2

The spacetime of general relativity and paths of particles

2.0 Introduction

Einstein's general theory of relativity postulates that gravitational effects may be explained by the curvature of spacetime (modeled by a four-dimensional pseudo-Riemannian manifold) and that gravity should not be regarded as a force in the conventional sense. To get a preliminary idea of what is involved, we shall follow the practice of a number of authors[1] and consider ants crawling over a curved surface, namely the skin of an apple.

Suppose then that an ant wishes to follow a straight path on the apple's skin. *The straightest path it could take would be achieved by its making its left-hand paces equal to its right-hand ones.* This would clearly generate a straight-line path if it were crawling on a plane, so it is natural to adopt a path generated on a curved surface in this way as the analog of a straight line. These paths are called *geodesics*. If the ant had inky feet, so that it left footprints, then making cuts along the left-hand and right-hand tracks would yield a thin strip of peel which could be removed. If this thin strip were laid flat on a plane it would be straight, confirming that a geodesic, as we have defined it, is the analog of a straight line.

Suppose now that we have several ants crawling over the apple (without colliding) and each follows a geodesic path, leaving a record of its progress on the apple's skin. (A single track rather than a double one: ink on the tip of its abdomen, rather than inky feet.) If we concentrate on a portion of the apple's skin *which is so small that it may be considered flat*, then the tracks of the ants would appear as straight lines on this "flat" portion (see Fig. 2.1). If, however, we take a larger view of things, then the picture is different. For example, suppose two ants leave from nearby points on a starting line at the same time, and move with the same constant speed, following geodesics which are initially perpendicular to the starting line (see Fig. 2.2). Their paths

[1]Notably Misner, Thorne, and Wheeler, on whose well-known illustration our Fig. 2.1 is based. See Misner, Thorne, and Wheeler, 1973, §1.1.

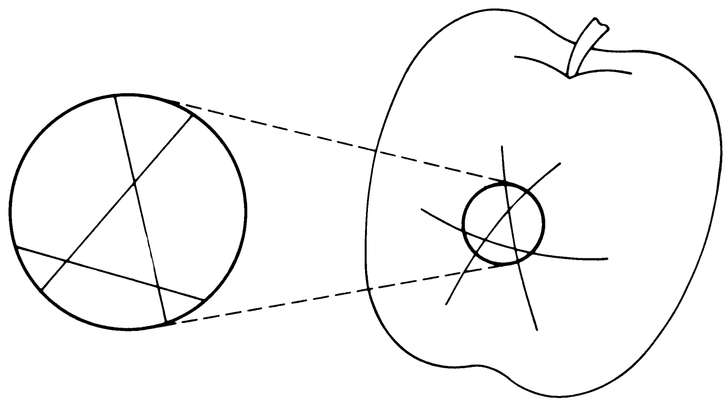

Fig. 2.1. Small portions of an apple's skin may be regarded as flat.

would initially be parallel, but because of the curvature of the apple's skin, they would start to converge. That is, their separation d does not remain constant. More generally, we can see that the relative acceleration of ants which follow neighboring geodesics with constant (but not necessarily equal) speeds is non-zero, if the surface over which they are crawling is curved. In this way, curvature may be detected implicitly by what is called *geodesic deviation.*

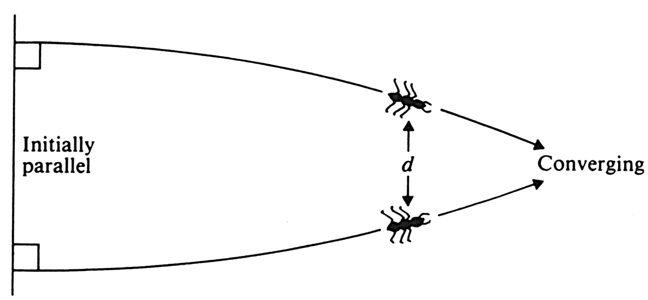

Fig. 2.2. Converging geodesics on an apple's skin.

An apple is not a perfect sphere: there is a dimple caused by the stalk. Should an ant pass near the stalk its geodesic path would suffer a marked deflection, like that of a comet passing near the Sun, and it would look as if the stalk attracted the ant. However, this is not the correct interpretation. The stalk modifies the curvature of the apple's skin in its vicinity, and this produces geodesics which give the effect of an attraction by the stalk.

This allegory may be interpreted in the following way. The curved surface which is the apple's skin represents the curved spacetime of Einstein's general theory, which bears the same relation to the flat spacetime of the special theory as does the apple's skin to a plane. Free particles (i.e., those moving under gravity alone, gravity no longer being a force) follow the straightest paths or geodesics in the curved spacetime, just as the ants follow geodesics on the apple's skin. Locally the spacetime of the general theory is like that of the special theory,[2] but on a larger view it is curved, *and this curvature may be detected implicitly by means of geodesic deviation*, just as the curvature of the apple's skin may be detected by noting the convergence of neighboring geodesics. The way in which the dimple around the stalk gives the impression of attraction corresponds to the fact that massive bodies modify the curvature of spacetime in their vicinity, and this modification affects the geodesics in such a way as to give the impression that free particles are acted on by a force, whereas in actual fact they are following the straightest paths in the curved spacetime.

Given that Einstein's general theory does not involve the idea of gravity as a force, how does the gravitational "force" that is a feature of the Newtonian theory arise? We remarked in the Introduction that in a local inertial frame (a freely falling, nonrotating reference system occupying a small region of spacetime) the laws of physics are those of special relativity, and in particular free particles (those moving under gravity alone) follow straight-line paths with constant speed, so for these frames there is no acceleration and consequently no "force." When discussing gravity in Newtonian terms, it is customary to insist that the frame used is nonrotating (so there are no centrifugal or Coriolis "forces"), but one does not normally use a frame that is freely falling, and it is this use of nonfreely falling frames that gives rise to gravitational forces. Just as the fictitious forces associated with rotation (the centrifugal and the Coriolis forces) can be transformed away (locally) by changing to a nonrotating frame, so can the fictitious force of gravity be transformed away by changing to a freely falling frame.

Newton's theory of gravity is nonrelativistic and uses a model for spacetime that combines three-dimensional Euclidean space with one-dimensional time. Getting the Newtonian theory as an approximation to Einstein's general theory of relativity therefore involves two things: passing from a relativistic to a nonrelativistic way of looking at things and interpreting the effects of the curvature of spacetime in the setting of three-dimensional Euclidean space plus one-dimensional time. The whole process is quite complicated, but is essential for a proper understanding of the relationship between Einstein's theory and the Newtonian theory. We shall perform this approximation later in this chapter and establish various points of contact between the two theories.

Before we can do this, we must explain how our model for spacetime can handle the paths of particles by including the handling of geodesics as part of our mathematical repertoire. The mathematics of geodesics is covered in the next few sections, along with the related concepts of parallelism and

[2] Compare remarks made in the Introduction.

absolute and covariant differentiation, needed for our discussion of curvature in Chapter 3. Note that in the present chapter we are concerned only with the motion of particles in a *given* spacetime: they are test particles responding to a *given* gravitational field. How that field arises is answered in the next chapter, where we relate the curvature of spacetime to the sources of the gravitational field.

Exercise 2.0

1. Ants follow geodesics on a surface which is an infinite cylinder (i.e., the outside of an infinitely long straight pipe).
   Do the geodesics deviate?
   By considering only the paths of itself and its neighbors, can an ant decide whether it is on a cylinder or a plane?

2.1 Geodesics

A geodesic in Euclidean space is simply a straight line, which can be characterized as the shortest curve between two points. Such a characterization could be extended to a geodesic in a manifold, where the metric tensor field gives us the length of a curve via the integral (1.82). However, this approach to geodesics presents some technical difficulties, particularly when the metric tensor field is indefinite (as in the spacetime of general relativity), since in that case we can have curves (or parts of curves) that have zero length. We therefore adopt another characterization of a straight line, namely its *straightness*, and use this as a guide to defining geodesics in a manifold.

What makes a straight line straight is the fact that its tangent vectors all point in the same direction. If we use the arc-length s measured from some base point on the line as a parameter, then the tangent vectors $\boldsymbol{\lambda} \equiv \dot{\mathbf{r}}(s)$ have constant length (as they are unit vectors: see Exercise 1.3.3), so we can express the fact that they have constant direction by stating that

$$d\boldsymbol{\lambda}/ds = \mathbf{0}. \tag{2.1}$$

Let us see what form this equation takes when we use arbitrary coordinates u^i and the related natural basis $\{\mathbf{e}_i\}$.

Putting $\boldsymbol{\lambda} = \lambda^i \mathbf{e}_i$ and using dots to denote differentiation with respect to s give

$$0 = d\boldsymbol{\lambda}/ds = d(\lambda^i \mathbf{e}_i)/ds = \dot{\lambda}^i \mathbf{e}_i + \lambda^i \dot{\mathbf{e}}_i. \tag{2.2}$$

Now

$$\dot{\mathbf{e}}_i = \partial_j \mathbf{e}_i \dot{u}^j$$

and in general the vector fields $\partial_j \mathbf{e}_i$ are nonzero. At each point of space, we can refer $\partial_j \mathbf{e}_i$ to the basis $\{\mathbf{e}_i\}$, so that

$$\partial_j \mathbf{e}_i = \mathit{\Gamma}^k_{ij} \mathbf{e}_k,$$

which gives rise to 27 quantities $\mathit{\Gamma}^k_{ij}$ defined at each point of space. After some manipulation and relabeling of dummy suffixes we then get

$$(\dot{\lambda}^i + \mathit{\Gamma}^i_{jk} \lambda^j \dot{u}^k) \mathbf{e}_i = \mathbf{0} \tag{2.3}$$

from equation (2.2). Since $\lambda^i = \dot{u}^i = du^i/ds$, we see that the components du^i/ds of the tangent vector to the straight line satisfy

$$\frac{d^2 u^i}{ds^2} + \mathit{\Gamma}^i_{jk} \frac{du^j}{ds} \frac{du^k}{ds} = 0. \tag{2.4}$$

For this last equation to have any meaning, we must obtain an expression for $\mathit{\Gamma}^i_{jk}$ in terms of known quantities.

We start by noting that[3]

$$\partial_j \mathbf{e}_i = \frac{\partial^2 \mathbf{r}}{\partial u^j \partial u^i} = \frac{\partial^2 \mathbf{r}}{\partial u^i \partial u^j} = \partial_i \mathbf{e}_j,$$

so that $\mathit{\Gamma}^k_{ij} \mathbf{e}_k = \mathit{\Gamma}^k_{ji} \mathbf{e}_k$. Forming the dot product with $\mathbf{e}^l$ then gives the symmetry property

$$\boxed{\mathit{\Gamma}^l_{ij} = \mathit{\Gamma}^l_{ji}.} \tag{2.5}$$

We then use $g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j$ to get

$$\partial_k g_{ij} = \partial_k \mathbf{e}_i \cdot \mathbf{e}_j + \mathbf{e}_i \cdot \partial_k \mathbf{e}_j = \mathit{\Gamma}^m_{ik} \mathbf{e}_m \cdot \mathbf{e}_j + \mathbf{e}_i \cdot \mathit{\Gamma}^m_{jk} \mathbf{e}_m.$$

So

$$\partial_k g_{ij} = \mathit{\Gamma}^m_{ik} g_{mj} + \mathit{\Gamma}^m_{jk} g_{im}. \tag{2.6}$$

Relabeling suffixes we have

$$\partial_i g_{jk} = \mathit{\Gamma}^m_{ji} g_{mk} + \mathit{\Gamma}^m_{ki} g_{jm} \tag{2.7}$$

and

$$\partial_j g_{ki} = \mathit{\Gamma}^m_{kj} g_{mi} + \mathit{\Gamma}^m_{ij} g_{km}. \tag{2.8}$$

Subtracting equation (2.8) from the sum of equations (2.6) and (2.7), and using the symmetry of both $\mathit{\Gamma}^m_{ij}$ and g_{ij} give

$$2\mathit{\Gamma}^m_{ki} g_{mj} = \partial_k g_{ij} + \partial_i g_{jk} - \partial_j g_{ki}.$$

Contracting with $\frac{1}{2} g^{lj}$ then gives

[3] Provided that we can change the order of partial differentiation, which is certainly the case if the coordinate functions $x(u^i)$, $y(u^i)$, $z(u^i)$ have continuous second partial derivatives.

$$\boxed{\Gamma^l_{ki} = \tfrac{1}{2} g^{lj}(\partial_k g_{ij} + \partial_i g_{jk} - \partial_j g_{ki}).} \tag{2.9}$$

Equation (2.4), with Γ^i_{jk} given by equation (2.9) is the *geodesic equation* for Euclidean space.

If we use a general parameter t to parameterize the straight line, then the geodesic equation has a more complicated form. However, for parameters related to the arc-length s by an equation of the form

$$t = As + B, \tag{2.10}$$

where A, B are constants ($A \neq 0$), the geodesic equation has basically the same form as when s is used:

$$\frac{d^2u^i}{dt^2} + \Gamma^i_{jk}\frac{du^j}{dt}\frac{du^k}{dt} = 0. \tag{2.11}$$

(See Exercise 2.1.1 for a justification of these claims.) These privileged parameters for which the geodesic equation has the form (2.11) (with Γ^i_{jk} given by equation (2.9)) are known as *affine parameters.* For an affine parameter, ds/dt is constant, so one is taken along the geodesic at a constant sort of rate. (If we think of t as time, then the geodesic is traversed at constant speed.)

Equation (2.11) represents a system of second-order differential equations whose general solution $u^i(t)$ gives the geodesics of Euclidean space (i.e., straight lines) in whatever coordinate system we happen to be using. To obtain a particular solution, six conditions are needed. These might take the form of specifying a starting point and a starting direction, or of specifying a starting point and an ending point for the geodesic.

Using the above ideas as a guide, we can define an *affinely parameterized geodesic* in an N-dimensional Riemannian or pseudo-Riemannian manifold as a curve given by $x^a(u)$ satisfying[4]

$$\boxed{\frac{d^2x^a}{du^2} + \Gamma^a_{bc}\frac{dx^b}{du}\frac{dx^c}{du} = 0,} \tag{2.12}$$

where the N^3 quantities Γ^a_{bc} are given by

$$\boxed{\Gamma^a_{bc} = \tfrac{1}{2} g^{ad}(\partial_b g_{dc} + \partial_c g_{bd} - \partial_d g_{bc}).} \tag{2.13}$$

These quantities are known as *connection coefficients*[5] and, like their three-dimensional Euclidean counterparts, they satisfy the symmetry relation

[4]Note the change of notation from u^i for coordinates and t for parameter in three-dimensional Euclidean space to x^a for coordinates and u for parameter in an N-dimensional manifold.

[5]The reason for this terminology is given in the next section.

$$\boxed{\Gamma^a_{bc} = \Gamma^a_{cb},} \tag{2.14}$$

as is clear from their defining equation. It can be shown that in moving along an affinely parameterized geodesic, the length of the tangent vector $\dot{x}^a \equiv dx^a/du$ remains constant (see Exercise 2.1.2), and it follows that if the geodesic is not null (which could be the case with an indefinite metric tensor field), then the affine parameter is related to the arc-length s by an equation of the form

$$u = As + B, \tag{2.15}$$

where A, B are constants ($A \neq 0$) (see Exercise 2.1.3). If the metric tensor field is indefinite, then we can have affinely parameterized *null* geodesics whose tangent vectors $\dot{x}^a$ satisfy $g_{ab}\dot{x}^a\dot{x}^b = 0$ and for which the arc-length s cannot be used as a parameter.

Example 2.1.1

In this example we show that, of all the circles of latitude on a sphere, only the equator is a geodesic. We take the radius of the sphere to be a, and use $u^1 = \theta$ and $u^2 = \phi$ (borrowed from spherical coordinates) as parameters.

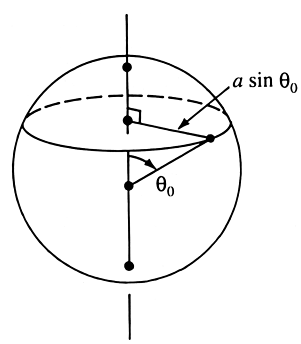

Fig. 2.3. A circle of latitude on a sphere.

The figure shows the circle of latitude given by $\theta = \theta_0$. Since its radius is $a \sin\theta_0$, we can parameterize it by saying that

$$u^1 \equiv \theta = \theta_0, \quad u^2 \equiv \phi = (a\sin\theta_0)^{-1}s,$$

where s is the arc-length measured round from the point where $\phi = 0$. So (for $A = 1, 2$)

$$u^A = \theta_0\delta^A_1 + \frac{s}{a\sin\theta_0}\delta^A_2, \quad \dot{u}^A = \frac{1}{a\sin\theta_0}\delta^A_2, \quad \text{and} \quad \ddot{u}^A = 0,$$

so, for the geodesic equation to be satisfied, we need

$$\ddot{u}^A + \Gamma^A_{BC}\dot{u}^B\dot{u}^C = 0 + \frac{1}{a^2\sin^2\theta_0}\Gamma^A_{22} = 0. \tag{2.16}$$

From Exercise 2.1.5 we have that the only nonzero connection coefficients are

$$\Gamma^1_{22} = -\sin\theta\cos\theta, \quad \Gamma^2_{12} = \Gamma^2_{21} = \cot\theta,$$

so equation (2.16) is satisfied for $A = 2$ (as $\Gamma^2_{22} = 0$), while for $A = 1$ it gives $-\cot\theta_0/a^2 = 0$, which is satisfied only if $\theta_0 = \pi/2$. So, of all the circles of latitude, only the equator is a geodesic.

In order to obtain the parametric equations $x^a = x^a(u)$ of an affinely parameterized geodesic, we must solve the system of differential equations (2.12). These equations are second-order, and require $2N$ conditions to determine a unique solution. Suitable conditions are given by specifying the coordinates x^a_0 of some point on the geodesic, and the components $\dot{x}^a_0$ of the tangent vector at that point. Bearing in mind the equations (2.13) which define the Γ^a_{bc}, it would seem to be a complicated procedure just to set up the geodesic equations, let alone solve them. Fortunately the equations may be generated by a very neat procedure which also produces the Γ^a_{bc} as a spin-off.

Consider the *Lagrangian* $L(\dot{x}^c, x^c) \equiv \frac{1}{2}g_{ab}(x^c)\dot{x}^a\dot{x}^b$, which we regard as a function of $2N$ independent variables x^c and $\dot{x}^c$. The *Euler–Lagrange equations* for a Lagrangian are the equations

$$\boxed{\frac{d}{du}\left(\frac{\partial L}{\partial \dot{x}^c}\right) - \frac{\partial L}{\partial x^c} = 0,} \tag{2.17}$$

and for the given Lagrangian these reduce to the geodesic equations (in covariant rather than contravariant form), as we now show.

Differentiating the Lagrangian we have

$$\frac{\partial L}{\partial \dot{x}^c} = \tfrac{1}{2}g_{ab}\delta^a_c\dot{x}^b + \tfrac{1}{2}g_{ab}\dot{x}^a\delta^b_c = g_{cb}\dot{x}^b$$

and

$$\frac{\partial L}{\partial x^c} = \tfrac{1}{2}\partial_c g_{ab}\dot{x}^a\dot{x}^b,$$

so equations (2.17) are

$$d(g_{cb}\dot{x}^b)/du - \tfrac{1}{2}\partial_c g_{ab}\dot{x}^a\dot{x}^b = 0,$$

or

$$g_{cb}\ddot{x}^b + \partial_a g_{cb}\dot{x}^a\dot{x}^b - \tfrac{1}{2}\partial_c g_{ab}\dot{x}^a\dot{x}^b = 0.$$

But

$$\partial_a g_{cb}\dot{x}^a\dot{x}^b = \tfrac{1}{2}\partial_a g_{cb}\dot{x}^a\dot{x}^b + \tfrac{1}{2}\partial_b g_{ca}\dot{x}^a\dot{x}^b,$$

so we have

$$g_{cb}\ddot{x}^b + \tfrac{1}{2}(\partial_a g_{cb} + \partial_b g_{ca} - \partial_c g_{ab})\dot{x}^a\dot{x}^b = 0.$$

That is, the Euler–Lagrange equations reduce to

$$g_{cb}\ddot{x}^b + \Gamma_{cab}\dot{x}^a\dot{x}^b = 0, \tag{2.18}$$

where $\Gamma_{cab} \equiv \frac{1}{2}(\partial_a g_{cb} + \partial_b g_{ca} - \partial_c g_{ab})$ and raising c gives

$$\ddot{x}^c + \Gamma^c_{ab}\dot{x}^a\dot{x}^b = 0, \tag{2.19}$$

which are the equations of an affinely parameterized geodesic.

Those familiar with the calculus of variations or Lagrangian mechanics will know that the Euler–Lagrange equations give the solution to the problem of finding the curve (with fixed endpoints) which extremizes the integral $\int_{u_1}^{u_2} L(\dot{x}^c, x^c)\, du$. While there is some connection with the characterization of a geodesic as an extremal of length, it should be noted that the integral involved does not give the length of the curve. For reasons stated earlier, we shall not pursue this approach any further, but simply regard the Euler–Lagrange equations as a useful device for generating the geodesic equations and the connection coefficients which may be extracted from them.

Demonstrating the equivalence of the geodesic and the Euler–Lagrange equations allows us to make a useful observation. If g_{ab} does not depend on some particular coordinate x^d, say, then equation (2.17) shows that

$$\frac{d}{du}\left(\frac{\partial L}{\partial \dot{x}^d}\right) = 0,$$

which implies that $\partial L/\partial \dot{x}^d$ is constant along an affinely parameterized geodesic. But $\partial L/\partial \dot{x}^d = g_{db}\dot{x}^b$, so we then have that $p_d \equiv g_{db}\dot{x}^b$ is constant along an affinely parameterized geodesic. The situation is exactly the same as in Lagrangian mechanics where, if the Lagrangian does not contain a particular generalized coordinate, then the corresponding generalized momentum is conserved, and borrowing a term from mechanics we call a coordinate which is absent from g_{ab} *cyclic* or *ignorable*.[6] Being able to say that $p_d =$ constant whenever x^d is cyclic gives us immediate integrals of the geodesic equations, which help with their solution. An example should make some of these ideas clear.

Example 2.1.2

The Robertson–Walker line element is used in cosmology (see Chap. 6). It is defined by

$$g_{\mu\nu}dx^\mu dx^\nu \equiv dt^2 - (R(t))^2\left((1-kr^2)^{-1}dr^2 + r^2 d\theta^2 + r^2\sin^2\theta\, d\phi^2\right),$$

where $\mu, \nu = 0, 1, 2, 3$ (our usual notation for spacetimes), k is a constant, and $x^0 \equiv t$, $x^1 \equiv r$, $x^2 \equiv \theta$, $x^3 \equiv \phi$.

[6] See, for example, Goldstein, Poole, and Safko, 2002, §2–6, or Symon, 1971, §9–6.

So the Lagrangian is

$$L(\dot{x}^\sigma, x^\sigma) \equiv \tfrac{1}{2}\left\{\dot{t}^{\,2} - (R(t))^2\left((1-kr^2)^{-1}\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2\right)\right\},$$

where dots denote differentiation with respect to an affine parameter u. Partial differentiation gives

$$\partial L/\partial\dot{t} = \dot{t},$$

$$\partial L/\partial\dot{r} = -R^2(1-kr^2)^{-1}\dot{r},$$

$$\partial L/\partial\dot{\theta} = -R^2r^2\dot{\theta},$$

$$\partial L/\partial\dot{\phi} = -R^2r^2\sin^2\theta\,\dot{\phi},$$

$$\partial L/\partial t = -RR'[(1-kr^2)^{-1}\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2]$$

(where $R' = dR/dt$),

$$\partial L/\partial r = -R^2(1-kr^2)^{-2}kr\dot{r}^2 - R^2r\dot{\theta}^2 - R^2r\sin^2\theta\,\dot{\phi}^2,$$

$$\partial L/\partial\theta = -R^2r^2\sin\theta\cos\theta\,\dot{\phi}^2,$$

$$\partial L/\partial\phi = 0.$$

Substitution of these derivatives in the Euler–Lagrange equations (2.17) gives

$$\begin{aligned}
&\ddot{t} + RR'[(1-kr^2)^{-1}\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2] = 0,\\
&-R^2(1-kr^2)^{-1}\ddot{r} - 2RR'(1-kr^2)^{-1}\dot{t}\dot{r}\\
&\qquad\quad - R^2(1-kr^2)^{-2}kr\dot{r}^2 + R^2r\dot{\theta}^2 + R^2r\sin^2\theta\,\dot{\phi}^2 = 0,\\
&-R^2r^2\ddot{\theta} - 2RR'r^2\dot{t}\dot{\theta} - 2R^2r\dot{r}\dot{\theta} + R^2r^2\sin\theta\,\cos\theta\,\dot{\phi}^2 = 0,\\
&-R^2r^2\sin^2\theta\,\ddot{\phi} - 2RR'r^2\sin^2\theta\,\dot{t}\dot{\phi}\\
&\qquad\quad - 2R^2r\sin^2\theta\,\dot{r}\dot{\phi} - 2R^2r^2\sin\theta\,\cos\theta\,\dot{\theta}\dot{\phi} = 0.
\end{aligned}$$

The above comprise the covariant version of the geodesic equations, as given by equation (2.18). Because $[g_{\mu\nu}]$ is diagonal, it is a simple matter to obtain the contravariant form of the geodesic equations (as given by equation (2.19)). All we have to do is to divide each equation as necessary, so as to make the coefficients of $\ddot{t}$, $\ddot{r}$, $\ddot{\theta}$ and $\ddot{\phi}$ equal to one. We thus arrive at the geodesic equations for the Robertson–Walker spacetime in standard form:

$$\begin{aligned}
&\ddot{t} + RR'[(1-kr^2)^{-1}\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2] = 0,\\
&\ddot{r} + 2R'R^{-1}\dot{t}\dot{r} + kr(1-kr^2)^{-1}\dot{r}^2\\
&\qquad\quad - r(1-kr^2)\dot{\theta}^2 - r(1-kr^2)\sin^2\theta\,\dot{\phi}^2 = 0,\\
&\ddot{\theta} + 2R'R^{-1}\dot{t}\dot{\theta} + 2r^{-1}\dot{r}\dot{\theta} - \sin\theta\,\cos\theta\,\dot{\phi}^2 = 0,\\
&\ddot{\phi} + 2R'R^{-1}\dot{t}\dot{\phi} + 2r^{-1}\dot{r}\dot{\phi} + 2\cot\theta\,\dot{\theta}\dot{\phi} = 0.
\end{aligned} \tag{2.20}$$

Comparing these with equations (2.19) allows us to pick out the connection coefficients. These are zero except for the following:

$$\begin{array}{lll}
\Gamma^0_{11} = RR'/(1-kr^2), & \Gamma^0_{22} = RR'r^2, & \Gamma^0_{33} = RR'r^2\sin^2\theta, \\
\Gamma^1_{01} = R'/R, & \Gamma^1_{11} = kr/(1-kr^2), & \Gamma^1_{22} = -r(1-kr^2), \\
\Gamma^1_{33} = -r(1-kr^2)\sin^2\theta, & & \\
\Gamma^2_{02} = R'/R, & \Gamma^2_{12} = 1/r, & \Gamma^2_{33} = -\sin\theta\cos\theta, \\
\Gamma^3_{03} = R'/R, & \Gamma^3_{13} = 1/r, & \Gamma^3_{23} = \cot\theta.
\end{array}$$

Note, for example, that in the second geodesic equation $2R'R^{-1}\dot{t}\dot{r}$ includes two terms of the sum $\Gamma^1_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$, namely $\Gamma^1_{01}\dot{x}^0\dot{x}^1$ and $\Gamma^1_{10}\dot{x}^1\dot{x}^0$, and one must remember to halve the multipliers of the cross terms $\dot{x}^\mu\dot{x}^\nu$ ($\mu \neq \nu$) when extracting the connection coefficients from the geodesic equations.

Note also that in the example above ϕ is a cyclic coordinate, so one may say immediately that $\partial L/\partial\dot{\phi}$ is constant; that is, $R^2r^2\sin^2\theta\,\dot{\phi} = A$. Differentiation with respect to u results in the last geodesic equation, showing that we do indeed have an integral.

Exercises 2.1

1. Show that if a general parameter $t = f(s)$ is used to parameterize a straight line in Euclidean space, then the geodesic equation takes the form
$$\frac{d^2u^i}{dt^2} + \Gamma^i_{jk}\frac{du^j}{dt}\frac{du^k}{dt} = h(s)\frac{du^i}{dt},$$
where $h(s) = -\dfrac{d^2t}{ds^2}\left(\dfrac{dt}{ds}\right)^{-2}$.

   Deduce that this reduces to the simple form (2.11) if, and only if, $t = As + B$, where A, B are constants ($A \neq 0$).

2. The aim of this exercise is to show that the length L of the tangent vector $\dot{x}^a$ to an affinely parameterized geodesic is constant.
   (a) Start by arguing that $\pm L^2 = g_{ab}\dot{x}^a\dot{x}^b$.
   (b) Differentiate this equation to obtain an expression for $\pm 2L\dot{L}$ in terms of the quantities g_{ab}, $\dot{g}_{ab}$, $\dot{x}^a$, and $\ddot{x}^a$.
   (c) Put $\dot{g}_{ab} = \partial_c g_{ab}\dot{x}^c$ and use the geodesic equation (2.12) to express the second derivatives $\ddot{x}^a$ in terms of the connection coefficients Γ^a_{bc} and the first derivatives $\dot{x}^a$.
   (d) Then use equation (2.13) to express the Γ^a_{bc} in terms of the metric tensor components and their derivatives.

(e) Simplify to obtain $2L\dot{L} = 0$, from which it follows that $\dot{L} = 0$ and L is constant.

(See Exercise 2.3.4 for a much shorter derivation of this result.)

3. Use the result of Exercise 2 to show that, for a non-null geodesic affinely parameterized by u, $u = As + B$, where A, B are constants ($A \neq 0$).

4. Show that for any geodesic (non-null or null) any two affine parameters u and u' are related by an equation of the form $u' = Au + B$, where A, B are constants with $A \neq 0$.

5. Use the result of Exercise 1.6.2(a) to show that, for a sphere of radius a parameterized in the usual way by $u^1 \equiv \theta$, $u^2 \equiv \phi$ (borrowed from spherical coordinates), the metric tensor components are given by

$$[g_{AB}] = \begin{bmatrix} a^2 & 0 \\ 0 & a^2 \sin^2\theta \end{bmatrix}.$$

Deduce that the only nonzero connection coefficients are

$$\Gamma^1_{22} = -\sin\theta\cos\theta, \quad \Gamma^2_{12} = \Gamma^2_{21} = \cot\theta.$$

6. Show that all lines of longitude on a sphere (curves given by $\phi =$ constant) are geodesics.

7. In a Robertson–Walker spacetime, a coordinate curve for which r, θ, ϕ are constant and t varies is given by

$$x^\mu(u) = u\delta^\mu_0 + r_0\delta^\mu_1 + \theta_0\delta^\mu_2 + \phi_0\delta^\mu_3,$$

where r_0, θ_0, ϕ_0 are constants and u is a parameter. Verify that all such coordinate curves are geodesics affinely parameterized by u.

(See Example 2.1.2 for the connection coefficients for a Robertson–Walker spacetime.)

2.2 Parallel vectors along a curve

Our way of characterizing a straight line in Euclidean space and (by extension) a geodesic in a manifold is related to the idea of parallelly transporting a vector along a curve.

Let γ be a curve in three-dimensional Euclidean space given parametrically by $u^i(t)$ and let P_0 with parameter t_0 be some initial point on γ where we give a vector $\boldsymbol{\lambda}_0$. We can think of transporting $\boldsymbol{\lambda}_0$ along γ *without any change to its length or direction* so as to obtain a parallel vector $\boldsymbol{\lambda}(t)$ at each point of γ (see Fig. 2.4). The result is a *parallel field of vectors along* γ generated

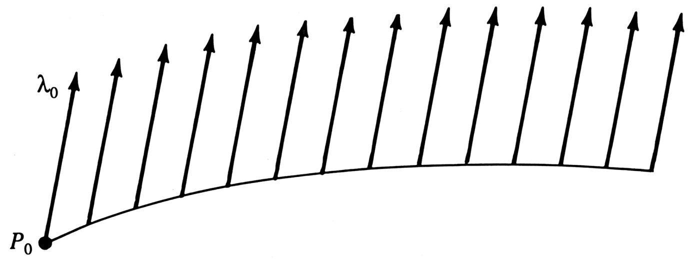

Fig. 2.4. A parallel field of vectors generated by parallel transport.

by the *parallel transport* of $\boldsymbol{\lambda}_0$ along γ. Since there is no change in the length or direction of $\boldsymbol{\lambda}(t)$ along γ, it satisfies the differential equation

$$d\boldsymbol{\lambda}/dt = \mathbf{0}, \tag{2.21}$$

for which $\boldsymbol{\lambda}(t_0) = \boldsymbol{\lambda}_0$ is the initial condition. If we work on equation (2.21) like we did on equation (2.1), then we can deduce from an equation like equation (2.3) that the components λ^i of the transported vector satisfy

$$\dot{\lambda}^i + \Gamma^i_{jk}\lambda^j \dot{u}^k = 0, \tag{2.22}$$

where the connection coefficients are given by equation (2.9).

Equation (2.22) is the component version of the equation for parallelly transporting a vector along a curve in Euclidean space. Its generalization for the parallel transport of a contravariant vector λ^a along a curve γ in an N-dimensional manifold with metric tensor field g_{ab} is clearly

$$\boxed{\dot{\lambda}^a + \Gamma^a_{bc}\lambda^b \dot{x}^c = 0,} \tag{2.23}$$

where the connection coefficients are given by equation (2.13) and $\dot{x}^a$ is the tangent vector arising from the parameterization $x^a(u)$ of γ. We now see that our definition of an affinely parameterized geodesic in the previous section amounts to saying that it is a curve characterized by the fact that its tangent vectors $\dot{x}^a$ form a parallel field of vectors along itself.

Parallel transport along curves in a curved manifold is significantly different from that along curves in flat Euclidean space in that it is *path-dependent*: the vector obtained by transporting a given vector from a point P to a remote point Q depends on the route taken from P to Q. This path dependence also shows up in transporting a vector around a closed loop, where on returning to the starting point the direction of the transported vector is (in general) different from the vector's initial direction. This path dependence can be demonstrated on a curved surface, in both practical and mathematical terms.

In Appendix B we describe the construction of a machine that gives a practical means of transporting a vector parallelly along a curve on a surface. It is a small two-wheeled vehicle carrying a pointer (which represents the vector) equipped with some rather clever gearing that receives input from the two wheels and outputs adjustments to the direction of the pointer. These adjustments ensure that the pointer is parallelly transported along the path taken by the vehicle. If we were to take this parallel transporter for walks on various surfaces, we would confirm that, for a curved surface, parallel transport is (in general) path-dependent, while, for a plane, it is path-independent. We would also observe that on completing a closed loop on a curved surface, the final direction of the pointer is (in general) different from its initial direction. The following example illustrates in mathematical terms this phenomenon for curves on a sphere.

Example 2.2.1

Consider a sphere of radius a, coordinatized in the usual way using $u^1 \equiv \theta$,

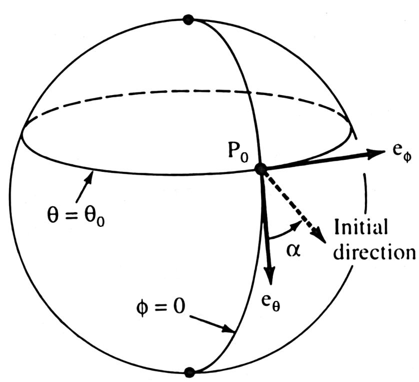

Fig. 2.5. Parallel transport around a circle of latitude.

$u^2 \equiv \phi$, where θ, ϕ are polar angles borrowed from spherical coordinates, with $0 \le \theta \le \pi$ and (for convenience) $0 \le \phi \le 2\pi$. Then

$$[g_{AB}] = \begin{bmatrix} a^2 & 0 \\ 0 & a^2 \sin^2\theta \end{bmatrix}$$

and the only nonzero connection coefficients are

$$\Gamma^1_{22} = -\sin\theta\cos\theta, \quad \Gamma^2_{12} = \Gamma^2_{21} = \cot\theta$$

(see Exercise 2.1.5). Let us transport a vector $\boldsymbol{\lambda}$ parallelly around the circle of latitude γ given by $\theta = \theta_0$ ($\theta_0 =$ const), starting and ending at the point P_0 where $\phi = 0$ or 2π (see Fig. 2.5). The circle is given parametrically by

$$u^A(t) = \theta_0 \delta_1^A + t\delta_2^A, \quad 0 \le t \le 2\pi,$$

so $\dot{u}^A = \delta_2^A$ and the equation for parallel transport becomes $\dot{\lambda}^A + \Gamma^A_{B2}\lambda^B = 0$, which is equivalent to the pair

$$\begin{cases} \dot{\lambda}^1 - \sin\theta_0 \cos\theta_0 \, \lambda^2 = 0 \\ \qquad \dot{\lambda}^2 + \cot\theta_0 \, \lambda^1 = 0 \end{cases}. \tag{2.24}$$

Suppose initially that $\boldsymbol{\lambda}$ is a unit vector whose direction makes an angle α east of south. Then

$$\lambda^1(0) = a^{-1}\cos\alpha, \quad \lambda^2(0) = (a\sin\theta_0)^{-1}\sin\alpha, \tag{2.25}$$

as may be checked by noting that these must satisfy

$$g_{AB}\lambda^A(0)\lambda^B(0) = 1 \quad \text{and} \quad g_{AB}\lambda^A(0)S^B = \cos\alpha,$$

where $S^A \equiv a^{-1}\delta_1^A$ is the south-pointing unit vector at P_0.

We have an initial-value problem comprising the pair of equations (2.24) with initial conditions (2.25). Its solution is

$$\begin{cases} \lambda^1 = a^{-1}\cos(\alpha - \omega t) \\ \lambda^2 = (a\sin\theta_0)^{-1}\sin(\alpha - \omega t) \end{cases}, \tag{2.26}$$

where $\omega = \cos\theta_0$, as may be checked (see Exercise 2.2.1). On completing the circuit of γ, the vector obtained by parallel transport has components

$$\begin{cases} \lambda^1(2\pi) = a^{-1}\cos(\alpha - 2\pi\omega) \\ \lambda^2(2\pi) = (a\sin\theta_0)^{-1}\sin(\alpha - 2\pi\omega) \end{cases}.$$

We see that $g_{AB}\lambda^A(2\pi)\lambda^B(2\pi) = 1$, so $\lambda^A(2\pi)$ is a unit vector (as it should be), but its direction is not that of the initial vector $\lambda^A(0)$ (unless ω happens to be zero, as on the equator). Noting that

$$\begin{aligned} g_{AB}\lambda^A(0)\lambda^B(2\pi) &= \cos\alpha\cos(\alpha - 2\pi\omega) + \sin\alpha\sin(\alpha - 2\pi\omega) \\ &= \cos(\alpha - (\alpha - 2\pi\omega)) \\ &= \cos 2\pi\omega, \end{aligned}$$

we see that the final vector makes an angle of $2\pi\omega$ with the initial vector, where $\omega \equiv \cos\theta_0$. For example, for $\theta_0 = 85°$ the vector has twisted through $31.4°$, whereas for $\theta_0 = 5°$ (near the North Pole) the angle between the final and initial vectors is $1.4°$.

The above example can be used to illustrate two further points concerning parallel transport. The first of these is that if the curve along which the vector is transported is a geodesic, then the angle between the transported vector and the tangent to the geodesic remains constant. This is clearly the case when the geodesic is a straight line in Euclidean space and we shall obtain it as a general result for a manifold in the next section. The verification of this result for the sphere is left as an exercise for the reader (see Exercise 2.2.3).

The second point concerns transporting a vector parallelly around a closed curve that is "small." If in the example above θ_0 is small, then γ is a small circle about the North Pole, $\omega \equiv \cos\theta_0 \approx 1$ and the angle between the initial direction and the final direction is approximately 2π, which amounts to a negligible discrepancy between the initial and final vectors. This illustrates the fact that, by sticking to a small portion of a curved surface, we tend not to pick up its curvature by parallel transport around a closed curve.[7] Locally the surface behaves much as if it were flat, and experimentation over an extended region gives us a better chance of detecting curvature. The same is true for manifolds in general.

The connection coefficients Γ^a_{bc} are said to define a *connection* on the manifold. The reason for this kind of terminology is because it provides us with a connection between tangent spaces at different points of a manifold, enabling us to associate a vector in the tangent space at one point with the vector parallel to it at another point. For widely separated points, this association depends on the path used to connect the points, but for neighboring points (separated by small coordinate differences) the association is unique (up to first order in the small coordinate differences), as we now go on to show.

Suppose that P with coordinates x^a and Q with coordinates $x^a + \delta x^a$ are nearby points. Let γ be any parameterized curve through P and Q, with P having parameter u and Q having parameter $u + \delta u$, and let $\bar{\lambda}^a \equiv \lambda^a + \delta\lambda^a$ be the vector at Q parallel to a given vector λ^a at P. Since the vector at Q is obtained by the parallel transport of λ^a at P along the short piece of curve from P to Q, we have that

$$\delta\lambda^a \approx \frac{d\lambda^a}{du}\delta u,$$

where (from equation (2.23)) $d\lambda^a/du = -\Gamma^a_{bc}\lambda^b dx^c/du$, which gives

$$\bar{\lambda}^a \approx \lambda^a - \Gamma^a_{bc}\lambda^b \frac{dx^c}{du}\delta u \approx \lambda^a - \Gamma^a_{bc}\lambda^b \delta x^c. \tag{2.27}$$

So to first order in δx^a, we have a linear mapping from the the tangent space T_P to the tangent space T_Q in which the vector at P with components λ^a is mapped into the parallel vector at Q with components $\bar{\lambda}^a = A^a_b \lambda^b$, where

$$A^a_b \equiv \delta^a_b - \Gamma^a_{bc}\delta x^c. \tag{2.28}$$

We shall make use of this mapping in the next section when defining absolute and covariant differentiation.

In adopting equation (2.23) as the equation defining the parallel transport of a contravariant vector along a curve in a manifold, we completely ignored the

[7] This is because it is a second-order effect. See Sec. 3.3.

question of coordinate independence. If we were to use a primed coordinate system and perform parallel transport by having $\lambda^{a'}$ satisfy

$$\dot{\lambda}^{a'} + \Gamma^{a'}_{b'c'}\lambda^{b'}\dot{x}^{c'} = 0, \tag{2.29}$$

where

$$\Gamma^{a'}_{b'c'} \equiv \tfrac{1}{2}g^{a'd'}(\partial_{b'}g_{d'c'} + \partial_{c'}g_{b'd'} - \partial_{d'}g_{b'c'}), \tag{2.30}$$

would we get the same parallel field of vectors along the curve? We can answer this question in an indirect sort of way by showing that (to first order in small coordinate differences) the mapping from T_P to T_Q given by equation (2.27) does not depend on the coordinate system used. Thus we need to show that if

$$\bar{\lambda}^{a'} \equiv \lambda^{a'} - \Gamma^{a'}_{b'c'}\lambda^{b'}\delta x^{c'},$$

then

$$\bar{\lambda}^{a'}(X^e_{a'})_Q = \bar{\lambda}^e,$$

where $(X^e_{a'})_Q$ denotes $\partial x^e/\partial x^{a'}$ evaluated at Q. In terms of values at P, we can say that (to first order)

$$(X^e_{a'})_Q = X^e_{a'} + X^e_{d'a'}\delta x^{d'},$$

where $X^e_{d'a'} = \partial^2 x^e/\partial x^{d'}\partial x^{a'}$, so we need to show that

$$(\lambda^{a'} - \Gamma^{a'}_{b'c'}\lambda^{b'}\delta x^{c'})(X^e_{a'} + X^e_{d'a'}\delta x^{d'}) = \lambda^e - \Gamma^e_{fg}\lambda^f\delta x^g.$$

But $\lambda^{a'}X^e_{a'} = \lambda^e$, so (to first order) the above condition reduces to

$$X^e_{d'a'}\lambda^{a'}\delta x^{d'} - \Gamma^{a'}_{b'c'}X^e_{a'}\lambda^{b'}\delta x^{c'} = -\Gamma^e_{fg}\lambda^f\delta x^g,$$

which is equivalent to

$$(\Gamma^{a'}_{b'c'} - \Gamma^d_{fg}X^{a'}_d X^f_{b'}X^g_{c'} - X^d_{c'b'}X^{a'}_d)X^e_{a'}\lambda^{b'}\delta x^{c'} = 0, \tag{2.31}$$

since $X^f_{b'}\lambda^{b'} = \lambda^f$ and (to first order) $X^g_{c'}\delta x^{c'} = \delta x^g$. Using the defining equations (2.13) and (2.30), we can show that the connection coefficients transform according to

$$\boxed{\Gamma^{a'}_{b'c'} = \Gamma^d_{fg}X^{a'}_d X^f_{b'}X^g_{c'} + X^d_{c'b'}X^{a'}_d} \tag{2.32}$$

(see Exercise 2.2.4), so condition (2.31) is satisfied and the coordinate independence of parallel transport is established. (See Exercise 2.2.6 for a more direct way of establishing this result.) Since we can express the definition of a geodesic in terms of parallel transport, it follows that this definition is also coordinate-independent.

We finish this section by establishing a few formulae involving the connection coefficients Γ^a_{bc} and the related quantities Γ_{abc} defined by

$$\boxed{\Gamma_{abc} \equiv \tfrac{1}{2}(\partial_b g_{ac} + \partial_c g_{ba} - \partial_a g_{bc}).} \tag{2.33}$$

The traditional names for Γ_{abc} and Γ^a_{bc} are *Christoffel symbols of the first and second kinds*, respectively, and the notation $\Gamma_{abc} \equiv [bc, a]$, $\Gamma^a_{bc} \equiv \left\{ {a \atop bc} \right\}$ is often used, especially in older texts.

From equation (2.13) we see that

$$\boxed{\Gamma^a_{bc} = g^{ad}\Gamma_{dbc}} \tag{2.34}$$

and a short calculation shows that

$$\boxed{\Gamma_{abc} = g_{ad}\Gamma^d_{bc}.} \tag{2.35}$$

Adding Γ_{bac} to Γ_{abc} gives

$$\boxed{\partial_c g_{ab} = \Gamma_{abc} + \Gamma_{bac},} \tag{2.36}$$

allowing us to express the partial derivatives of the metric tensor components in terms of the connection coefficients. If we denote the value of the determinant $|g_{ab}|$ by g, then the cofactor of g_{ab} in this determinant is gg^{ab}. (Note that g is not a scalar: changing coordinates changes the value of g at any point.) It follows that $\partial_c g = (\partial_c g_{ab})gg^{ab}$, so from equations (2.36) and (2.34) we have

$$\partial_c g = gg^{ab}(\Gamma_{abc} + \Gamma_{bac}) = g(\Gamma^b_{bc} + \Gamma^a_{ac}) = 2g\Gamma^a_{ac}.$$

So the contraction Γ^a_{ab} of the connection coefficients is given by

$$\Gamma^a_{ab} = \tfrac{1}{2}g^{-1}\partial_b g = \tfrac{1}{2}\partial_b \ln|g|\,, \tag{2.37}$$

the modulus signs being needed as g is not necessarily positive in the indefinite case. Alternative expressions are

$$\Gamma^a_{ab} = \partial_b \ln|g|^{1/2} \quad \text{and} \quad \Gamma^a_{ab} = |g|^{-1/2}\,\partial_b\,|g|^{1/2}\,. \tag{2.38}$$

Exercises 2.2

1. Verify that the initial-value problem comprising the pair of equations (2.24) with initial conditions (2.25) has a solution given by equations (2.26).

2. For what circle(s) of latitude is the final direction of the transported vector in Example 2.2.1 exactly opposite to that of the initial direction?

3. Noting the result of Example 2.1.1, verify that for parallel transport along a geodesic the angle between the transported vector of Example 2.2.1 and the tangent to the geodesic is constant.

4. Verify that the connection coefficients transform according to equation (2.32).

5. Show that an alternative form for the transformation formula (2.32) is

$$\Gamma^{a'}_{b'c'} = \Gamma^{d}_{ef} X^{a'}_{d} X^{e}_{b'} X^{f}_{c'} - X^{e}_{b'} X^{f}_{c'} X^{a'}_{ef}. \tag{2.39}$$

6. By transforming the left-hand side of equation (2.23) to a primed coordinate system, show that this defining equation for the parallel transport of a contravariant vector along a curve is coordinate-independent.

2.3 Absolute and covariant differentiation

In this section, we turn our attention to the effect of differentiation on tensor fields on a manifold M. Initially we shall consider fields defined along a curve, rather than throughout a region U or throughout the whole manifold M. Here we can regard the components of the field as functions of the parameter u used to label points on the curve, and we can consider their derivatives with respect to u. As we shall see, these derivatives are *not* the components of a tensor, which may come as a surprise to those used to differentiating the velocity components of a particle with respect to time t (which acts as a parameter along the particle's path) to obtain its acceleration. To make differentiation respect the tensor character of fields it needs to be modified, which, for differentiation along curves, leads to the idea of the *absolute derivative*. Having made this modification for fields along curves, we shall then go on to consider tensor fields defined throughout a region covered by a coordinate system, where the components can be regarded as functions of the coordinates. For these there is a corresponding modification of partial differentiation, called *covariant differentiation*, which is defined so that it respects tensor character. Both absolute and covariant differentiation depend on the notion of parallelism introduced in the previous section. These ideas play a crucial role in the formulation of the general theory of relativity and, because of this, this section and the following one are particularly important.

Suppose that we have a vector field $\lambda^a(u)$ defined along a curve γ given parametrically by $x^a(u)$. As we remarked above, the N quantities $d\lambda^a/du$ are *not* the components of a vector. To see this we use another (primed) coordinate system and look at the corresponding primed quantities $d\lambda^{a'}/du$ to see how they are related to the unprimed quantities $d\lambda^a/du$. These primed quantities are given by

$$d\lambda^{a'}/du = d(X^{a'}_{b}\lambda^b)/du = X^{a'}_{b}(d\lambda^b/du) + X^{a'}_{bc}(dx^c/du)\lambda^b, \tag{2.40}$$

and the term involving $X^{a'}_{bc} \equiv \partial^2 x^{a'}/\partial x^b \partial x^c$ would be absent if the $d\lambda^a/du$ were the components of a vector. The reason for the presence of this term is that in the defining equation

$$\frac{d\lambda^a}{du} \equiv \lim_{\delta u \to 0} \frac{\lambda^a(u+\delta u) - \lambda^a(u)}{\delta u}, \tag{2.41}$$

we take differences of components at *different* points of γ, and here is the origin of our problem. Because in general the transformation coefficients depend on position, we have $(X^{a'}_b)_u \neq (X^{a'}_b)_{u+\delta u}$, which means that these differences in components are not the components of a vector (at either of the points in question). In the limit the difference between $(X^{a'}_b)_u$ and $(X^{a'}_b)_{u+\delta u}$ shows up as $X^{a'}_{bc}$. For differentiation to yield a vector, we must take component differences at the *same* point of γ, and we can do this by exploiting the notion of parallelism introduced in the previous section.

Let P be the point on γ with parameter value u and Q be a neighboring point with parameter value $u+\delta u$. Then $\lambda^a(u+\delta u)$ is a vector at Q, as is the vector $\bar{\lambda}^a$ obtained by the parallel transport of $\lambda^a(u)$ at P to Q. The difference $\lambda^a(u+\delta u)-\bar{\lambda}^a$ is then a vector at Q, and so is the quotient $(\lambda^a(u+\delta u)-\bar{\lambda}^a)/\delta u$. It is the limit of this quotient (as $\delta u \to 0$) that gives the *absolute derivative* $D\lambda^a/du$ of $\lambda^a(u)$ along γ. Now

$$\lambda^a(u+\delta u) \approx \lambda^a(u) + \frac{d\lambda^a}{du}\delta u$$

and, from equation (2.27),

$$\bar{\lambda}^a \approx \lambda^a(u) - \Gamma^a_{bc}\lambda^b(u)\delta x^c,$$

so

$$\frac{\lambda^a(u+\delta u) - \bar{\lambda}^a}{\delta u} \approx \frac{d\lambda^a}{du} + \Gamma^a_{bc}\lambda^b(u)\frac{\delta x^c}{\delta u}.$$

As $\delta u \to 0$, the point Q tends to P, and the limit of the quotient is

$$\boxed{\frac{D\lambda^a}{du} \equiv \frac{d\lambda^a}{du} + \Gamma^a_{bc}\lambda^b\frac{dx^c}{du},} \tag{2.42}$$

where all quantities are evaluated at the same point P of γ. Thus the absolute derivative of a vector field λ^a along a curve γ (which is a vector field along γ) involves not only the total derivative $d\lambda^a/du$ (which is not a vector field along γ), but also the connection coefficients Γ^a_{bc}.

The claim that the absolute derivative is a vector field along γ is justified by the way in which it is defined. It can also be justified by checking that

$$\left(\frac{d\lambda^{a'}}{du} + \Gamma^{a'}_{b'c'}\lambda^{b'}\frac{dx^{c'}}{du}\right) = X^{a'}_d\left(\frac{d\lambda^d}{du} + \Gamma^d_{ef}\lambda^e\frac{dx^f}{du}\right), \tag{2.43}$$

using the transformation equations (2.40) and (2.39) for $d\lambda^a/du$ and Γ^a_{bc}. Both of these involve second derivatives of the form $X^{a'}_{bc}$, but in such a way that they cancel when used to transform the quantities $D\lambda^a/du$. (See Exercise 2.3.1.)

We now see that equation (2.23) for parallelly transporting a contravariant vector along a curve can be written as $D\lambda^a/du = 0$ and that $\lambda^a(u)$ form a parallel field of vectors along γ if, and only if, $D\lambda^a/du = 0$. By extending the definition of absolute differentiation to general tensor fields $\tau^{a_1\ldots a_r}_{b_1\ldots b_s}(u)$ defined along a curve γ, we can also extend the notion of parallel transport along γ by requiring that $D\tau^{a_1\ldots a_r}_{b_1\ldots b_s}/du = 0$.

There are two approaches which may be taken to defining the absolute derivative of general tensor fields along a curve. One is to extend the notion of parallelism between neighboring tangent spaces T_P and T_Q to one between the space of type (r, s) tensors at P and the space of type (r, s) at Q, while the other is to demand that the operation of absolute differentiation satisfies certain reasonable conditions which allow us to extend the concept to general tensor fields along curves. We shall take the latter course, and impose the following conditions on the differential operator D/du applied to tensor fields defined along a curve parameterized by u:

(a) When applied to a tensor field, D/du yields a tensor field of the same type.
(b) D/du is a linear operation.
(c) D/du obeys Leibniz' rule with respect to tensor products.
(d) For any scalar field ϕ, $D\phi/du = d\phi/du$.

Condition (b) is a normal requirement of a differential operator and simply means that we are allowed to say things like $D(\sigma^{ab}_c + \tau^{ab}_c)/du = D\sigma^{ab}_c/du + D\tau^{ab}_c/du$, and $D(k\tau^a_{bc})/du = k(D\tau^a_{bc}/du)$ for constant k, while condition (c) allows us to say things like $D(\sigma^a_b\tau^{cd})/du = (D\sigma^a_b/du)\tau^{cd} + \sigma^a_b(D\tau^{cd}/du)$.

We now show how, by using conditions (a)–(d) and the expression already obtained for a contravariant vector field, we can obtain expressions for the absolute derivatives of tensor fields of any type. We shall do this in detail for some simple fields of specific type, from which we shall be able to infer the general pattern for a field of any type.

The absolute derivative of a scalar field

From condition (d) above, we have

$$\boxed{D\phi/du \equiv d\phi/du.} \tag{2.44}$$

The absolute derivative of a contravariant vector field

With the dot-notation for derivatives, equation (2.42) takes the form

$$\boxed{D\lambda^a/du \equiv \dot{\lambda}^a + \Gamma^a_{bc}\lambda^b\dot{x}^c.} \tag{2.45}$$

The absolute derivative of a covariant vector field

If μ_a is a covariant vector field along a curve γ, then for any contravariant vector field λ^a along γ, $\lambda^a \mu_a$ is a scalar field, so using equation (2.44) we have

$$d(\lambda^a \mu_a)/du = D(\lambda^a \mu_a)/du. \tag{2.46}$$

Then using Leibniz' rule (condition (c)) on the contracted tensor product we get

$$\begin{aligned} \frac{d\lambda^a}{du}\mu_a + \lambda^a \frac{d\mu_a}{du} &= \frac{D\lambda^a}{du}\mu_a + \lambda^a \frac{D\mu_a}{du} \\ &= \mu_a \left(\frac{d\lambda^a}{du} + \Gamma^a_{bc}\lambda^b \frac{dx^c}{du} \right) + \lambda^a \frac{D\mu_a}{du}, \end{aligned}$$

which implies that

$$\lambda^a \frac{D\mu_a}{du} = \lambda^a \frac{d\mu_a}{du} - \Gamma^a_{bc}\lambda^b \frac{dx^c}{du}\mu_a = \lambda^a \left(\dot{\mu}_a - \Gamma^d_{ac}\mu_d \dot{x}^c \right).$$

Since this holds for arbitrary vector fields λ^a, we deduce that

$$\boxed{D\mu_a/du \equiv \dot{\mu}_a - \Gamma^d_{ac}\mu_d \dot{x}^c,} \tag{2.47}$$

and in this way our conditions yield the absolute derivative of a covariant vector field. (As we note below, when forming the absolute derivative of a tensor field, a Γ term with a minus sign is included for each subscript. As a reminder, we can extend our mnemonic to "co-below and minus.")

The absolute derivative of a type $(\mathbf{2}, \mathbf{0})$ tensor field

As a guide to obtaining an expression for the absolute derivative of a type $(2, 0)$ tensor field, we consider the special case in which $\tau^{ab} = \lambda^a \mu^b$, where λ^a, μ^a are contravariant vector fields along the curve. Then using condition (c) we have

$$D\tau^{ab}/du = D(\lambda^a \mu^b)/du = (D\lambda^a/du)\mu^b + \lambda^a (D\mu^b/du).$$

Inserting appropriate expressions for $D\lambda^a/du$ and $D\mu^b/du$, and recombining $\lambda^a \mu^b$ as τ^{ab} results in

$$\boxed{D\tau^{ab}/du \equiv \dot{\tau}^{ab} + \Gamma^a_{cd}\tau^{cb}\dot{x}^d + \Gamma^b_{cd}\tau^{ac}\dot{x}^d,} \tag{2.48}$$

which we take to be the formula for the absolute derivative of a type $(2, 0)$ tensor field.

The absolute derivative of a type $(0,2)$ tensor field
Similarly, by considering $\tau_{ab} = \lambda_a \mu_b$, we can arrive at

$$\boxed{D\tau_{ab}/du \equiv \dot{\tau}_{ab} - \Gamma^c_{ad}\tau_{cb}\dot{x}^d - \Gamma^c_{bd}\tau_{ac}\dot{x}^d} \tag{2.49}$$

as the formula for the absolute derivative of a type $(0,2)$ tensor field. (See Exercise 2.3.2.)

The absolute derivative of a type $(1,1)$ tensor field
Likewise, by considering $\tau^a_b = \lambda^a \mu_b$, we get

$$\boxed{D\tau^a_b/du \equiv \dot{\tau}^a_b + \Gamma^a_{cd}\tau^c_b\dot{x}^d - \Gamma^c_{bd}\tau^a_c\dot{x}^d} \tag{2.50}$$

for the absolute derivative of a type $(1,1)$ tensor field. (Again, see Exercise 2.3.2.)

The pattern should now be clear. The absolute derivative of a type (r,s) tensor field $\tau^{a_1\ldots a_r}_{b_1\ldots b_s}$ along a curve γ is given by the sum of the total derivative $\dot{\tau}^{a_1\ldots a_r}_{b_1\ldots b_s}$ of its components, r terms of the form $\Gamma^{a_k}_{cd}\tau^{\ldots c\ldots}_{\ldots}\dot{x}^d$ and s terms of the form $-\Gamma^c_{b_k d}\tau^{\ldots}_{\ldots c\ldots}\dot{x}^d$. For example,

$$D\tau^{ab}_c/du \equiv \dot{\tau}^{ab}_c + \Gamma^a_{de}\tau^{db}_c\dot{x}^e + \Gamma^b_{de}\tau^{ad}_c\dot{x}^e - \Gamma^d_{ce}\tau^{ab}_d\dot{x}^e. \tag{2.51}$$

As we remarked above, we can extend the notion of parallel transport to a tensor of any type, simply by requiring that its absolute derivative along the curve be zero. Again we emphasize that the *parallel transport of tensors is in general path-dependent*. Scalar fields are, of course, excepted, since $D\phi/du = 0$ implies that $d\phi/du = 0$, which in turn implies that ϕ is constant along the curve.

We are now in a position to introduce the *covariant derivative* of a tensor field, which is closely related to the absolute derivative. For absolute differentiation, the tensor fields involved need only be defined along the curve in question. The covariant derivative arises where we have a tensor field defined throughout M (or throughout a region of M).

Suppose, for example, we have a contravariant vector field λ^a defined throughout a region U. If γ is a curve in U, we can restrict λ^a to γ, and define its absolute derivative:

$$D\lambda^a/du \equiv \dot{\lambda}^a + \Gamma^a_{bc}\lambda^b\dot{x}^c. \tag{2.52}$$

But $\dot{\lambda}^a = \dfrac{\partial\lambda^a}{\partial x^c}\dot{x}^c$, so this may be written

$$\frac{D\lambda^a}{du} = \left(\frac{\partial\lambda^a}{\partial x^c} + \Gamma^a_{bc}\lambda^b\right)\dot{x}^c.$$

The bracketed expression on the right of this last equation does not depend on γ but only on the components λ^a and their derivatives at the point in question, and the equation is true for arbitrary tangent vectors $\dot{x}^a$ at the point in question. The usual argument involving the quotient theorem entitles us to deduce that $\dfrac{\partial \lambda^a}{\partial x^c} + \Gamma^a_{bc}\lambda^b$ are the components of a type $(1,1)$ tensor field. This tensor field is the *covariant derivative* of the vector field λ^a, and we denote it by $\lambda^a{}_{;c}$.

It is convenient at this point to introduce some more notation. We have already used ∂_a as an abbreviation for $\partial/\partial x^a$ and we shall continue to use it when dealing with covariant derivatives. We shall also use a comma followed by a subscript a written after the object on which it is acting to mean the same thing. So the covariant derivative of λ^a may be written as

$$\boxed{\lambda^a{}_{;c} = \partial_c \lambda^a + \Gamma^a_{bc}\lambda^b = \lambda^a{}_{,c} + \Gamma^a_{bc}\lambda^b.} \tag{2.53}$$

This notation extends naturally to repeated derivatives. For example, we write $\partial^2\lambda^a/\partial x^b \partial x^c$ as $\partial_b\partial_c\lambda^a$ or as $\lambda^a{}_{,cb}$. In a similar way we shall use $\lambda^a{}_{;cb}$ to denote the repeated covariant derivative $(\lambda^a{}_{;c})_{;b}$.

Returning now to covariant differentiation, we see that the argument above may be applied to a type (r,s) tensor field so as to define its covariant derivative, and it is clear that the resulting tensor field is of type $(r, s+1)$. It follows that covariant differentiation satisfies conditions analogous to (a)–(d) stipulated for absolute differentiation. Expressions for the covariant derivatives of general lower-rank tensor fields are noted below.

Covariant derivatives of lower-rank tensor fields

For a scalar field ϕ, covariant differentiation is simply partial differentiation:

$$\boxed{\phi_{;a} \equiv \partial_a \phi.} \tag{2.54}$$

For a contravariant vector field λ^a, we have

$$\boxed{\lambda^a{}_{;b} \equiv \partial_b \lambda^a + \Gamma^a_{cb}\lambda^c.} \tag{2.55}$$

For a covariant vector field μ_a, we have

$$\boxed{\mu_{a;c} \equiv \partial_c \mu_a - \Gamma^b_{ac}\mu_b.} \tag{2.56}$$

For a type $(2,0)$ tensor field τ^{ab}, we have

$$\boxed{\tau^{ab}{}_{;c} \equiv \partial_c \tau^{ab} + \Gamma^a_{dc}\tau^{db} + \Gamma^b_{dc}\tau^{ad}.} \tag{2.57}$$

For a type $(0, 2)$ tensor field τ_{ab}, we have

$$\boxed{\tau_{ab;c} \equiv \partial_c \tau_{ab} - \Gamma^d_{ac}\tau_{db} - \Gamma^d_{bc}\tau_{ad}.} \tag{2.58}$$

For a type $(1, 1)$ tensor field τ^a_b, we have

$$\boxed{\tau^a_{b;c} \equiv \partial_c \tau^a_b + \Gamma^a_{dc}\tau^d_b - \Gamma^d_{bc}\tau^a_d.} \tag{2.59}$$

Again, the mnemonic "co-below and minus" is a useful reminder for the sign of a Γ term.

The essential property common to both covariant and absolute differentiation is that when the operation is applied to a tensor field it produces a tensor field, while the operations of partial and total differentiation do not (i.e., the partial derivatives and total derivatives of tensor components do not obey transformation laws of the kind (1.73)). Another way in which the covariant derivative differs from the partial derivative is that in repeated differentiation the order matters. Thus for a vector field λ^a, we must acknowledge that even if $\lambda^a{}_{,bc} = \lambda^a{}_{,cb}$ holds, in general, $\lambda^a{}_{;bc} \neq \lambda^a{}_{;cb}$. We shall have more to say on this matter in the next chapter. We finish this section by considering the derivatives of the metric tensor field and its associated fields, noting in particular a special property that they possess.

Using equation (2.35) and the fact that $\Gamma^a_{bc} = \Gamma^a_{cb}$, we can rewrite equation (2.36) as

$$\partial_c g_{ab} - \Gamma^d_{ca} g_{db} - \Gamma^d_{cb} g_{ad} = 0,$$

which shows that $g_{ab;c} = 0$. That is, *the covariant derivative of the metric tensor field is identically zero.* The Kronecker tensor field with components δ^a_b and the contravariant metric tensor field with components g^{ab} also have covariant derivatives that are zero, as we now show. For the Kronecker tensor field, we simply note that

$$\delta^a_{b;c} = \partial_c \delta^a_b + \Gamma^a_{dc}\delta^d_b - \Gamma^d_{bc}\delta^a_d = 0 + \Gamma^a_{bc} - \Gamma^a_{bc} = 0,$$

while for the contravariant metric tensor field, we use the result just established to argue that

$$\begin{aligned} 0 = \delta^a_{b;c} &= (g^{ad} g_{db})_{;c} \\ &= g^{ad}{}_{;c} g_{db} + g^{ad} g_{db;c} \quad \text{(by Leibniz' rule)} \\ &= g^{ad}{}_{;c} g_{db} \quad \text{(as } g_{db;c} = 0\text{).} \end{aligned}$$

Then contraction with g^{be} gives

$$0 = g^{ad}{}_{;c} \delta^e_d = g^{ae}{}_{;c},$$

as claimed. Along any curve γ, where we can regard the components g_{ab} as functions of the parameter u, we have that $Dg_{ab}/du = g_{ab;c}\dot{x}^c = 0$, establishing that the absolute derivative of the metric tensor field along γ is zero. We can argue similarly that the absolute derivatives of the Kronecker tensor field and the contravariant metric tensor field are also zero along any curve.

To sum up, we have shown that *the metric tensor field g_{ab}, the Kronecker tensor field with components δ^a_b and the contravariant metric tensor field g^{ab} have covariant derivatives that are zero:*

$$\boxed{g_{ab;c} = 0, \quad \delta^a_{b;c} = 0, \quad g^{ab}{}_{;c} = 0;} \tag{2.60}$$

and that along any curve γ their absolute derivatives are also zero:

$$\boxed{Dg_{ab}/du = 0, \quad D\delta^a_b/du = 0, \quad Dg^{ab}/du = 0.} \tag{2.61}$$

These special properties of the metric tensor field and its associated fields allow us to establish the important result that *inner products are preserved under parallel transport*. What we mean by this is that if two vector fields λ^a, μ^a are parallelly transported along a curve γ, then the inner product $g_{ab}\lambda^a\mu^b$ is constant along γ. We prove this by noting that

$$\begin{aligned} d(g_{ab}\lambda^a\mu^b)/du &= D(g_{ab}\lambda^a\mu^b)/du \\ &= (Dg_{ab}/du)\lambda^a\mu^b + g_{ab}(D\lambda^a/du)\mu^b + g_{ab}\lambda^a(D\mu^b/du) \\ &= 0, \end{aligned}$$

since $D\lambda^a/du = D\mu^b/du = 0$ (because the vectors are parallelly transported) and $Dg_{ab}/du = 0$ (established above). It follows that the length of a parallelly transported vector is constant, and also that the angle between two parallelly transported vectors is constant. Since the tangent vector to an affinely parameterized geodesic is parallelly transported along the geodesic, we can deduce that *if a vector is parallelly transported along a geodesic, then the angle between the transported vector and the tangent to the geodesic remains constant.* (See the remarks after Example 2.2.1 and Exercise 2.2.3.)

Having defined covariant differentiation, we can extend the familiar notion of the divergence of a vector field in Euclidean space to vector and tensor fields on a manifold. For a contravariant vector field λ^a we define its *divergence* to be the scalar field $\lambda^a{}_{;a}$. This definition is reasonable, for in a Cartesian coordinate system in Euclidean space $g_{ij} = \delta_{ij}$, so $\partial_k g_{ij} = 0$ giving $\Gamma^i_{jk} = 0$, and $\lambda^i{}_{;i}$ reduces to $\lambda^i{}_{,i}$. The *divergence* of a covariant vector field μ_a is defined to be that of the associated contravariant vector field $\mu^a \equiv g^{ab}\mu_b$. For a type (r, s) tensor field we may define $(r + s)$ divergences,

$$\tau^{a_1\ldots c\ldots a_r}_{b_1\ldots b_s}{}_{;c}, \quad (\tau^{a_1\ldots a_r}_{b_1\ldots c\ldots b_s} g^{cd})_{;d},$$

although these will not be distinct if the tensor field possesses symmetries. We can use this approach to calculate the divergence of a vector field in Euclidean space using curvilinear coordinate systems, as in the following example.

Example 2.3.1
In spherical coordinates the position vector field is $\mathbf{r} = r\mathbf{e}_r = r\mathbf{e}_1$ (on labeling the coordinates according to $u^1 \equiv r$, $u^2 \equiv \theta$, $u^3 \equiv \phi$), so its components are $r^i = r\delta^i_1$. Its divergence can then be calculated by saying

$$\begin{aligned}
\nabla \cdot \mathbf{r} = r^i{}_{;i} &= \partial_i r^i + \Gamma^i_{ji} r^j = \partial_i(r\delta^i_1) + \Gamma^i_{ji}(r\delta^j_1) = \partial r/\partial r + r\Gamma^i_{1i} \\
&= 1 + \tfrac{1}{2} r g^{-1} \partial_1 g \qquad \text{(using equation (2.37))} \\
&= 1 + \tfrac{1}{2} r (r^4 \sin^2\theta)^{-1} \partial(r^4 \sin^2\theta)/\partial r \quad (\text{as } g = \det[g_{ij}] = r^4 \sin^2\theta) \\
&= 1 + 2 = 3.
\end{aligned}$$

Exercises 2.3

1. Check formula (2.43).
   (Most of the work was done in Exercise 2.2.6.)

2. Obtain formulae (2.49) and (2.50), using methods similar to that used in deriving the result (2.48).

3. Show that equation (2.12) for an affinely parameterized geodesic can be written as $D\dot{x}^a/du = 0$.

4. Prove that the length of the tangent vector $\dot{x}^a$ to an affinely parameterized geodesic is constant.

2.4 Geodesic coordinates

It can be seen from equation (2.13) that if we could introduce a coordinate system throughout which the metric tensor components were constant, then the connection coefficients would be zero, and the mathematics of parallel transport, absolute differentiation, and covariant differentiation would be much simpler. It is possible to introduce such a coordinate system in Euclidean space, for example, by using Cartesian coordinates in which $g_{ij} = \delta_{ij}$, but in a general curved manifold it is not. However, it is possible to introduce a system of coordinates in which $\Gamma^a_{bc} = 0$ at a given point O, and such systems have their uses in simplifying some calculations involving the connection coefficients (see, e.g., Sec. 3.2 where the Bianchi identity is established). Such coordinates are generally referred to as *geodesic coordinates* with origin O, but this is not always appropriate, as they need not be based on geodesics.

Suppose we start with some system of coordinates in which O has coordinates x^a_O. Let us define a new system of (primed) coordinates by means of the equation

$$x^{a'} \equiv x^a - x^a_O + \tfrac{1}{2}(\Gamma^a_{bc})_O(x^b - x^b_O)(x^c - x^c_O), \tag{2.62}$$

where $(\Gamma^a_{bc})_O$ are the connection coefficients at O, as given in the original (unprimed) coordinate system. Differentiation with respect to x^d gives

$$\begin{aligned} X^{a'}_d &= \delta^a_d + \tfrac{1}{2}(\Gamma^a_{bc})_O\delta^b_d(x^c - x^c_O) + \tfrac{1}{2}(\Gamma^a_{bc})_O(x^b - x^b_O)\delta^c_d \\ &= \delta^a_d + (\Gamma^a_{dc})_O(x^c - x^c_O), \end{aligned}$$

so $(X^{a'}_d)_O = \delta^a_d$ and $\det[X^{a'}_d]_O \neq 0$. This means that equation (2.62) defines a new system of coordinates in some neighborhood U' of O, as claimed (see Sec. 1.7). A second differentiation gives

$$X^{a'}_{ed} = (\Gamma^a_{dc})_O\delta^c_e = (\Gamma^a_{de})_O,$$

showing that $(X^{a'}_{ed})_O = (\Gamma^a_{de})_O$. If we now use the transformation equation of Exercise 2.2.5 (noting that $(X^d_{a'})_O = \delta^d_a$ as a consequence of $(X^{a'}_d)_O = \delta^a_d$), we get

$$(\Gamma^{a'}_{b'c'})_O = (\Gamma^d_{ef})_O\delta^a_d\delta^e_b\delta^f_c - \delta^e_b\delta^f_c(\Gamma^a_{fe})_O = (\Gamma^a_{bc})_O - (\Gamma^a_{bc})_O = 0.$$

So in the new (primed) coordinate system the connection coefficients at O are zero, and we have a system of geodesic coordinates with origin O.

Geodesic coordinates can be used to construct a system of *local Cartesian coordinates* about a point O. These are an approximation to Cartesian coordinates, valid near O in a region of limited extent where the curvature of the manifold can be neglected. To get at such a system, we make use of a second coordinate transformation that brings the metric tensor at O to a simple diagonal form, while keeping the connection coefficients at O zero. To this end, we introduce a third (double-primed) system of coordinates about O defined by

$$x^{a''} = p^a_b x^{b'}, \tag{2.63}$$

where p^a_b are constants such that the matrix $P \equiv [p^a_b]$ is nonsingular. Differentiation of equation (2.63) shows that

$$X^{a''}_{c'} = p^a_b\delta^b_c = p^a_c,$$

so that the matrix version of

$$(g_{a''b''})_O = (g_{c'd'})_O(X^{c'}_{a''})_O(X^{d'}_{b''})_O$$

is

$$G''_O = P^{\mathrm{T}} G'_O P.$$

This means that, in matrix terms, G''_O is obtained from G'_O by a similarity transformation using the matrix P (see, e.g., Birkhoff and Mac Lane, 1977, §2–6). Matrix theory tells us that there exists a matrix P that brings G''_O to diagonal form in which each diagonal entry is either $+1$ or -1. If the metric tensor is positive definite, then all these entries are $+1$, but if it is indefinite, some will be $+1$ and others will be -1. In the latter case it is usual to use a diagonalizing matrix P that gives a diagonal form for G''_O with all the positive entries preceding the negative ones, so that

$$[g_{a''b''}]_O = G''_O = \mathrm{diag}(1, \ldots, 1, -1, \ldots, -1).$$

This second transformation has $X^{a''}_{d'c'} = 0$, so it follows from the equation of Exercise 2.2.5 (adapted for primed and double-primed coordinates) that if $(\Gamma^{a'}_{b'c'})_O = 0$, then $(\Gamma^{a''}_{b''c''})_O = 0$, which is what we required of it. Note that $x^{a'}_O = 0$ (from equation (2.62)), so $x^{a''}_O = 0$ (from equation (2.63)), showing that the point O is the "origin" of the double-primed coordinate system.

Dropping the double primes, we see that about O we have introduced a system of coordinates in which

$$x^a_O = 0, \qquad (\Gamma^a_{bc})_O = 0,$$

and

$$[g_{ab}]_O = \mathrm{diag}(1, \ldots, 1, -1, \ldots, -1)$$

(where the negative entries are absent in the positive definite case), so that

$$\Gamma^a_{bc} \approx 0, \qquad [g_{ab}] \approx \mathrm{diag}(1, \ldots, 1, -1, \ldots, -1), \tag{2.64}$$

in some neighborhood of O. These are *local Cartesian coordinates*, and the extent of the region in which the approximation (2.64) is valid depends (in a way to be made precise later) on the curvature of the manifold in the vicinity of O.

The implication of this for general relativity is that about each point of spacetime we can introduce a coordinate system in which

$$\Gamma^\mu_{\nu\sigma} \approx 0, \qquad g_{\mu\nu} \approx \eta_{\mu\nu}, \tag{2.65}$$

where $[\eta_{\mu\nu}] = \mathrm{diag}(1, -1, -1, -1)$, showing that locally the spacetime of general relativity looks like that of special relativity. This observation is a key factor in our discussion of the spacetime of general relativity in the next section.

Exercise 2.4

1. Show that, as a result of the coordinate transformation leading to geodesic coordinates (equation (2.62)), $(g_{a'b'})_O = (g_{ab})_O$.

2.5 The spacetime of general relativity

The spacetime of special relativity is discussed in Appendix A. In the language of Section 1.9, it is a four-dimensional pseudo-Riemannian manifold with the property that there exist global coordinate systems in which the metric tensor takes the form

$$[\eta_{\mu\nu}] \equiv \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

and we call such coordinate systems *Cartesian.* As explained in Section 1.2, we use x^μ to label points in spacetime, where the Greek suffixes, μ, ν, etc., have the range 0, 1, 2, 3, there being a certain convenience in counting from zero rather than one. As is customary in relativity we shall frequently refer to a point in spacetime as an *event.* Cartesian coordinates are related to the more familiar coordinates, t, x, y, z of special relativity by $x^0 \equiv ct$, $x^1 \equiv x$, $x^2 \equiv y$, $x^3 \equiv z$, c being the speed of light. We may, of course, use non-Cartesian coordinates, where the metric tensor $g_{\mu\nu} \neq \eta_{\mu\nu}$ but the essential feature of the spacetime of special relativity is that we may always introduce a Cartesian coordinate system about any point, so that $g_{\mu\nu} = \eta_{\mu\nu}$, and this coordinate system is global in the sense that it covers the whole of spacetime.

One of our guiding requirements for the spacetime of general relativity is that locally it should be like the spacetime of special relativity. We therefore assume that it is a four-dimensional pseudo-Riemannian manifold with the property that about any point there exists a system of local Cartesian coordinates in which the metric tensor field $g_{\mu\nu}$ is approximately $\eta_{\mu\nu}$. Note that we do *not* assert the existence of coordinate systems in which $g_{\mu\nu} = \eta_{\mu\nu}$ exactly, and this is the essential difference between the spacetimes of general and special relativity.

As explained in the previous section, we can construct a coordinate system about any point P of general-relativistic spacetime in which $(\Gamma^\mu_{\nu\sigma})_P = 0$, and $(x^\mu)_P = (0, 0, 0, 0)$. This means that $(\partial_\sigma g_{\mu\nu})_P = 0$, and so for points near to P, where the coordinates x^μ are small, Taylor's theorem gives

$$g_{\mu\nu} \approx \eta_{\mu\nu} + \tfrac{1}{2}(\partial_\alpha \partial_\beta g_{\mu\nu})_P x^\alpha x^\beta, \tag{2.66}$$

and this approximation is valid for small x^μ.

If we are sufficiently close to P for the second term on the right of equation (2.66) to be neglected, we have a coordinate system in which $g_{\mu\nu} = \eta_{\mu\nu}$ approximately, and the extent of the region in which this approximation is valid will depend on the sizes of the second derivatives $(\partial_\alpha \partial_\beta g_{\mu\nu})_P$, and also on the accuracy of our measuring procedures. It should be stressed that in special relativity we have *global* Cartesian coordinate systems, where $g_{\mu\nu} = \eta_{\mu\nu}$ *exactly*, whereas in general relativity we have only *local* Cartesian coordinate systems of limited extent, where $g_{\mu\nu} = \eta_{\mu\nu}$ *approximately.* We distinguish the two by saying that the spacetime of special relativity is *flat*, while that of general relativity is *curved*. The above discussion shows that the departure from flatness is connected with the nonvanishing of the second derivatives $\partial_\alpha \partial_\beta g_{\mu\nu}$, and we shall see the significance of this in Chapter 3, when we give a more formal definition of flatness in terms of the curvature tensor.

The purpose of the above discussion was to show, by introducing local Cartesian coordinates, the sense in which the spacetime of general relativity is locally like that of special relativity. However, it is not sensible to work in terms of local Cartesian coordinates as these involve approximations which amount to neglecting gravity, nor is it often convenient, since more suitable coordinates may be defined in a natural way. *We therefore use general coordinates, and formulate things in ways which are valid in any coordinate system.*

Another feature of the above discussion is that it gives us a means of generalizing to general relativity results which are valid in special relativity. For example, it is shown in Appendix A that in a Cartesian coordinate system of special-relativistic spacetime, Maxwell's equations may be written in the form

$$\begin{aligned} F^{\mu\nu}{}_{,\nu} &= \mu_0 j^\mu, \\ F_{\mu\nu,\sigma} + F_{\nu\sigma,\mu} + F_{\sigma\mu,\nu} &= 0. \end{aligned} \tag{2.67}$$

where a comma denotes partial differentiation. We may adopt

$$\begin{aligned} F^{\mu\nu}{}_{;\nu} &= \mu_0 j^\mu, \\ F_{\mu\nu;\sigma} + F_{\nu\sigma;\mu} + F_{\sigma\mu;\nu} &= 0. \end{aligned} \tag{2.68}$$

where a semicolon now denotes covariant differentiation, as the general-relativistic version of these, for in a local Cartesian coordinate system (where $g_{\mu\nu} = \eta_{\mu\nu}$ approximately, and we can neglect $\Gamma^\mu_{\nu\sigma}$) equations (2.68) reduce to equations (2.67). There are really two points to note here. The first is that if any physical quantity can be defined as a Cartesian tensor in special relativity, then we can give its definition in general relativity by defining it in exactly the same way in a local Cartesian coordinate system; its components in any other coordinate system are then given by the usual transformation formulae (1.73). Given this first point, the second is that any Cartesian tensor equation valid in special relativity may be converted to an equation valid in general relativity in any coordinate system, simply by replacing partial differentiation with respect to coordinates by covariant differentiation, total derivatives along curves by absolute derivatives, and $\eta_{\mu\nu}$ by $g_{\mu\nu}$. (Compare remarks made in the Introduction.)

As an example of this, consider the path of a particle (with mass) in special relativity. Its world velocity is $u^\mu \equiv dx^\mu / d\tau$ (see Sec. A.5), where the proper time τ for the particle is defined by (see Sec. A.0)

$$c^2 d\tau^2 \equiv \eta_{\mu\nu} dx^\mu dx^\nu.$$

Its equation of motion is then (equation (A.29))

$$dp^\mu/d\tau = f^\mu,$$

where $p^\mu \equiv mu^\mu$, m being the proper mass of the particle and f^μ the 4-force acting on it. The generalization of these ideas to general relativity gives $u^\mu \equiv dx^\mu/d\tau$ as the definition of the world velocity of the particle, where now the proper time τ is defined by

$$\boxed{c^2 d\tau^2 \equiv g_{\mu\nu} dx^\mu dx^\nu} \tag{2.69}$$

and

$$Dp^\mu/d\tau = f^\mu, \tag{2.70}$$

as the equation of motion, where $p^\mu \equiv mu^\mu$, and the definitions of m and f^μ are taken over from special relativity as explained above. Moreover, these equations are valid in any coordinate system.

As in special relativity we assume that a clock measures its own proper time. In particular, if the particle is a pulsating atom, the proper time interval between events on the atom's path where successive pulses occur is constant.

In the case of a free particle for which $f^\mu = 0$, equation (2.70) reduces to $D(dx^\mu/d\tau)/d\tau = 0$, or

$$\boxed{\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\sigma}\frac{dx^\sigma}{d\tau}\frac{dx^\nu}{d\tau} = 0.} \tag{2.71}$$

This reinforces our assertion that the path of a free particle is a geodesic in spacetime, and establishes that the proper time experienced by the particle is an affine parameter along it. This result is often stated as an explicit postulate of general relativity (the *geodesic postulate*), but it emerges here as a natural consequence of the way in which we generalize special-relativistic concepts. It is a perfectly natural generalization, for the path of a free particle in the flat spacetime of special relativity is a straight line and this generalizes to a geodesic in curved spacetime.

The path of a photon (or any other zero-rest-mass particle) in the spacetime of special relativity is also a straight line, and this also generalizes to a geodesic in curved spacetime. However, there is no change in proper time along the path of a photon, so τ cannot be used as a parameter. But we can still use an affine parameter u so that the analog of equation (2.71) for a photon is

$$\boxed{\frac{d^2x^\mu}{du^2} + \Gamma^\mu_{\nu\sigma}\frac{dx^\sigma}{du}\frac{dx^\nu}{du} = 0.} \tag{2.72}$$

The fact that the photon's speed is c finds expression as

$$\boxed{g_{\mu\nu}\frac{dx^\mu}{du}\frac{dx^\nu}{du} = 0,} \tag{2.73}$$

which generalizes the relation $\eta_{\mu\nu}\dfrac{dx^\mu}{du}\dfrac{dx^\nu}{du} = 0$ (equivalent to $c^2dt^2 - dx^2 - dy^2 - dz^2 = 0$) of special relativity.

We have already remarked on the characterization of a vector λ^μ as

$$\begin{cases} \text{timelike} \\ \text{null} \\ \text{spacelike} \end{cases} \text{if } g_{\mu\nu}\lambda^\mu\lambda^\nu \begin{cases} > 0 \\ = 0 \\ < 0 \end{cases}$$

(see Sec. 1.9). One should note that at any point of spacetime the null cone of vectors given by $g_{\mu\nu}\lambda^\mu\lambda^\nu = 0$ lies in the tangent space at that point and not in the manifold. This fact is not readily appreciated in the flat spacetime of special relativity, because its basic linear structure allows one to regard the tangent space at each point as being embedded in the spacetime.

At any point on the path of a particle (with mass) its world velocity is a tangent vector to the path, and equation (2.69) tells us that this tangent vector is timelike. So a particle with mass follows a timelike path through spacetime, and in particular a free particle follows a timelike geodesic. A photon, however, follows a null geodesic, as equation (2.73) tells us that the tangent vectors to its path are null. Spacelike paths and spacelike geodesics may also be defined, but these have no physical significance.[8]

In moving from the flat spacetime of special relativity to the curved spacetime of general relativity we hope somehow to incorporate the effects of gravity, and the point of view we are adopting is that gravity is not a force, and that gravitational effects may be explained in terms of the curvature of spacetime. It should therefore be understood that by free particles we mean particles moving under gravity alone. Comparing equation (2.71) with its special-relativistic analog $d^2x^\mu/d\tau^2 = 0$ indicates that the connection coefficients play an important role in explaining gravitational effects. Since these are given by derivatives of the metric tensor field, we see that it is this tensor field which, in a sense, carries the gravitational content of spacetime. For the moment we shall take the metric tensor field as given, and postpone until Chapter 3 the question of how it is determined by the distribution of matter and energy in spacetime. In the rest of this chapter we take a closer look at equations (2.70) and (2.71), and relate them to some familiar Newtonian ideas.

Exercises 2.5

1. Is the world velocity of a stationary chair (in the lab) timelike or spacelike? Is its world line a geodesic?

2. Deduce the geodesic equation (2.71) from equation (2.70).

[8] Unless one believes in tachyons.

2.6 Newton's laws of motion

Newton's first law that "every body perseveres in its state of rest, or of uniform motion in a right [straight] line, unless it is compelled to change that state by forces impressed thereon" clearly has its counterpart in the statement that "every particle follows a geodesic in spacetime."[9] Indeed, in a local inertial coordinate system where we may neglect the $\Gamma^{\mu}_{\nu\sigma}$, the geodesic equation reduces to $d^2x^\mu/d\tau^2 = 0$. For nonrelativistic speeds $d\tau/dt$ is approximately one, so the geodesic equation yields $d^2x^i/dt^2 = 0$ $(i = 1, 2, 3)$, the familiar Newtonian equation of motion of a free particle. Newton's second law that

Newton	Einstein
Free particles move in straight lines through space.	Free particles follow geodesics through spacetime.
$\mathbf{F} = m\dfrac{d^2\mathbf{x}}{dt^2}$	$f^\mu = m\left(\dfrac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\sigma}\dfrac{dx^\nu}{d\tau}\dfrac{dx^\sigma}{d\tau}\right)$
To every action there is always opposed an equal reaction.	The third law is true for non-gravitational forces, just as in Newtonian physics (but see text for gravitational interaction).

Table 2.1. Newton's laws and their relativistic counterparts.

"the alteration of motion is ever proportional to the motive force impressed; and is made in the right line in which that force is impressed" is usually rendered as the 3-vector equation

$$d\mathbf{p}/dt = \mathbf{F},$$

where $\mathbf{p}$ is the momentum and $\mathbf{F}$ the applied force. This clearly has its counterpart in equation (2.70).

Newton's third law that "to every action there is always opposed an equal reaction: or the mutual actions of two bodies upon each other are always equal, and directed to contrary parts" is true in general relativity also. However, we must be careful, because Newton's gravitational force is now replaced by Einstein's idea that a massive body causes curvature of the spacetime around it, and a free particle responds by moving along a geodesic in that spacetime. It should be noted that this viewpoint ignores any curvature produced by the particle following the geodesic. That is, the particle is a *test particle*, and there

[9] The versions of Newton's laws quoted here are from Andrew Motte's translation (London, 1729) of Newton's *Principia*.

is no question of its having any effect on the body producing the gravitational field.

The gravitational interaction of two large bodies is not directly addressed by Einstein's theory, although it is of importance in astronomy, as for example in the famous pair of orbiting neutron stars PSR 1913+16. Approximation methods for such cases were studied in the 1980s,[10] but are beyond the scope of our book.

2.7 Gravitational potential and the geodesic

Suppose we have a coordinate system in which the metric tensor field is given by

$$g_{\mu\nu} \equiv \eta_{\mu\nu} + h_{\mu\nu}, \tag{2.74}$$

where the $h_{\mu\nu}$ are small, but not so small that they may be neglected. Our aim in this section is to obtain a Newtonian approximation to the geodesic equation given by the metric tensor field (2.74) valid for a particle whose velocity components dx^i/dt $(i = 1, 2, 3)$ are small compared with c. We shall assume that the gravitational field, as expressed by $h_{\mu\nu}$, is quasi-static in the sense that $\partial_0 h_{\mu\nu} \equiv c^{-1}\partial h_{\mu\nu}/\partial t$ is negligible when compared with $\partial_i h_{\mu\nu}$.

If instead of the proper time τ we use the coordinate time t (defined by $x_0 \equiv ct$) as a parameter, then the geodesic equation giving the path of a free particle has the form

$$\frac{d^2x^\mu}{dt^2} + \Gamma^\mu_{\nu\sigma}\frac{dx^\nu}{dt}\frac{dx^\sigma}{dt} = h(t)\frac{dx^\mu}{dt}, \tag{2.75}$$

where

$$h(t) \equiv -\frac{d^2t}{d\tau^2}\left(\frac{dt}{d\tau}\right)^{-2} = \frac{d^2\tau}{dt^2}\left(\frac{d\tau}{dt}\right)^{-1}. \tag{2.76}$$

This can be deduced by an argument like that used in Exercise 2.1.1 and by noting that $\dfrac{d}{d\tau}\left(\dfrac{dt}{d\tau}\right) = \dfrac{dt}{d\tau}\dfrac{d}{dt}\left(\dfrac{d\tau}{dt}\right)^{-1}$. On dividing by c^2, the spatial part of equation (2.75) may be written

$$\frac{1}{c^2}\frac{d^2x^i}{dt^2} + \Gamma^i_{00} + 2\Gamma^i_{0j}\left(\frac{1}{c}\frac{dx^j}{dt}\right) + \Gamma^i_{jk}\left(\frac{1}{c}\frac{dx^j}{dt}\right)\left(\frac{1}{c}\frac{dx^k}{dt}\right) = \frac{1}{c}h(t)\left(\frac{1}{c}\frac{dx^i}{dt}\right), \tag{2.77}$$

and the last term on the left is clearly negligible.

If we put $h^{\mu\nu} \equiv \eta^{\mu\sigma}\eta^{\nu\rho}h_{\sigma\rho}$, then a short calculation shows that, to first order in the small quantities $h_{\mu\nu}$ and $h^{\mu\nu}$,

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu} \quad \text{and} \quad \Gamma^\mu_{\nu\sigma} = \tfrac{1}{2}\eta^{\mu\rho}(\partial_\nu h_{\sigma\rho} + \partial_\sigma h_{\nu\rho} - \partial_\rho h_{\nu\sigma}). \tag{2.78}$$

[10] See Damour and Deruelle, 1986.

So to first order,

$$\begin{aligned}\Gamma^i_{00} &= \tfrac{1}{2}\eta^{i\rho}(\partial_0 h_{0\rho} + \partial_0 h_{0\rho} - \partial_\rho h_{00}) \\ &= -\tfrac{1}{2}\eta^{ij}\partial_j h_{00} = \tfrac{1}{2}\delta^{ij}\partial_j h_{00},\end{aligned}$$

on neglecting $\partial_0 h_{\mu\nu}$ in comparison with $\partial_i h_{\mu\nu}$. Also to first order,

$$\begin{aligned}\Gamma^i_{0j} &= \tfrac{1}{2}\eta^{i\rho}(\partial_0 h_{j\rho} + \partial_j h_{0\rho} - \partial_\rho h_{0j}) \\ &= -\tfrac{1}{2}\delta^{ik}(\partial_j h_{0k} - \partial_k h_{0j}),\end{aligned}$$

again on neglecting $\partial_0 h_{\mu\nu}$.

We have now approximated all the terms on the left-hand side of equation (2.77), and there remains the right-hand side to deal with. Working to the same level of approximation as above, and neglecting squares and products of $c^{-1}dx^i/dt$, we find from

$$\left(\frac{d\tau}{dt}\right)^2 = \frac{1}{c^2}g_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}$$

that

$$d\tau/dt = (1 + h_{00})^{1/2} = 1 + \tfrac{1}{2}h_{00}, \tag{2.79}$$

so

$$d^2\tau/dt^2 = \tfrac{1}{2}ch_{00,0}$$

and

$$\frac{1}{c}h(t) = \tfrac{1}{2}h_{00,0}(1 - \tfrac{1}{2}h_{00}) = \tfrac{1}{2}h_{00,0}$$

from equation (2.76).

It follows that the right-hand side of equation (2.77) is negligible, and our approximation gives

$$\frac{1}{c^2}\frac{d^2x^i}{dt^2} + \tfrac{1}{2}\delta^{ij}\partial_j h_{00} - \delta^{ik}(\partial_j h_{0k} - \partial_k h_{0j})\frac{1}{c}\frac{dx^j}{dt} = 0.$$

Introducing the mass m of the particle and rearranging gives

$$m\frac{d^2x^i}{dt^2} = -m\delta^{ij}\partial_j(\tfrac{1}{2}c^2 h_{00}) + mc\delta^{ik}(\partial_j h_{0k} - \partial_k h_{0j})\frac{dx^j}{dt}. \tag{2.80}$$

Let us now interpret this in Newtonian terms. The left-hand side is mass × acceleration, so the right-hand side is the "gravitational force" on the particle. The first term on the right is the force $-m\nabla V$ arising from a potential V given by $V \equiv \frac{1}{2}c^2 h_{00}$, while the second term on the right is velocity-dependent and clearly smacks of rotation.[11] This is not surprising, for the principle of equivalence asserts that the forces of acceleration, such as the velocity-dependent

[11] Most authors assume that the derivatives $\partial_i h_{\mu\nu}$ are small along with the $h_{\mu\nu}$, and therefore do not obtain these velocity-dependent rotational terms. However, the fact that the $h_{\mu\nu}$ are small does not mean that their derivatives are also small (see Sec. 2.9).

Coriolis force[12] which would arise from using a rotating reference system, are on the same footing as gravitational forces. If we agree to call a nearly inertial coordinate system in which $\partial_j h_{0k} - \partial_k h_{0j}$ is zero *nonrotating*, then we have for a slowly moving particle in a nearly inertial, nonrotating, coordinate system, in which the quasi-static condition holds, the approximation

$$d^2x^i/dt^2 = -\delta^{ij}\partial_j V, \tag{2.81}$$

where

$$V \equiv \tfrac{1}{2}c^2 h_{00} + \text{const.} \tag{2.82}$$

This is the Newtonian equation of motion for a particle moving in a gravitational field of potential V, provided we make the identification (2.82). This gives

$$g_{00} = 2V/c^2 + \text{const},$$

and if we choose the constant to be 1, then g_{00} reduces to its flat spacetime value when $V = 0$. This gives

$$\boxed{g_{00} = 1 + 2V/c^2} \tag{2.83}$$

as the relation between g_{00} and the Newtonian potential V in this approximation.

Exercises 2.7

1. Check approximations (2.78) and (2.79).

2. Show that equation (2.81) is equivalent to $m\mathbf{a} = \mathbf{F} = -m\nabla V$, where $\mathbf{a}$ is the acceleration and $\mathbf{F}$ is the force on the particle.

2.8 Newton's law of universal gravitation

Newton's law of universal gravitation does not survive intact in general relativity, which is after all a new theory replacing the Newtonian theory. However, we should be able to recover it as an approximation.

The Schwarzschild solution is an exact solution of the field equations of general relativity, and it may be identified as representing the field produced by a massive body. This solution is derived in the next chapter (see Sec. 3.7), and its line element is

$$c^2d\tau^2 = (1 - 2GM/rc^2)c^2dt^2 - (1 - 2GM/rc^2)^{-1}dr^2 - r^2d\theta^2 - r^2\sin^2\theta\, d\phi^2,$$

[12] See, for example, Goldstein, Poole, and Safko, 2002, §4–10.

where M is the mass of the body and G the gravitational constant. For small values of GM/rc^2 this is close to the line element of flat spacetime in spherical coordinates, and r then behaves like radial distance. If we were to put

$$x^0 \equiv ct, \quad x^1 \equiv r \sin\theta \cos\phi, \quad x^2 \equiv r \sin\theta \sin\phi, \quad x^3 \equiv r\cos\theta,$$

we would obtain a line element whose metric tensor had the form $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, where, for large values of rc^2/GM, the $h_{\mu\nu}$ are small and $g_{00} = 1 - 2GM/rc^2$. This gives $h_{00} = -2GM/rc^2$, and according to the results of the last section, a Newtonian potential $V = -GM/r$. The 3-vector form of equation (2.81) gives

$$m\, d^2\mathbf{r}/dt^2 = -m\nabla V = -GMmr^{-2}\hat{\mathbf{r}},$$

where $\mathbf{r} \equiv (x^1, x^2, x^3)$, m is the mass of the test particle, and $\hat{\mathbf{r}}$ is a unit vector in the direction of $\mathbf{r}$. The "force" on the test particle is in agreement with that given by Newton's law, and in this way the law is recovered as an approximation valid for large values of rc^2/GM and slowly moving particles.

2.9 A rotating reference system

The principle of equivalence (see the Introduction) implies that the "fictitious" forces of accelerating coordinate systems are essentially in the same category as the "real" forces of gravity. Put another way, if the geodesic equation contains gravity in the $\Gamma^{\mu}_{\nu\sigma}$ it must also contain any accelerations which may have been built in by choice of coordinate system. In a curved spacetime it is not always easy, and often impossible, to sort these forces out, but in flat spacetime we have only the fictitious forces of acceleration and these should be included in the $\Gamma^{\mu}_{\nu\sigma}$. As an example of this, let us consider a rotating reference system in flat spacetime.

Starting with a nonrotating system K with coordinates (T, X, Y, Z) and line element

$$c^2 d\tau^2 = c^2 dT^2 - dX^2 - dY^2 - dZ^2, \tag{2.84}$$

let us define new coordinates (t, x, y, z) by (see Fig. 2.6)

$$\begin{aligned} T &\equiv t, \\ X &\equiv x\cos\omega t - y \sin\omega t, \\ Y &\equiv x \sin\omega t + y\cos\omega t, \\ Z &\equiv z. \end{aligned} \tag{2.85}$$

Note that at this point we are only defining a change of coordinates and we are not too concerned (yet) about their physical meanings. Points given by x, y, z constant rotate with angular speed ω about the Z axis of K, and this

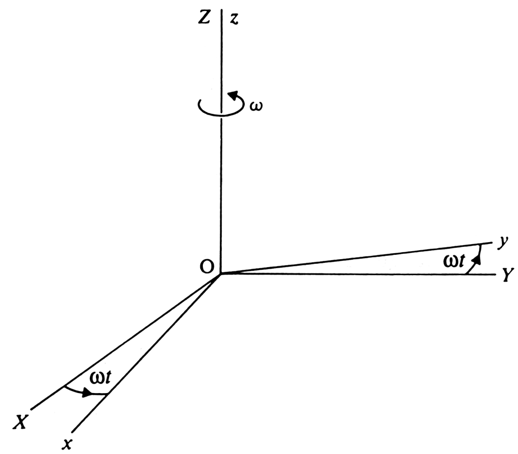

Fig. 2.6. Coordinate system $K'(t, x, y, z)$ rotating relative to the coordinate system $K(T, X, Y, Z)$.

defines the rotating system K' (see Fig. 2.6). In terms of the new coordinates the line element is

$$c^2 d\tau^2 = [c^2 - \omega^2(x^2 + y^2)]dt^2 + 2\omega y\, dx\, dt - 2\omega x\, dy\, dt - dx^2 - dy^2 - dz^2, \quad (2.86)$$

and the geodesic equations are

$$\begin{aligned} \ddot{t} &= 0, \\ \ddot{x} - \omega^2 x \dot{t}^2 - 2\omega \dot{y}\dot{t} &= 0, \\ \ddot{y} - \omega^2 y \dot{t}^2 + 2\omega \dot{x}\dot{t} &= 0, \\ \ddot{z} &= 0. \end{aligned} \quad (2.87)$$

where dots denote differentiation with respect to proper time (see Exercises 2.9.1 and 2.9.2). These constitute the equation of motion of a free particle (with mass).

The first of equations (2.87) implies that $dt/d\tau$ is constant, so the remaining equations may be written as

$$\begin{aligned} d^2x/dt^2 - \omega^2 x - 2\omega\, dy/dt &= 0, \\ d^2y/dt^2 - \omega^2 y + 2\omega\, dx/dt &= 0, \\ d^2z/dt^2 &= 0. \end{aligned}$$

Introducing the mass m of the particle and rearranging gives

$$\begin{aligned} &m\, d^2x/dt^2 = m\omega^2 x + 2m\omega\, dy/dt, \\ &m\, d^2y/dt^2 = m\omega^2 y - 2m\omega\, dx/dt, \\ &m\, d^2z/dt^2 = 0. \end{aligned} \quad (2.88)$$

or, in 3-vector notation,

$$m\,d^2\mathbf{r}/dt^2 = -m\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) - 2m\boldsymbol{\omega}\times(d\mathbf{r}/dt), \tag{2.89}$$

where $\mathbf{r} \equiv (x, y, z)$ and $\boldsymbol{\omega} \equiv (0, 0, \omega)$.

An observer using t for time would interpret the left-hand side of equation (2.89) as mass $\times$ acceleration, and would therefore assert the existence of a "gravitational force" as given by the right-hand side. This "force" is, of course, the sum of the centrifugal force $-m\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r})$ and the Coriolis force $-2m\boldsymbol{\omega}\times(d\mathbf{r}/dt)$, and this would seem to bear out our assertion that the geodesic equation does indeed include the forces of acceleration in the $\Gamma^{\mu}_{\nu\sigma}$. However, such an observer would be using the time associated with the nonrotating system K, because $t \equiv T$ and T is the time measured by clocks at rest in K. It is possible to define a time for K' based on a system of clocks at rest in K', but we shall not follow that course, as it would involve replacing equations (2.88) and (2.89) by more complicated ones that tend to conceal the Coriolis and centrifugal forces. Note that t is *exactly* the proper time for an observer situated at the common origin O of the two systems, so observers close to O who are at rest in the rotating system would accept equations (2.88) and (2.89) as approximately valid and recognize the terms on the right as forces of acceleration.

We can relate the situation described above to the approximation methods of Section 2.7 by putting $x^0 \equiv ct$, $x^1 \equiv x$, $x^2 \equiv y$, $x^3 \equiv z$, and noting that the line element (2.86) then gives $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, where

$$[h_{\mu\nu}] \equiv \begin{bmatrix} -\omega^2(x^2+y^2)/c^2 & \omega y/c & -\omega x/c & 0 \\ \omega y/c & 0 & 0 & 0 \\ -\omega x/c & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The $h_{\mu\nu}$ are small, provided we restrict ourselves to the region near the z axis where $\omega^2(x^2+y^2)/c^2$ is small. Moreover, $\partial_0 h_{\mu\nu} = 0$, so the quasi-static condition is fulfilled. However, our system is rotating, so we must use the approximation (2.80) rather than (2.81). We see that

$$\tfrac{1}{2}c^2 h_{00} = -\tfrac{1}{2}\omega^2(x^2+y^2),$$

and a straightforward calculation (see Exercise 2.9.4) gives

$$[A^i_j] \equiv \begin{bmatrix} 0 & 2\omega & 0 \\ -2\omega & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{2.90}$$

where $A^i_j \equiv c\delta^{ik}(\partial_j h_{0k} - \partial_k h_{0j})$. Hence the approximation (2.80) gives

$$\begin{aligned} m\,d^2x/dt^2 &= m\omega^2 x + 2m\omega\,dy/dt, \\ m\,d^2y/dt^2 &= m\omega^2 y - 2m\omega\,dx/dt, \\ m\,d^2z/dt^2 &= 0. \end{aligned}$$

These equations are identical with equations (2.88), and may be rearranged to exhibit the centrifugal and Coriolis forces, as before.

Exercises 2.9

1. Check the form of the line element (2.86) and verify that

$$[g^{\mu\nu}] = c^{-2} \begin{bmatrix} c^2 & \omega y c & -\omega x c & 0 \\ \omega y c & \omega^2 y^2 - c^2 & -\omega^2 x y & 0 \\ -\omega x c & -\omega^2 x y & \omega^2 x^2 - c^2 & 0 \\ 0 & 0 & 0 & -c^2 \end{bmatrix}.$$

.

2. Obtain the geodesic equations (2.87) in three different ways:
   (a) By using the Euler–Lagrange equations (and $[g^{\mu\nu}]$ from Exercise 1).
   (b) By extracting $[g_{\mu\nu}]$ from the line element (2.86), and then calculating the $\Gamma^{\mu}_{\nu\sigma}$ (again using $[g^{\mu\nu}]$ from Exercise 1).
   (c) By substituting for T, X, Y, Z in $\ddot{T} = \ddot{X} = \ddot{Y} = \ddot{Z} = 0$, using equations (2.85).

3. Cylindrical coordinates (ρ, ϕ, z) may be introduced into the rotating system K' by putting $x \equiv \rho \cos\phi$, $y \equiv \rho \sin\phi$. Show that in terms of these the geodesic equations are

$$\begin{aligned} \ddot{t} &= 0, \\ \ddot{\rho} - \rho\omega^2 \dot{t}^2 - \rho\dot{\phi}^2 - 2\omega\rho\dot{\phi}\dot{t} &= 0, \\ \ddot{\phi} + 2\rho^{-1}\dot{\rho}\dot{\phi} + 2\omega\rho^{-1}\dot{\rho}\dot{t} &= 0, \\ \ddot{z} &= 0, \end{aligned}$$

   so that corresponding to equations (2.88) one has

$$\begin{aligned} & m\left[\frac{d^2\rho}{dt^2} - \rho\left(\frac{d\phi}{dt}\right)^2\right] = m\rho\omega^2 + 2m\omega\rho\frac{d\phi}{dt}, \\ & m\left(\rho\frac{d^2\phi}{dt^2} + 2\frac{d\rho}{dt}\frac{d\phi}{dt}\right) = -2m\omega\frac{d\rho}{dt}, \\ & m\frac{d^2 z}{dt^2} = 0. \end{aligned}$$

   Interpret these in terms of the radial, transverse, and axial components of acceleration, centrifugal, and Coriolis forces.

4. Check that the matrix $[A^i_j]$ is as given by equation (2.90), and that the approximation (2.80) does give equations (2.88).

Problems 2

1. Obtain the geodesic equations (using arc-length s as a parameter) for the hyperbolic paraboloid of Example 1.6.1.

   Deduce that all parametric curves are geodesics.

2. Using polar coordinates $\rho \equiv u^1$, $\phi \equiv u^2$, obtain the geodesic equations for the plane and verify that the ray $\phi = \phi_0$ (ϕ_0 = constant) is a geodesic. Use the equations of parallel transport to show that if λ^A is parallelly transported along this ray from its initial value λ_0^A at (ρ_0, ϕ_0), then
$$\lambda^1 = \lambda_0^1, \quad \lambda^2 = (\rho_0/\rho)\lambda_0^2.$$
   Verify that its length is constant and that it makes a constant angle with the ray.

3. If in spherical coordinates we set $\theta = \theta_0$, where θ_0 is a constant between 0 and $\pi/2$, we get a cone, and the remaining coordinates (r, ϕ) act as parameters on the cone. Show that the line element of the cone is
$$ds^2 = dr^2 + \omega^2 r^2 d\phi^2,$$
   where $\omega \equiv \sin\theta_0$, and that the Euler–Lagrange equations for geodesics on the cone yield
$$\ddot{r} - \omega^2 r \dot{\phi}^2 = 0, \quad \dot{\phi} = k/r^2,$$
   where k is constant. By eliminating the parameter, show that the geodesics satisfy
$$\frac{d}{d\phi}\left(\frac{1}{r^2}\frac{dr}{d\phi}\right) = \frac{\omega^2}{r}.$$
   Use the substitution $u = 1/r$ to solve this equation and hence show that the geodesics are given by $1 = Ar\cos(\omega\phi) + Br\sin(\omega\phi)$, where A and B are constants of integration.

   Use this result to show that (as intuition suggests) if the cone is cut along a generator and flattened to lie in a plane, then the geodesics are straight lines on the resulting flat surface.

4. The *curl* of a covariant field λ_a is the skew-symmetric tensor field A_{ab} defined by
$$A_{ab} \equiv \lambda_{a;b} - \lambda_{b;a}.$$
   Show that $A_{ab} = \lambda_{a,b} - \lambda_{b,a}$.

5. If A_{ab} is a skew-symmetric type $(0, 2)$ tensor field, prove that
$$B_{abc} \equiv A_{ab,c} + A_{bc,a} + A_{ca,b}$$
   are the components of a type $(0, 3)$ tensor field.

   (Hint: Put $A_{ab,c} = A_{ab;c} + \Gamma^d_{ac} A_{db} + \Gamma^d_{bc} A_{ad}$.)

6. Show that when spherical coordinates are used the line element of flat spacetime is

$$c^2 d\tau^2 = c^2 dt^2 - dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta\, d\phi^2.$$

7. The line element of a static spherically symmetric spacetime is

$$c^2 d\tau^2 = A(r)dt^2 - B(r)dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta\, d\phi^2.$$

Use the Euler–Lagrange equations to obtain the geodesic equations, and hence show that the only nonvanishing connection coefficients are:

$$\begin{array}{lll} \Gamma^0_{01} = A'/2A, & \Gamma^1_{00} = A'/2B, & \Gamma^1_{11} = B'/2B, \\ \Gamma^1_{22} = -r/B, & \Gamma^1_{33} = -(r\sin^2\theta)/B, & \Gamma^2_{12} = 1/r, \\ \Gamma^2_{33} = -\sin\theta\cos\theta, & \Gamma^3_{13} = 1/r, & \Gamma^3_{23} = \cot\theta, \end{array}$$

where primes denote derivatives with respect to r, and

$$x^0 \equiv t, \quad x^1 \equiv r, \quad x^2 \equiv \theta, \quad x^3 \equiv \phi.$$

8. One can conceive of an observer in a swivel chair located above the Sun, looking down on the plane of the Earth's orbit. If the chair rotates at the rate of one revolution a year, then to the observer the Earth appears stationary. If for some reason all heavenly bodies other than the Earth and the Sun are invisible, how does the observer explain why the Earth does not collapse in towards the Sun, there being no detectable orbit?

http://www.springer.com/978-0-387-26078-5

A Short Course in General Relativity
Foster, J.; Nightingale, J.D.
2006, X, 292 p. 51 illus., Softcover
ISBN: 978-0-387-26078-5