2.5 Spacetime curvature

Figure 2.4 repeats the previous example of non-Euclidean geometry on the surface of a sphere. We have the triangle ABC of Figure 2.2(a) whose three angles are each 90°. Consider what happens to a vector (shown by a dotted arrow) as it is parallelly transported along the three sides of this triangle. As shown in the figure, this vector is originally perpendicular to AB when it starts its journey at A. When it reaches B it lies along CB. So it keeps pointing along this line as it moves from B to C. At C it is again perpendicular to AC. So, as it moves along CA from C to A, it maintains this perpendicularity with the result that when it arrives at A it is pointing along AB. In other words, one circuit around this triangle has resulted in a change of direction of the vector by 90°, although at each stage it was being moved parallel to itself!

A similar experiment with a triangle drawn on a flat piece of paper will tell us that there is no resulting change in the direction of the vector when it moves parallel to itself around the triangle. So our physical triangle behaves differently from the flat Euclidean triangle.

The phenomenon illustrated in Figure 2.4 can also be described as follows. If we had moved our vector from A to C along two different routes – along AC and along AB followed by BC – we would have found it pointing in two different directions. In fact, if we had taken any arbitrary curves from A to C we would have found that the parallel transport of a vector from A to C varies from curve to curve; that is, the outcome depends on the path of transport from A to C.

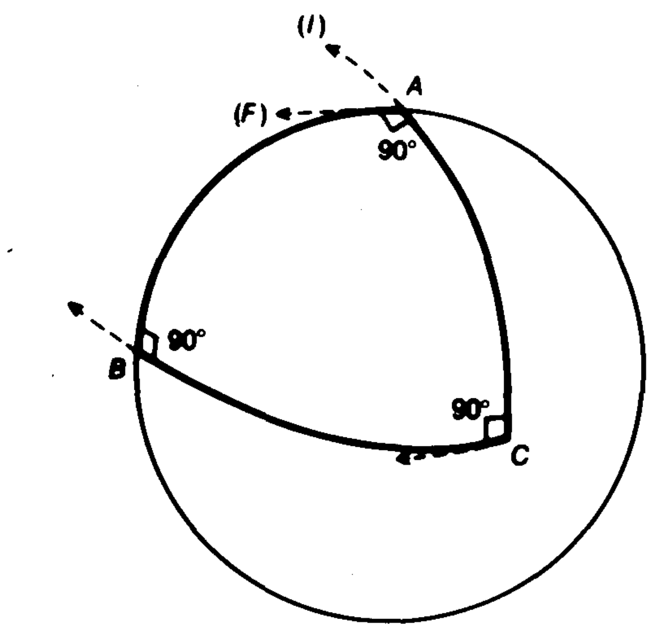

Fig. 2.4 Parallel transport on a spherical surface.

This is one of the properties that distinguishes a curved space from a flat space. Let us consider it in more general terms for our four-dimensional spacetime. Let a vector B_i at P be transported parallelly to Q and let us ask for the condition that the answer should be *independent* of the curve joining P to Q. We have seen that under parallel transport from a point $\{x^i\}$ to a neighbouring point $\{x^i + \delta x^i\}$, the components of the vector change according to (2.32). If it were possible to transport B_i from P to Q without the result depending on which path is taken, then we would be able to generate a vector field $B^i(x^k)$, satisfying the differential equation

$$\frac{\partial B_i}{\partial x^k} = \Gamma^l_{ik} B_l. \tag{2.44}$$

So the answer to our question depends on whether we can find a nontrivial solution to (2.44).

The necessary condition for the existence of a solution is easily derived. We differentiate (2.44) with respect to x^n to get

$$\begin{aligned}\frac{\partial^2 B_i}{\partial x^n \partial x^k} &= \frac{\partial}{\partial x^n}(\Gamma^l_{ik} B_l) = \frac{\partial \Gamma^l_{ik}}{\partial x^n} B_l + \Gamma^l_{ik} \frac{\partial B_l}{\partial x^n} \\ &= \left(\frac{\partial \Gamma^m_{ik}}{\partial x^n} + \Gamma^l_{ik}\Gamma^m_{ln}\right) B_m.\end{aligned}$$

We now interchange the order of differentiation with respect to x^n and x^k and use the result $B_{i,nk} = B_{i,kn}$. We then get the required necessary condition as

$$R_i{}^m{}_{kn} \equiv \frac{\partial \Gamma^m_{ik}}{\partial x^n} - \frac{\partial \Gamma^m_{in}}{\partial x^k} + \Gamma^l_{ik}\Gamma^m_{ln} - \Gamma^l_{in}\Gamma^m_{lk} = 0. \tag{2.45}$$

It is not obvious simply from the above expression that $R_i{}^m{}_{kn}$ should be a tensor. Yet our result, in order to be significant, must clearly hold whatever coordinates we employ to derive it. So we do expect $R_i{}^m{}_{kn}$ to be a tensor. A simple calculation shows that, for any twice differentiable vector field B_i,

$$B_{i;nk} - B_{i;kn} \equiv R_i{}^m{}_{kn} B_m. \tag{2.46}$$

Since the left-hand side is a tensor, so is the right-hand side, and, B_m being an arbitrary vector, we have by the *quotient law* (see Exercise 10) the result that $R_i{}^m{}_{kn}$ are the components of a tensor.

This tensor, known as the *Riemann Christoffel tensor* (or, more commonly, the *Riemann tensor*), plays an important role in specifying the geometrical properties of spacetime. Although we have derived (2.45) as a necessary condition, a slightly more sophisticated technique shows that

(2.45) is also the sufficient condition that a vector field $B_i(x^k)$ can be defined over the spacetime by parallel transport.

Spacetime is said to be *flat* if its Riemann tensor vanishes everywhere. Otherwise, it is said to be *curved*. Exercises 26 and 27 illustrate two other ways in which this tensor distinguishes the properties of a curved spacetime from those of a flat spacetime.

2.5.1 Symmetries of R_{iklm}

It is more convenient to lower the second index of the Riemann tensor to study its symmetry properties. Since the symmetry or antisymmetry of a tensor does not depend on what coordinates are used, it is more convenient to write (2.45) in the locally inertial coordinates (2.43). We then get

$$R_{iklm} = \tfrac{1}{2}(g_{kl,im} + g_{im,kl} - g_{km,il} - g_{il,km}). \tag{2.47}$$

From this expression the following symmetries are immediately obvious:

$$R_{iklm} = -R_{kilm} = -R_{ikml} = -R_{lmik}. \tag{2.48}$$

We also get relations of the following type:

$$R_{iklm} + R_{imkl} + R_{ilmk} \equiv 0. \tag{2.49}$$

If we take all these symmetries into account, we find that of the $4^4 = 256$ components of the Riemann tensor, only 20 at most are independent! Moreover, we will soon see that there are identities linking their derivatives.

2.5.2 The Ricci and Einstein tensors

By the process of contraction we can construct lower rank tensors from R_{iklm}. The tensor

$$R_{kl} = g^{im} R_{iklm} \equiv R^m{}_{klm} \tag{2.50}$$

is called the *Ricci tensor*. If we use the locally inertial coordinate system, we see immediately that

$$R_{kl} = R_{lk}. \tag{2.51}$$

Owing to the symmetries of (2.48), there are no other independent second-rank tensors that can be constructed out of R_{iklm}.

By further contraction we get a scalar:

$$R = R_{ik} \equiv R^k_k. \tag{2.52}$$

R is called the *scalar curvature*. The tensor

$$G_{ik} \equiv g^{kl}R_{kl} - \tfrac{1}{2}g_{ik}R \qquad (2.53)$$

will turn out to have a special role to play in Einstein's general relativity. This tensor is called the *Einstein tensor*.

2.5.3 Bianchi identities

The expression (2.47) suggests another symmetry for the components of R_{iklm}. This symmetry is not algebraic, but involves calculus. In covariant language we may express it as follows:

$$R_{iklm;n} + R_{iknl;m} + R_{ikmn;l} \equiv 0. \qquad (2.54)$$

These relations are known as the *Bianchi identities*. Their proof is most easily given in the locally inertial system, as in (2.47).

But multiplying (2.54) by $g^{im}g^{kn}$, and using (2.50)–(2.52), we can deduce from these identities another that is of importance to relativity:

$$(R^{ik} - \tfrac{1}{2}g^{ik}R)_{;k} \equiv 0. \qquad (2.55)$$

In other words, *the Einstein tensor G^{ik} has zero divergence.*

2.6 Geodesics

So far we have talked about non-Euclidean geometries without mentioning whether they have the equivalents of straight lines in Euclidean geometry. We now show how equivalent concepts do exist in the Riemannian geometry under consideration here.

There are two properties of a straight line that can be generalized: the property of 'straightness' and the property of 'shortest distance'. Straightness means that as we move along the line, its direction does not change. Let us see how we can generalize this concept first.

Let $x^i(\lambda)$ be the parametric representation of a curve in spacetime. Its tangent vector is given by

$$u^i = \frac{\mathrm{d}x^i}{\mathrm{d}\lambda}. \qquad (2.56)$$

Our straightness criterion demands that u^i should not change along the curve. In going from λ to $\lambda + \delta\lambda$, the change in u^i is given by

$$\Delta u^i = \frac{\mathrm{d}u^i}{\mathrm{d}\lambda}\,\delta\lambda + \Gamma^i_{kl}u^k\,\delta x^l.$$

The second expression on the right-hand side arises from the change produced by parallel transport through a coordinate displacement δx^l.

However, $\delta x^l = u^l \, \delta\lambda$. Therefore the condition of no change of direction u^i implies $\Delta u^i = 0$; that is,

$$\frac{\mathrm{d}u^i}{\mathrm{d}\lambda} + \Gamma^i_{kl} u^k u^l = 0. \tag{2.57}$$

This is the condition that our curve must satisfy in order to be straight.

The second property of a straight line in Euclidean geometry is that it is the curve of shortest distance between two points. Let us generalize this property in the following way. Let the curve, parametrized by λ, connect two points P_1 and P_2 of spacetime, with parameters λ_1 and λ_2 respectively. Then the 'distance' of P_2 from P_1 is defined as

$$s(P_2, P_1) = \int_{\lambda_1}^{\lambda_2} \left(g_{ik} \frac{\mathrm{d}x^i}{\mathrm{d}\lambda} \frac{\mathrm{d}x^k}{\mathrm{d}\lambda} \right)^{1/2} \mathrm{d}\lambda \equiv \int_{\lambda_1}^{\lambda_2} L \, \mathrm{d}\lambda. \tag{2.58}$$

We now demand that $s(P_2, P_1)$ be 'stationary' for small displacements of the curve connecting P_1 and P_2, these displacements vanishing at P_1 and P_2.

This is a standard problem in the calculus of variations, and its solution leads to the familiar Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \left(\frac{\partial L}{\partial \dot{x}^i} \right) - \frac{\partial L}{\partial x^i} = 0, \tag{2.59}$$

where $\dot{x}^i \equiv \mathrm{d}x^i/\mathrm{d}\lambda$ and $L \equiv [g_{ik}(\mathrm{d}x^i/\mathrm{d}\lambda)(\mathrm{d}x^k/\mathrm{d}\lambda)]^{1/2}$ is a function of x^i and $\dot{x}^i$. It is easy to see that (2.59) leads to

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \left(g_{ik} \frac{1}{L} \frac{\mathrm{d}x^k}{\mathrm{d}\lambda} \right) - \tfrac{1}{2} g_{mn,i} \frac{1}{L} \frac{\mathrm{d}x^m}{\mathrm{d}\lambda} \frac{\mathrm{d}x^n}{\mathrm{d}\lambda} = 0.$$

If we substitute

$$\mathrm{d}s = L \, \mathrm{d}\lambda \tag{2.60}$$

and use (2.39), we get the above equation in the form

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}s^2} + \Gamma^i_{kl} \frac{\mathrm{d}x^k}{\mathrm{d}s} \frac{\mathrm{d}x^l}{\mathrm{d}s} = 0. \tag{2.61}$$

There are a few loose ends to be sorted out in the above derivation. First, L would be real only for timelike curves. Thus if we want to use a real parameter along the curve, then for spacelike curves we must replace $\mathrm{d}s$ by

$$\mathrm{d}\sigma = i \, \mathrm{d}s, \qquad i = (-1)^{1/2}. \tag{2.62}$$

For null curves, $L = 0$. The above treatment therefore breaks down. It is then more convenient to replace the integral (2.58) by another:

$$I = \int_{\lambda_1}^{\lambda_2} L^2 \, \mathrm{d}\lambda, \tag{2.63}$$

and consider $\delta I = 0$. We can always choose a new parameter $\lambda' = \lambda'(\lambda)$ such that the equation of the curve takes the same form as (2.61), with λ' replacing s.

It is easy to see that (2.61) is the same as (2.57). Although s in (2.61) has the special meaning 'length along the curve', while λ in (2.57) appears to be general, it is not difficult to see that if (2.57) is satisfied then λ must be a constant multiple of s. This is because (2.57) has the first integral

$$g_{ik} \frac{\mathrm{d}x^i}{\mathrm{d}\lambda} \frac{\mathrm{d}x^k}{\mathrm{d}\lambda} = C, \qquad C = \text{constant.} \tag{2.64}$$

These curves of 'stationary distance' are called *geodesics*. For timelike curves $C > 0$, for spacelike curves $C < 0$, while for null curves $C = 0$. λ is called an *affine parameter*.

Example Let us calculate the null geodesics from $t = 0$, $r = 0$ to the point $t = T$, $r = R$, $\theta = \theta_1$, $\phi = \phi_1$ in the de Sitter spacetime

$$\mathrm{d}s^2 = c^2\,\mathrm{d}t^2 - \mathrm{e}^{2Ht}[\mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)],$$

where H = constant. It is not difficult to verify that the θ and ϕ equations of (2.61) are satisfied by $\theta = \theta_1$, $\phi = \phi_1$. That is, our straight line moves in the fixed (θ, ϕ) direction. The t equation simplifies to

$$\frac{\mathrm{d}^2 t}{\mathrm{d}\lambda^2} + \frac{H}{c^2}\mathrm{e}^{2Ht}\left(\frac{\mathrm{d}r}{\mathrm{d}\lambda}\right)^2 = 0.$$

The first integral (2.64) gives, on the other hand,

$$c^2\left(\frac{\mathrm{d}t}{\mathrm{d}\lambda}\right)^2 = \mathrm{e}^{2Ht}\left(\frac{\mathrm{d}r}{\mathrm{d}\lambda}\right)^2.$$

The two equations can be easily solved to give

$$t = \frac{1}{H}\ln\left(1 + \frac{\lambda}{\lambda_0}\right), \qquad r = \frac{c}{H}\frac{\lambda}{\lambda + \lambda_0},$$

where λ_0 is determined from the boundary condition that when $r = R$, $t = T$. Note that a solution is possible only if R and T are related by the condition

$$R = \frac{c}{H}(1 - \mathrm{e}^{-HT}).$$

We next consider the special role played by geodesics in general relativity.

2.7 The principle of equivalence

Having described the machinery of vectors and tensors, and having outlined the salient features of Riemannian geometry, we now make our first contact with physics and introduce the so-called *principle of equivalence*, which has played the key role in general relativity.

Let us go back to the purely mathematical result embodied in the relations shown in (2.43) and attempt to describe their physical meaning. These relations tell us that special (locally inertial) coordinates exist in the neighbourhood of any point P in spacetime that behave like the coordinates (t, x, y, z) of special relativity. Physically, these coordinates imply a special frame of reference in which a momentary illusion is created at P and in a small neighbourhood of P that the geometry is of special relativity. The illusion is momentary and local to P because we have seen that the relations of (2.43) cannot be made to hold everywhere and at all times.

In view of the assertion made in section 2.1 that gravitation manifests itself as non-Euclidean geometry, we would have to argue that in the above locally inertial frame gravitation has been transformed away momentarily and in a small neighbourhood of P. How does this happen in practice? Consider Einstein's celebrated example of the freely falling lift. A person inside such a lift feels weightless. The accelerated frame of reference of the lift provides the locally inertial frame in the small neighbourhood of the falling person. Similarly, a spacecraft circling around the Earth is in fact freely falling in the Earth's gravity, and the astronauts inside it feel weightless.

It should be emphasized that this feeling of weightlessness in a falling lift or a spacecraft is limited to local regions: there is no universal frame that transforms away Earth's gravity everywhere, at all times. If we demand that the relations of (2.43) hold at all points of spacetime, we would need to have $\partial \Gamma^i_{kl}/\partial x^m = 0$ everywhere, leading to $R^i_{klm} = 0$ – that is, to a flat spacetime. Thus a curved spacetime with a non-vanishing Riemann tensor is necessary to describe the genuine effects of gravitation. (See Exercise 27.)

The *weak principle of equivalence* states that effects of gravitation can be transformed away locally and over small intervals of time by using suitably accelerated frames of reference. Thus it is the physical statement of the mathematical relations given by (2.43). It is possible, however, to go from here to a much stronger statement, the so-called *strong principle of equivalence*, which states that any physical interaction (other than gravitation, which has now been identified with geometry) behaves in a locally inertial frame as if gravitation were absent. For example, Maxwell's equations will have their familiar form (of special relativity) in a locally inertial frame. Thus an observer performing a local experiment in a freely falling lift would measure the speed of light to be c.

The strong principle of equivalence enables us to extend any physical law that is expressed in the covariant language of special relativity to the more general form it would have in the presence of gravitation. The law is usually expressed in vectors, tensors, or spinors in the Minkowski spacetime of special relativity. All we have to do is to write it in terms of the corresponding entities in curved spacetime. Thus in the flat spacetime of special relativity, the Maxwell electromagnetic field F^{ik} is related to the current vector j^k by

$$F^{ik}_{\ ,i} = 4\pi j^k. \tag{2.65}$$

In curved spacetime the ordinary tensor derivative is replaced by the covariant derivative:

$$F^{ik}_{\ ;i} = 4\pi j^k. \tag{2.66}$$

Notice that the effect of gravitation enters through the Γ^i_{kl} terms that are present in (2.66). This generalization of (2.65) to (2.66) is called the *minimal coupling* of the field with gravitation, since it is the simplest one possible.

So in order to describe how other interactions behave in the presence of gravitation, we use the covariance under the general coordinate transformation as the criterion to be satisfied by their equations. It is immediately clear from the example of the electromagnetic field that a light ray describes a null geodesic.

In the same vein we can now describe a moving object that is acted on by no other interaction except gravitation – for example, a probe moving in the gravitational field of the Earth. *In the absence of gravity*, this object would move in a straight line with uniform velocity; that is, with the equation of motion,

$$\frac{\mathrm{d}u^i}{\mathrm{d}s} = 0, \qquad u^i = \text{4-velocity}. \tag{2.67}$$

In the presence of gravity, (2.67) is modified to our geodesic equation (2.61).

We end this section with another example that provides a clue about how gravitational effects show up in spacetime geometry according to general relativity. Consider the Minkowski spacetime with the standard line element

$$\mathrm{d}s^2 = c^2\,\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2. \tag{2.68}$$

If we make the coordinate transformation for a constant g,

$$x = \frac{c^2}{g}\left(\cosh\frac{gt'}{c} - 1\right) + x'\cosh\frac{gt'}{c}, \tag{2.69}$$
$$y = y', \qquad z = z', \qquad t = \frac{c}{g}\sinh\frac{gt'}{c} + \frac{x'}{c},$$

this leads to the line element

$$ds^2 = \left(1 + \frac{gx'}{c^2}\right)^2 dt'^2 - dx'^2 - dy'^2 - dz'^2. \tag{2.70}$$

What interpretation can we give to (2.70)? The origin of the (x', y', z') system has a world line whose parametric form in the old coordinates is given by

$$x = \frac{c^2}{g}\left(\cosh\frac{gt'}{c} - 1\right), \qquad y = 0, \qquad z = 0, \qquad t = \frac{c}{g}\sinh\frac{gt'}{c}. \tag{2.71}$$

Using the kinematics of special relativity, we can easily see that (2.71) describes the motion of a point that has a uniform acceleration g in the x-direction, a point that is momentarily at rest at the origin of (x, y, z) at $t = 0$. We may interpret the line element (2.70) and the new coordinate system as describing the spacetime in the rest frame of the uniformly accelerated observer.

Direct calculation shows that not all Γ^i_{kl} are zero in (2.70) at $x' = 0$, $y' = 0$, $z' = 0$. The frame is therefore non-inertial. For the neighbourhood of the origin, the metric component

$$g_{00} \approx 1 + \frac{2gx'}{c^2} = 1 + \frac{2\phi}{c^2}, \tag{2.72}$$

where ϕ is the Newtonian gravitational potential for a uniform gravitational field that induces an acceleration due to gravity $= -g$. We have here the reverse situation to that of the falling lift: we seem to have generated a pseudogravitational field by choosing a suitably accelerated observer. The prefix 'pseudo-' is used because the gravitational field is not real – it is an illusory effect arising from the choice of coordinates. The Riemann tensor is zero. Nevertheless, the relation (2.72) is also suggestive of the real gravitational field, as we will see in section 2.9.

2.8 Action principle and the energy tensors

Before examining relativity proper, let us see how we can write the laws of physics in the covariant language in Riemannian spacetime using the strong principle of equivalence. We take the familiar example of charged particles interacting with the electromagnetic field. The physical laws can be derived from an action principle. First we write the action in Minkowski spacetime:

$$\mathcal{A} = -\sum_a cm_a \int \mathrm{d}s_a - \frac{1}{16\pi c}\int F_{ik}F^{ik}\,\mathrm{d}^4x - \sum_a \frac{e_a}{c}\int A_i\,\mathrm{d}a^i; \tag{2.73}$$

here A_i are the components of the 4-potential, which are related to the field tensor F_{ik} by

$$A_{k,i} - A_{i,k} = F_{ik}, \tag{2.74}$$

while e_a and m_a are the charge and rest mass of particle a, whose coordinates are given by a^i and the proper time by s_a with

$$\mathrm{d}s_a^2 = \eta_{ik}\,\mathrm{d}a^i\,\mathrm{d}a^k. \tag{2.75}$$

How do we generalize (2.73) to Riemannian spacetime? First, we note that η_{ik} in (2.75) are replaced by g_{ik}. Next, starting from the covariant vector A_i, we generate F_{ik} by the covariant generalization of (2.74):

$$A_{k;i} - A_{i;k} = F_{ik}. \tag{2.76}$$

However, since the expression (2.76) is antisymmetric in (i, k), the extra terms involving the Christoffel symbols vanish and we are back to (2.74)! The volume integral in (2.73) is modified to

$$\int F_{ik}F^{ik}(-g)^{1/2}\,\mathrm{d}^4x. \tag{2.77}$$

The extra factor $(-g)^{1/2}$ has crept in because the combination

$$(-g)^{1/2}\,\mathrm{d}x^1\,\mathrm{d}x^2\,\mathrm{d}x^3\,\mathrm{d}x^0 = \tfrac{1}{24}e_{ijkl}\,\mathrm{d}x^i\,\mathrm{d}x^j\,\mathrm{d}x^k\,\mathrm{d}x^l$$

acts as a scalar. We therefore have the generalized form of (2.73):

$$\mathcal{A} = -\sum_a cm_a \int \mathrm{d}s_a - \frac{1}{16\pi c}\int F_{ik}F^{ik}\,(-g)^{1/2}\,\mathrm{d}^4x - \sum \frac{e_a}{c}\int A_i\,\mathrm{d}a^i. \tag{2.78}$$

The variation of the world line of particle a gives its equation of motion,

$$\frac{\mathrm{d}^2a^i}{\mathrm{d}s_a^2} + \Gamma^i_{kl}\frac{\mathrm{d}a^k}{\mathrm{d}s_a}\frac{\mathrm{d}a^l}{\mathrm{d}s_a} = \frac{e_a}{m_a}F^i_l\frac{\mathrm{d}a^l}{\mathrm{d}s_a}, \tag{2.79}$$

while the variation of A_i gives the field equations (2.66).

The transition from (2.73) to (2.78) has, however, introduced an additional independent feature into the action, besides the particle world lines and the potential vector. The new feature is the spacetime geometry typified by the metric tensor g_{ik}. What will happen if we demand that the

g_{ik} are also dynamical variables and that the action $\mathscr{A}$ remains stationary for small variations of the type

$$g_{ik} \to g_{ik} + \delta g_{ik}? \tag{2.80}$$

From the generalized action principle, should we expect to get the equations that determine the spacetime geometry? Let us investigate.

A glance at the action (2.78) shows that the last term does not contribute anything under (2.80) if we keep the worldlines and A_i fixed in spacetime. The first two terms, however, do make contributions. Let us consider them in that order. First note that

$$\delta(\mathrm{d}s_a^2) = \delta g_{ik}\, \mathrm{d}a^i\, \mathrm{d}a^k.$$

That is,

$$\delta(\mathrm{d}s_a) = \tfrac{1}{2}\delta g_{ik} \frac{\mathrm{d}a^i}{\mathrm{d}s_a} \frac{\mathrm{d}a^k}{\mathrm{d}s_a}\, \mathrm{d}s_a.$$

Therefore,

$$\delta \sum_a c m_a \int \mathrm{d}s_a = \tfrac{1}{2} \sum_a c \int m_a \frac{\mathrm{d}a^i}{\mathrm{d}s_a} \frac{\mathrm{d}a^k}{\mathrm{d}s_a}\, \mathrm{d}s_a\, \delta g_{ik}. \tag{2.81}$$

Let us consider this variation in a small 4-volume $\mathscr{V}$ near a point P. If we consider a locally inertial coordinate system near P we can identify the above expression in a more familiar form. Let us first identify

$$p^i_{(a)} = c m_a \frac{\mathrm{d}a^i}{\mathrm{d}s_a}$$

as the 4-momentum of particle a. Then $cp^0_{(a)} = E_a =$ energy of the particle, and we get

$$\tfrac{1}{2} c m_a \frac{\mathrm{d}a^i}{\mathrm{d}s_a} \frac{\mathrm{d}a^k}{\mathrm{d}s_a}\, \mathrm{d}s_a = \frac{c^2}{2E_a} p^i_{(a)} p^k_{(a)}\, \mathrm{d}t_a = \frac{c}{2E_a} p^i_{(a)} p^k_{(a)}\, \mathrm{d}x^0_a.$$

Figure 2.5 shows the volume $\mathscr{V}$ as a shaded region in the neighbourhood of P, t being the local time coordinate and x^μ $(\mu = 1, 2, 3)$ the local rectangular coordinates. The expression (2.81) can then be looked upon as a volume integral over $\mathscr{V}$ of the form

$$\delta \sum_a c m_a \int \mathrm{d}s_a = \frac{1}{2c} \int_{\mathscr{V}} \delta g_{ik} \underset{(m)}{T^{ik}}\, \mathrm{d}^4 x, \tag{2.82}$$

where $\underset{(m)}{T^{ik}}$ is the sum of expressions

$$\frac{c^2}{E_a} p^i_{(a)} p^k_{(a)}$$

for each particle a that crosses a unit volume of the shaded region near P.

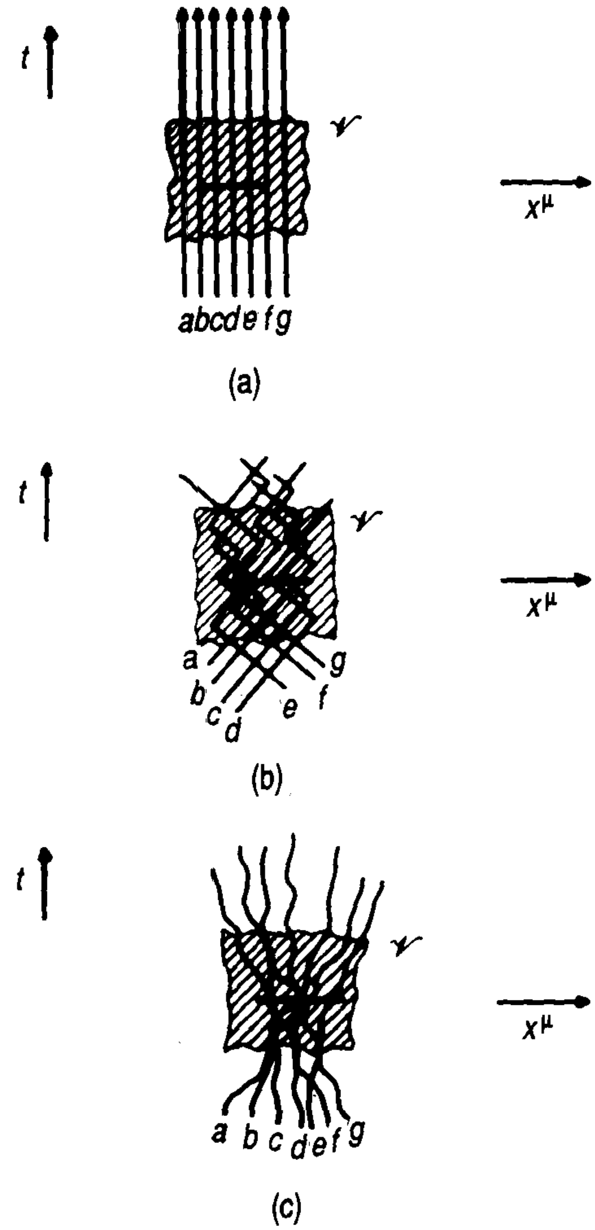

Fig. 2.5 Three cases of particle motion in the locally inertial region T near a typical point P of spacetime. The thick line on the x^μ-axes in each case represents a unit 3-volume. All particles $a, b, c, d, \ldots$ crossing this volume are counted for computing T^{ik}. (a) Particle world lines $a, b, c, \ldots$ are nearly parallel. This is the dust approximation. (b) The particles move at random with speeds near the speed of light, frequently changing directions through collisions. This is the relativistic case. (c) The intermediate situation, in which the particles collide, change directions, and generate pressures, but their motions are nonrelativistic. This is the case of a fluid.

2.8.1 Energy tensor of matter

This expression for T_{ik} is none other than the usual expression for the energy tensor of matter. Since we will need this tensor frequently, it is derived below for three different types of matter.

Dust: This is the simplest situation, in which all the world lines going through the shaded region in Figure 2.5(a) are more or less parallel, indicating that the particles of matter are moving without any relative motion in the neighbourhood of P. If we write the typical 4-velocity as u^i and using a Lorentz transformation to make $u^i = (1, 0, 0, 0)$ (that is, transforming to the rest frame of the dust) then the only non-zero component of the energy tensor is

$$T^{00} = \sum_a m_a c^2 = \rho_0 c^2,$$

where the summation is over a unit volume in the neighbourhood of P. Here ρ_0 is the rest mass density of dust. In any other Lorentz frame we get

$$\underset{(m)}{T^{ik}} = \rho_0 c^2 u^i u^k, \tag{2.83}$$

an expression that is easily generalized to any (non-Lorentzian) coordinate system.

Relativistic particles: This situation represents the opposite extreme. Here we have highly relativistic particles moving at random through $\mathcal{V}$ (see Figure 2.5(b)). The 4-momentum of a typical particle is then approximated to the form

$$p^i = \left(\frac{E}{c}, \mathbf{P}\right), \qquad E^2 = c^2 P^2 + m^2 c^4 \approx c^2 P^2, \qquad P = |\mathbf{P}|.$$

Using the fact that the particles are moving randomly, we find that the energy tensor has pressure components also:

$$\begin{aligned} T^{00} &= \sum E = \varepsilon, \\ T^{11} = T^{22} = T^{33} &= \sum \frac{P^2 c^2}{3E}. \end{aligned} \tag{2.84}$$

The factor $\frac{1}{3}$ comes from randomizing in all directions. These are the only nonzero pressure components. Here ε is the energy density. Thus for extreme relativistic particles we get

$$\underset{(m)}{T^{ik}} = \mathrm{diag}\,(\varepsilon, \tfrac{1}{3}\varepsilon, \tfrac{1}{3}\varepsilon, \tfrac{1}{3}\varepsilon). \tag{2.85}$$

This form is also applicable to randomly moving neutrinos or photons.

Fluid: This situation is illustrated in Figure 2.5(c) and consists of a collection of particles with small (nonrelativistic) random motions. If we choose the frame in which the fluid as a whole is at rest as the frame of reference, we can evaluate the components of $\underset{(m)}{T^{ik}}$ as follows. Let a typical particle have the momentum vector given by

$$p^0 = \frac{mc^2}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}},\ p^\mu = \frac{m\mathbf{v}}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}} \qquad (\mu = 1, 2, 3). \tag{2.86}$$

Then

$$T^{00} = \sum mc^2 \left(1 - \frac{v^2}{c^2}\right)^{-1/2} \approx \sum mc^2 \left(1 + \frac{v^2}{2c^2}\right) = \rho c^2,$$

$$T^{11} = T^{22} = T^{33} = \tfrac{1}{3} \sum mv^2 \left(1 - \frac{v^2}{c^2}\right)^{-1/2} \approx p. \tag{2.87}$$

Here ρ and p are the density and pressure of the fluid. In a frame of reference in which the fluid as a whole has a 4-velocity u^i, the energy tensor becomes

$$\underset{(m)}{T^{ik}} = (p + \rho c^2)u^i u^k - p\eta^{ik}. \tag{2.88}$$

The generally covariant form of (2.88) is obviously

$$\underset{(m)}{T^{ik}} = (p + \rho c^2)u^i u^k - pg^{ik}. \tag{2.89}$$

Note that ρ is not just the rest-mass density, but also includes energy density of internal motion, as seen in (2.87).

We may now relax our restriction to the locally inertial coordinate system at P. The generalized form of (2.82) is then

$$\delta \sum_a cm_a \int ds_a = \frac{1}{2c} \int \underset{(m)}{T^{ik}} (-g)^{1/2}\, \delta g_{ik}\, d^4x. \tag{2.90}$$

2.8.2 Energy tensor of the electromagnetic field

We next consider the variation of the second term of (2.73). If we keep A_i fixed, the F_{ik}, as given by (2.76) or (2.74), remain unchanged under the variation of g_{ik}. Hence

$$\delta(F_{ik}F^{ik}(-g)^{1/2}) = F_{ik}F_{lm}\delta(g^{il}g^{km}(-g)^{1/2}).$$

From (2.24) we get

$$\delta g^{ik} g_{kl} = -g^{ik}\delta g_{kl},$$

that is,

$$\delta g^{ik} = -g^{lm}g^{kn}\delta g_{mn}. \tag{2.91}$$

Also, from (2.40) we have

$$\delta\,(-g)^{1/2} = \tfrac{1}{2} g^{ik}(-g)^{1/2}\,\delta g_{ik}. \tag{2.92}$$

Substituting these expressions into the variation of the second term of the action gives

$$\delta \frac{1}{16\pi c}\int_{\mathcal{V}} F_{ik}F^{ik}(-g)^{1/2}\,\mathrm{d}^4x = \frac{1}{2c}\int_{\mathcal{V}} \underset{(em)}{T^{ik}}\,(-g)^{1/2}\,\delta g_{ik}\,\mathrm{d}^4x, \tag{2.93}$$

with the electromagnetic energy tensor given by

$$\underset{(m)}{T^{ik}} = \frac{1}{4\pi}\left(\tfrac{1}{4}F_{mn}F^{mn}g^{ik} - F^{i}{}_{l}F^{lk}\right) \tag{2.94}$$

It is obvious from our two examples that the energy tensor of any term in the action of the form Λ is related to the variation of Λ by

$$\delta\Lambda = \frac{1}{2c}\int \underset{(\Lambda)}{T^{ik}}(-g)^{1/2}\,\delta g_{ik}\,\mathrm{d}^4x. \tag{2.95}$$

In theories defined only in Minkowski space the appearance of energy tensors is somewhat *ad hoc*. They do not enter explicitly into any dynamic or field equations. They appear only through their divergences, the typical conservation of energy and momentum being given by

$$T^{ik}{}_{,k} = 0. \tag{2.96}$$

In our curved spacetime framework the T^{ik} find a natural expression through the variation of g_{ik}. It was this variation of the metric tensor that led Hilbert to derive the field equations of general relativity shortly after Einstein had proposed them from heuristic considerations. We now turn our attention to this topic.

2.9 Gravitational equations

The preceding section showed that the variation of the action $\mathcal{A}$ with respect to g_{ik} leads us to the energy tensor of various interactions. We still do not have dynamic equations that tell us how to determine the g_{ik} in terms of the distribution of matter and energy. It was Einstein's conjecture that the energy tensors should act as the 'sources' of gravity. Following the general trend of nineteenth-century physics, especially the Maxwell equations, Einstein looked for an expression that would act like a wave equation for g_{ik}, with T_{ik} as the source. It is immediately clear that the standard wave equation in the covariant form

$$g^{mn}g_{ik;mn} = \kappa T_{ik}, \tag{2.97}$$

where κ is a constant, will not do, for the left-hand side vanishes identically. Is there a second-rank tensor symmetric in its indices (like the T_{ik}) that involves second derivatives of g_{ik}? Clearly, if the tensor is to bring out the special feature of curvature of spacetime, it must be related to the Riemann tensor. After trial and error, Einstein finally arrived at the tensor G_{ik} of (2.53). His field equations of general relativity, published in 1915, took the form

$$R_{ik} - \tfrac{1}{2}g_{ik}R \equiv G_{ik} = -\kappa T_{ik}. \tag{2.98}$$

These equations have the added advantage that in view of the Bianchi identities in (2.55) we must have

$$T^{ik}{}_{;k} \equiv 0. \tag{2.99}$$

That is, the law of conservation of energy and momentum follows naturally from (2.98).

Although there are 10 Einstein equations for 10 unknown g_{ik}, the divergence condition of (2.99) reduces the number of independent equations to 6. This underdeterminacy of the problem is due to the general covariance of the theory: if g_{ik} is a solution, then so is any tensor transform of g_{ik} obtained through a change of coordinates.

The expression (2.99) follows for any T^{ik} obtained from an action principle by the variation of g_{ik} (see Exercise 33). It is therefore pertinent to ask whether the Einstein tensor can also be derived from an action principle. This problem was solved by Hilbert soon after Einstein proposed his equations of gravitation. Hilbert's problem can be posed as follows. Consider the variation of the term

$$\int_{\mathcal{V}} R\,(-g)^{1/2}\,\mathrm{d}^4x$$

for $g^{ik} \to g^{ik} + \delta g^{ik}$ with the restriction that δg^{ik} and $\delta g^{ik}{}_{,l}$ vanish on the boundary of $\mathcal{V}$. It can be shown (see Exercise 34 and 35) that

$$\begin{aligned}\delta\int_{\mathcal{V}} R(-g)^{1/2}\,\mathrm{d}^4x &= \int_{\mathcal{V}} \delta g^{ik}(R_{ik} - \tfrac{1}{2}g_{ik}R)(-g)^{1/2}\,\mathrm{d}^4x \\ &= -\int_{\mathcal{V}} \delta g_{ik}(R^{ik} - \tfrac{1}{2}g^{ik}R)(-g)^{1/2}\,\mathrm{d}^4x. \end{aligned} \tag{2.100}$$

Thus it follows that Einstein's equations can be derived from an action principle if we add to $\mathcal{A}$ the term

$$\frac{1}{2\kappa c}\int_{\mathcal{V}} R(-g)^{1/2}\,\mathrm{d}^4x. \tag{2.101}$$

If to the scalar R we add a constant (2λ, say) that is trivially a scalar, we get a modified set of field equations:

$$R_{ik} - \tfrac{1}{2}g_{ik}R + \lambda g_{ik} = -\kappa T_{ik}. \tag{2.102}$$