ as given functions of the spacetime variables, $g_{\mu\nu}(x^\sigma)$. In this case the geometry is rigidly determined and the effect of the physical object under study on the geometrical structure may be neglected. Under these circumstances, we may take over the special relativistic laws by substituting

$$d \to D, \partial_\mu \to D_\mu, d\Omega \to \sqrt{-g}d\Omega$$

where

$$\partial_\mu A^\nu = \frac{\partial A^\nu}{\partial x^\mu}, D_\mu A = A^\mu{}_{;\nu}.$$

As an example, consider the motion of a free particle. Its time track in Minkowski space is characterized by the equations

$$d^2x^{\mathrm{i}}/ds^2 = 0, i = 0, 1, 2, 3. \tag{2.69}$$

These equations imply a straight line in the four-dimensional Minkowski space, which in turn corresponds to a uniform rectilinear motion in three-dimensional space. These equations can be derived from the stationary condition that the integral $\int$ ds, taken along the motion between two points P and Q is stationary if one makes a small variation of the path keeping the end points fixed:

$$\delta \int_P^Q ds = 0 \tag{2.70}$$

where the variation vanishes at the end points P and Q.

If the particle is subject to the action of a gravitational field, its equation of motion is no longer a straight line, because the spacetime is curved. But the particle still follows a stationary trajectory, the geodesic. As shown above, the geodesic equation can be obtained from the same variational principle (2.70), provided that the Minkowski metric is replaced by the curved space metric $g_{\mu\nu}$ and $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$. Instead of computing explicitly the variation in (2.70), we can simply obtain the geodesic equations as the covariant generalization of (2.69):

$$D^2x^{\mathrm{i}}/D\mathrm{s}^2 = 0,$$

which is equivalent to

$$\mathrm{d}^2x^{\mathrm{i}}/\mathrm{ds}^2 + \Gamma^{\alpha}_{\mu\nu}(dx^{\mu}/ds)(dx^{\nu}/ds) = 0,$$

the geodesic equations.

2.11 The Metric Tensor and the Classical Gravitational Potential

The presence of a gravitational field modifies the structure of spacetime. Any gravitational field is just a change in the metric of space-time, as determined by the metric tensor $g_{\mu\nu}$. Through the geodesic equations of motion, we can now provide the expressions governing the union of geometry and gravitation. To this end, let us compare the Newtonian equation of motion of a particle in a gravitational field and its geodesic equations of motion in a curved-space geometry:

$$d^2x^{\alpha}/ds^2 + \partial\phi/\partial x^{\alpha} = 0 \tag{2.71}$$

$$d^2x^{\alpha}/ds^2 + \Gamma^{\alpha}{}_{\mu\nu}(dx^{\mu}/ds)(dx^{\nu}/ds) = 0 \tag{2.72}$$

where ϕ is the Newtonian gravitational potential. These two equations have a fundamental similarity in that both are independent of the mass of the moving body under consideration. Thus, both equations satisfy the principle of equivalence. Now since the derivative $\mathrm{d}^2x^{\alpha}/ds^2$ is the four-acceleration of the particle, the quantity $-m\Gamma^{\alpha}{}_{\mu\nu}\, u^{\mu}u^{\nu}$ may be interpreted as the gravitational force, and then the components of the metric tensor $\mathrm{g}_{\mu\nu}$ play the role of the Newtonian gravitational potential ϕ (as the Christoffel symbols are constructed from the derivatives of the $g_{\mu\nu}$). We must first show that this interpretation is consistent with the Newtonian equations of motion; namely, we must show that in the limit of ordinary velocities the geodesic equations reduce to the Newtonian equations. To see this, let the velocity $dx^{\alpha}/dt \ll c$. Then $ds^2 = g_{00}c^2dt^2$, and $ds = \sqrt{g_{00}}cdt$, so that

$$d^2x^{\alpha}/dt^2 + \Gamma^{\alpha}{}_{00}c^2 = 0$$

where

$$\Gamma^{\alpha}{}_{00} = \partial g_{00}/\partial x^{\alpha}.$$

From this we see that in this limit $g_{00} = K + 2\phi/c^2$. Since in flat space $g_{00} = 1$, we have $K = 1$ and

$$\mathrm{g}_{00} = 1 + 2\phi/c^2 \tag{2.73}$$

This shows that the identification postulated above is plausible, i.e., the metric tensor $\mathrm{g}_{\mu\nu}$ plays the role of Newtonian gravitational potential.

We should be careful to note that the physical content of the two equations, (2.71 and 2.72), are entirely different. In Newton's equation we have a field ϕ, which causes the motion. The particle is under a force and its velocity changes in time. In the geodesic equations, on the other hand, there is no physical agent such as ϕ. The particle follows a geodesic that is determined by the geometry of the space-time. This change in interpretation is actually a conceptual simplification, since inertia and gravitation are unified and the concept of external force is eliminated from the theory of gravitation.

2.12 Some Useful Calculation Tools

We conclude this chapter by giving a number of useful aids in manipulation of tensor quantities. First of all, it is very helpful to bear in mind that the covariant derivative of the metric tensor $g_{\mu\nu}$ is zero. That is, the metric tensor $g_{\mu\nu}$ may be considered as a constant during covariant differentiation.

We now derive an expression for the contracted Christoffel symbol $\Gamma^{\mu}{}_{\alpha\mu}$ that will be very useful later on. From (2.39) we have

$$\Gamma^{\mu}_{\alpha\mu} = \frac{1}{2} g^{\mu\nu} \left(\frac{\partial g_{\nu\alpha}}{\partial x^{\mu}} + \frac{\partial g_{\nu\mu}}{\partial x^{\alpha}} - \frac{\partial g_{\alpha\mu}}{\partial x^{\nu}} \right).$$

Changing the positions of the dummy indexes μ and ν in the first term and remembering $g_{\mu\nu} = g_{\nu\mu}$, we see that the first and third terms then cancel each other, so that

$$\Gamma^{\mu}_{\alpha\mu} = \frac{1}{2} g^{\mu\nu} \frac{\partial g_{\mu\nu}}{\partial x^{\alpha}}, \tag{2.74}$$

which can be simplified. To do this we calculate the differential dg of the determinant g made up from the components of the metric tensor $g_{\mu\nu}$; dg can be obtained by taking the differential of each component of the tensor $g_{\mu\nu}$ and multiplying it by its coefficient in the determinant, i.e., by the corresponding minor:

$$dg = dg_{\mu\nu} M^{\mu\nu}$$

where $M^{\mu\nu}$ is the minor of the component $g_{\mu\nu}$. Now,

$$g^{\mu\nu} = \frac{M^{\mu\nu}}{g}, \quad M^{\mu\nu} = g^{\mu\nu} g$$

Thus,

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

The expression on the far right of the above equation follows from

$$d(g_{\mu\nu} g^{\mu\nu}) = d(\delta^{\mu}_{\mu}) = d(4) = 0.$$

We then have

$$\frac{\partial g}{\partial x^{\alpha}} = g\, g^{\mu\nu} \frac{\partial g_{\mu\nu}}{\partial x^{\alpha}} = -g\, g_{\mu\nu} \frac{\partial g^{\mu\nu}}{\partial x^{\alpha}}. \tag{2.75}$$

The use of (2.75) enables us to write (2.74) in the form

$$\Gamma^{\mu}_{\alpha\mu} = \frac{1}{2} g^{\mu\nu} \frac{\partial g_{\mu\nu}}{\partial x^{\alpha}} = \frac{1}{2g} \frac{\partial g}{\partial x^{\alpha}} = \frac{\partial \ln \sqrt{-g}}{\partial x^{\alpha}}. \tag{2.76}$$

This expression is very useful. First consider the covariant divergence $A^{\mu}{}_{;\mu}$:

$$A^{\mu}{}_{;\mu} = \frac{\partial A^{\mu}}{\partial x^{\mu}} + \Gamma^{\mu}_{\alpha\mu} A^{\alpha}.$$

Substituting the expression (2.76) for $\Gamma^{\mu}_{\alpha\mu}$, we obtain

$$A^{\mu}_{;\mu} = \frac{\partial A^{\mu}}{\partial x^{\mu}} + A^{\alpha} \frac{\partial \ln \sqrt{-g}}{\partial x^{\alpha}} = \frac{1}{\sqrt{-g}} \frac{\partial \left(\sqrt{-g} A^{\mu}\right)}{\partial x^{\mu}}. \tag{2.77}$$

We now consider the covariant divergence of a contravariant tensor of the second rank $T^{\mu\nu}_{;\alpha}$. From (2.27), we have

$$T^{\mu\nu}{}_{;\alpha} = \frac{\partial T^{\mu\nu}}{\partial x^{\nu}} + \Gamma^{\mu}_{\beta\nu} T^{\beta\nu} + \Gamma^{\mu}_{\beta\alpha} T^{\mu\beta}.$$

Changing the positions of the dummy indexes μ and β in the third term on the right-hand side, we obtain

$$T^{\mu\nu}{}_{;\nu} = \frac{\partial T^{\mu\nu}}{\partial x^{\nu}} + \Gamma^{\mu}_{\beta\nu} T^{\beta\nu} + \Gamma^{\beta}_{\nu\beta} T^{\mu\nu}.$$

Substituting the expression (2.76) for $\Gamma^{\beta}{}_{\nu\beta}$, we obtain

$$T^{\mu\nu}{}_{;\nu} = \frac{\partial T^{\mu\nu}}{\partial x^{\nu}} + \Gamma^{\mu}{}_{\beta\nu} T^{\beta\nu} + \frac{\partial \ln \sqrt{-g}}{\partial x^{\nu}} T^{\mu\nu}$$

or

$$T^{\mu\nu}{}_{;\nu} = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^{\nu}} (\sqrt{-g} T^{\mu\nu}) + \Gamma^{\mu}{}_{\beta\nu} T^{\beta\nu}. \tag{2.78}$$

Similarly, for a mixed tensor, (2.28) leads to

$$T^{\beta}_{\alpha;\beta} = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^{\beta}} \left(T^{\beta}_{\alpha} \sqrt{-g}\right) - \Gamma^{\beta}_{\alpha\sigma} T^{\alpha}_{\beta}. \tag{2.79}$$

For an antisymmetric tensor $F^{\beta\nu} = -F^{\nu\beta}$, then

$$\Gamma^{\mu}_{\beta\nu} F^{\beta\nu} = \Gamma^{\mu}_{\nu\beta} F^{\nu\beta} = -\Gamma^{\mu}_{\beta\nu} F^{\beta\mu}. \tag{2.80}$$

The expression in the middle is obtained by interchanging the dummy indexes β and ν, and the expression on the far right follows from the $F^{\beta\nu} = -F^{\nu\beta}$, and $\Gamma^{\mu}{}_{\nu\beta} = \Gamma^{\mu}{}_{\beta\nu}$. From (2.80) it follows that

$$\Gamma^{\mu}_{\beta\nu} F^{\beta\mu} = 0.$$

Thus, for an antisymmetric tensor $F^{\alpha\beta}$, the last term of (2.78) vanishes and the covariant divergence is

$$F^{\alpha\beta}_{;\beta} = \frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^{\beta}}(F^{\alpha\beta}\sqrt{-g}). \tag{2.81}$$

For a symmetric tensor $S^{\alpha\beta}$, rearrangement of the last term of (2.79) gives

$$S^{\beta}_{\alpha;\beta} = \frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^{\beta}}(S^{\beta}_{\alpha}\sqrt{-g}) - \frac{1}{2}\frac{\partial g_{\beta\nu}}{\partial x^{\alpha}}S^{\beta\nu}. \tag{2.82}$$

In Cartesian coordinates, the curl

$$\frac{\partial A_{\mu}}{\partial x^{\nu}} - \frac{\partial A_{\nu}}{\partial x^{\mu}}$$

is an antisymmetric tensor. In curvilinear coordinates this tensor is $A_{\mu;\nu} - A_{\nu;\mu}$:

$$A_{\mu;\nu} - A_{\nu;\mu} = A_{\mu,\nu} - \Gamma^{\rho}_{\mu\nu}A_{\rho} - (A_{\nu,\mu} - \Gamma^{\rho}_{\nu\mu}A_{\rho})$$

Since $\Gamma^{\rho}{}_{\mu\nu} = \Gamma^{\rho}{}_{\nu\mu}$, we have

$$A_{\mu;\nu} - A_{\nu;\mu} = \frac{\partial A_{\mu}}{\partial x^{\nu}} - \frac{\partial A_{\nu}}{\partial x^{\mu}}. \tag{2.83}$$

This result may be stated: covariant curl equals ordinary curl, but it holds only for a covariant vector. For a contravariant vector we could not form the curl because the suffixes would not balance.

Finally, we transform to curvilinear coordinates the sum

$$\frac{\partial^2\varphi}{\partial x_{\alpha}\partial x^{\alpha}}$$

of the second derivatives of a scalar φ. In curvilinear coordinates this sum goes over to $\varphi^{;\alpha}_{;\alpha}$. But covariant differentiation of a scalar reduces to ordinary differentiation:

$$\varphi_{;\alpha} = \partial\varphi/\partial x^{\alpha}.$$

Raising the index α, we have

$$\varphi^{;\alpha} = g^{\alpha\beta}\partial\varphi/\partial x^{\beta},$$

which is a contravariant tensor of rank one. Using formula (2.77), we find

$$\varphi^{;\alpha}_{;\alpha} = \frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^{\alpha}}\left(\sqrt{-g}\, g^{\alpha\beta}\frac{\partial\varphi}{\partial x^{\beta}}\right). \tag{2.84}$$

In view of Eq. (77), Gauss' theorem for transformation of the integral of a vector over a hypersurface into an integral over a four-volume can now be written as

$$\oint A^{\alpha}\sqrt{-g}\,dS_{\alpha} = \int A^{\alpha}{}_{;\alpha}\sqrt{-g}\,d\Omega. \tag{2.85}$$

2.13 Problems

2.1. Write the terms in each of the following indicated sums

$$a\,.\,a_{jk}x^k \quad b\,.\,A_{pq}B^{qr} \quad c.\overline{g}_{rs} = g_{jk}\frac{\partial x^j}{\partial \overline{x}^r}\frac{\partial x^k}{\partial \overline{x}^s}$$

2.2. Write the transformation law for the following tensors: $a\,.\,A^i{}_{jk} \quad b\,.\,B^{pq}{}_{ijk}$

2.3. A quantity $A(j,k,m,n)$, which is a function of the coordinates x^i, transforms to another coordinate system $\bar{x}^i$ according to the rule

$$\overline{A}(p,q,r,s) = \frac{\partial x^j}{\partial \overline{x}^p}\frac{\partial \overline{x}^q}{\partial x^k}\frac{\partial \overline{x}^r}{\partial x^m}\frac{\partial \overline{x}^s}{\partial x^n}A(j,k,m,n).$$

Is the quantity a tensor? If so, write the tensor in suitable notation and give the covariant and contravariant rank.

2.4. Show that the property of symmetry (or antisymmetry) with respect to indexes of a tensor is invariant under coordinate transformation.

2.5. A covariant tensor has components xy, $2y - z^2$, xz in rectangular coordinates; find its covariant components in spherical coordinates.

2.6. Prove that the contraction of the outer product of the two tensors A^p and B_p is a scalar.

2.7. Determine the Christoffel symbols of the second kind in rectangular and cylindrical coordinates.

2.8. The line element on the surface of a sphere of radius a in Euclidean space is given by $ds^2 = a^2(d\theta^2 + \sin^2\theta d\phi^2)$. For this space calculate $\Gamma^i{}_{kl}, i, k, l = 1, 2$ (with $\theta = x^1, \phi = x^2$).

2.9. Find the covariant derivative of $A^i{}_j B^{km}{}_n$ with respect to x^q.

2.10. Prove that

$$\nabla \cdot A^p = \frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^k}\left(\sqrt{-g}A^k\right) \quad \text{and} \quad \nabla^2\phi = \frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^k}\left(\sqrt{-g}g^{kr}\frac{\partial \phi}{\partial x^r}\right).$$

2.11. Determine the force acting on a particle in a constant gravitational field.

2.12. Prove that the covariant divergence of the Einstein tensor vanishes.

2.13. The distance s between two points on a curve $x^\mu = x^\mu(\lambda)$ is given by

$$s = \int ds = \int_1^2 \sqrt{g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}}.$$

Show that the necessary condition that s be an extremum is that

$$\frac{\partial L}{\partial x^\nu} - \frac{d}{d\lambda}\frac{\partial L}{\partial \dot{x}^\nu} = 0$$

where

$$L = g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}, \text{ and } \dot{x}^\nu = \frac{dx^\nu}{d\lambda}.$$

2.14. Show that great circles drawn on the surface of a sphere are geodesics.

2.15. Show that, with (2.43), in a Euclidean space geodesics are straight lines.

References

Dirac PAM (1975) *General Theory of Relativity.* (John Wiley & Sons, New York)
Landau LD, Lifshitz EM (1975) *The Classical Theory of Fields.* (Pergamon Press, Oxford UK)