Chapter 1
RELATIVISTIC GRAVITY

1.1 What is a black hole?

Black holes arise because gravity affects the way light waves travel through space. Newtonian dynamics does not treat the effect of gravity on light, but we can use our Newtonian intuition to guess what sort of interaction there might be, given that there is one. The propagation of light is governed by its speed c, so it is natural to look for a characteristic speed in Newtonian gravity. The obvious one is the escape velocity $v_{\rm esc} = (2GM/R)^{1/2}$ from a body of mass M and radius R. So we expect that light emitted from the body will fail to escape to a distant observer when $v_{\rm esc} > c$, that is when the mass of a body of density ρ exceeds the value

$$M \sim (c^2/G)^{3/2}\rho^{-1/2},$$

where we have used the approximation $\rho \sim M/R^3$.

Before the discovery in the twentieth century of white dwarf stars, and later neutron stars, the densest matter known was something like lead. So the mass M_* of the smallest object having a normal density, of around say $\rho_* \sim 5000$ kg m^{-3}, that could prevent its light from escaping to a distant observer would appear to be

$$M_* \sim 10^8(\rho_*/\rho)^{1/2}M_\odot, \tag{1.1}$$

where $M_\odot$ is the solar mass. The mass M_* is very much greater than that of any known star. Such an object would be invisible or, as we would now say, would be a black hole.

These Newtonian arguments were known to Mitchell (the successor to Cavendish at Cambridge, and the first person to carry out Cavendish's experiment to measure G) and to Laplace, who was the first to carry out detailed studies of the many-body problem in the solar system.

Two things have changed this picture. The first was the discovery of a relativistic theory of gravity, Einstein's general theory of relativity, in which the behaviour of light under the influence of gravity is treated unambiguously. The second was the discovery of matter having densities in the range 10^9 to 10^{17} kg m^{-3} in the form of white dwarfs and neutron stars. The existence of bodies of such high densities suggests that stellar mass black holes might exist. We can add to these another class of astronomical object, the active galactic nuclei (or AGNs), in which the central objects have normal densities but masses of the order of $10^8\ M_\odot$. This combination of mass and density make AGNs candidates for black holes.

In a relativistic theory of gravity, by definition, the local speed of light is always c, the speed of light in a vacuum in the absence of gravity. The Newtonian picture of the emitted light being slowed down and turned back by gravity is therefore not appropriate. However, we can get a truer picture from an application of the equivalence principle. The presence of a gravitational field introduces a relative acceleration between freely-falling frames of reference. The equivalence principle then leads to an approximate relation between time interval $d\tau$ measured at radius R from a body of mass M and the corresponding time interval $d\tau'$ measured at infinity (Will, 1993)

$$d\tau \approx \left(1 - \frac{2GM}{Rc^2}\right)^{1/2} d\tau'. \tag{1.2}$$

(The exact relation depends on the full theory of gravity.) This says that a clock at radius r runs slow compared to a clock at infinity. There is a corresponding redshift z of light emitted at frequency ω and received at ω' given by $\omega/\omega' = 1 + z = d\tau'/d\tau$. This suggests that $z \to \infty$ as the Newtonian potential $2GM/R \to c^2$ and hence that the light from the surface of a body at this potential would be redshifted to invisibility. Thus the body would be a black hole. The relativistic condition $2GM/R = c^2$ is, of course, analytically the same as the Newtonian condition $v_{\text{esc}} = c$. But the condition for the Newtonian approximation to be valid is that $GM/R \ll c^2$, so we require the full theory of gravity to investigate this behaviour consistently and to treat black holes correctly. Furthermore, the gravitational potential on the surface of a neutron star is about $0.1c^2$ so we need a general relativistic theory of stellar evolution to be confident of understanding evolution beyond this stage.

The replacement of Newton's theory of gravity by Einstein's general theory of relativity does not alter the relationship between mass and density in equation (1.1), except that now the density is to be interpreted as the average density within the boundary of the (non-rotating) black hole. But it does alter our picture of the spacetime of a black hole and how it gravitates. A relativistic black hole has no material surface; all of its matter has collapsed into a singularity that is surrounded by a spherical boundary called its event horizon. The event horizon is a one-way surface: particles and light rays can enter the black hole from outside but nothing can escape from within the horizon of the hole into the external universe. An outgoing photon that originates outside the event horizon can propagate to infinity but in so doing it suffers a gravitational redshift: in Newtonian language it looses energy in doing work against the gravitational potential. This redshift is larger the closer the point of emission is to the horizon. On the other hand a photon or particle emitted inside the horizon in any direction must inevitably encounter the singularity and be annihilated. (This is strictly true only in the simplest type of black hole: in more general black holes destruction is not inevitable and the fate of a particle or photon inside the hole is more complicated.) A photon emitted at the horizon towards a distant observer stays there indefinitely. For this reason one can think of the horizon as made up of outwardly directed photons.

1.2 Why study black holes?

The importance of black holes for gravitational physics is clear: their existence is a test of our understanding of strong gravitational fields, beyond the point of small corrections to Newtonian physics, and a test of our understanding of astrophysics, particularly of stellar evolution. Current theories show that black holes are an almost inevitable consequence of the way that massive stars evolve: we therefore expect to find black holes amongst the stars in the Galaxy, and it appears that we do.

There is also a surprising and quite unexpected reason why black holes turn out to be important: this is for the potential insight they offer into the connection between quantum physics and gravity. We shall see that black holes appear formally to satisfy the laws of thermodynamics, with Mc^2 in the role of internal energy, the acceleration due to gravity in the guise of temperature and the black hole area as entropy. But this turns out to be more than a formal analogy. When we include the effects of quantum physics we find that black holes behave as real objects with a non-zero temperature and entropy: in particular they radiate like black bodies. The analogy with thermodynamics is therefore not just a formal one, but black holes really do obey the laws of thermodynamics.

We can now turn this argument around. Since black hole radiation involves a mixture of gravity and quantum physics this connection necessarily leads us into the territory of quantum gravity, and, since quantum gravity is the missing link in a complete picture of the fundamental forces, to 'theories-of-everything'. Any theory-of-everything has to be consistent with thermodynamics, and hence with black hole thermodynamics. Therefore any theory-of-everything should be able to predict the thermodynamic properties of black holes from *ab initio* statistical calculations. It is therefore interesting that theories that treat 'strings' as fundamental entities have been partially successful in this regard. It appears that black holes will play a central role in our understanding of fundamental physics.

1.3 Elements of general relativity

It is assumed that the reader has had a first acquaintance with a course on general relativity, for example from one of the many excellent introductory textbooks (for example, Kenyon, 1990, Hartle, 2003). In this section we shall present some of the main ideas of the theory, but only in the form of a brief review.

1.3.1 The principle of equivalence

The principle of equivalence tells us that local experiments (those carried out in the immediate vicinity of an event) cannot distinguish between an accelerated frame of reference and the presence of a gravitational field. In both cases we observe that bodies subject to no non-gravitational forces fall with equal acceleration. (We use the double negative in 'no non-gravitational forces' to emphasise that gravity may or may not be present, but no other forces are acting.) This creates difficulties for the Newtonian approach to dynamics because that requires us to choose a non-accelerated (or 'inertial') frame of reference. In Newtonian physics we get round this problem by designating the distant stars as a non-accelerated reference frame. This is a non-local, non-causal solution and therefore unsatisfactory. It is obviously non-local, and it is non-causal because there is no mechanism by which this reference frame is singled out, except by the fact that it gives the right answers (for example, for the motion of the planets in the Solar System).

Einstein was struck by the observation that all bodies fall with the same acceleration in a gravitational field. Newtonian gravity offers no explanation for this *universality of free-fall*. So Einstein used the universality of free fall to enunciate the principle of equivalence and made this the basis of his general relativistic theory of gravity. The principle of equivalence implies that in free fall we cannot detect the presence of a gravitational field by local experiments. (In free fall all bodies move inertially whether or not gravity is present.) Therefore in a local freely falling frame of reference we already know the laws of physics in the presence of gravity: they are the same as if gravity were absent!

1.3.2 The Newtonian affine connection

From a practical point of view it is not very easy to use the principle of equivalence directly for calculations. This is because in the presence of gravity the local freely falling frame of reference is changing from event to event and we are not experienced at doing calculations in ever changing frames of reference. Rather we need to translate this point of view into a fixed, but arbitrary, reference frame. Although reference frames and coordinate systems are not the same thing, (because the axes of a reference frame are not required to be tangents to coordinate lines), for the present purposes we shall ignore the distinction between them.

Imagine therefore that the system of coordinates $(\xi^0, \xi^1, \xi^2, \xi^3) = (\xi^\mu)$, $(\mu = 0, 1, 2, 3)$ corresponds momentarily to the natural choice of the freely-falling observer, with ξ^0 the Newtonian time (up to a factor of c). Suppose further that another set of coordinates (x^μ) are defined in a global patch (although not necessarily the whole) of spacetime. Each set of coordinates is given in terms of the other by $x^\mu = x^\mu(\xi^\nu)$ and $\xi^\mu = \xi^\mu(x^\nu)$, $(\mu, \nu = 0, 1, 2, 3)$. Note that on the left of these equations the variable stands for an independent coordinate and on the right for a function. The eliding of these distinct meanings by use of the same symbol is common practice and useful for keeping track of dependencies provided that care is taken.

According to Newtonian physics a test body subject to no non-gravitational forces will move along a spacetime trajectory defined by

$$\frac{d^2\xi^\mu}{d\tau^2} = 0. \tag{1.3}$$

Transforming to our (x^μ) coordinates this becomes, after some calculation of partial derivatives,

$$\frac{d^2x^\mu}{d\tau^2} + \frac{\partial x^\mu}{\partial \xi^\alpha}\frac{\partial^2\xi^\alpha}{\partial x^\nu \partial x^\rho}\frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau} = 0, \tag{1.4}$$

where we are employing the usual summation convention implying a sum over repeated indices. The Greek indices are assumed to range over the values 0,1,2,3 throughout. To derive (1.4) from Eq. (1.3) we have used

$$\frac{\partial x^\alpha}{\partial \xi^\mu}\frac{\partial \xi^\mu}{\partial x^\beta} = \delta^\alpha_\beta, \tag{1.5}$$

which follows from differentiation of $x^\alpha(\xi^\mu(x^\beta)) = x^\alpha$. Eq. (1.4) is of the form

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\rho}\frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau} = 0. \tag{1.6}$$

This equation for the motion of a test body holds whether or not gravity is present. If gravity is present the only place it can appear in this equation is through the quantity $\Gamma^\mu_{\nu\rho}$, called the affine connection. Looked at in this way, gravity therefore does not enter through an additional force term on the right hand side of (1.3). The only difference between the presence or absence of a gravitational field is that in the latter case it will be possible to recover the Eq. (1.3) everywhere in a single inertial coordinate system, not just locally in a freely-falling frame (because in the absence of gravity a local freely-falling frame is automatically a global inertial frame).

1.3.3 Newtonian gravity

To make the connection with the usual form of the equation of motion of a test body in Newtonian gravity we must be able to choose coodinates in which the affine connection takes an appropriate form. In fact, we must have, in some suitably chosen frame of reference (x^μ),

$$\Gamma^0_{ij} = 0; \quad \Gamma^i_{0j} = 0; \quad \Gamma^i_{00} = \frac{\partial \phi}{\partial x^i},$$

where $i, j = 1, 2, 3$, and ϕ is the Newtonian gravitational potential, since with these values for $\Gamma^\mu_{\nu\rho}$ we recover from (1.5) the equations of motion in a Newtonian gravitational field:

$$x^0 = ct = c\tau;$$
$$\frac{d^2x^i}{d\tau^2} = -\frac{\partial \phi}{\partial x^i}.$$

This also tells us that the affine connection is related to the distribution of matter through the extension to a general coordinate system of Poisson's equation $\nabla^2\phi = 4\pi G\rho$. Since this involves second derivatives of ϕ the relation between the affine connection and matter must involve the derivatives of the affine connection. The appropriate combinations of derivatives can be shown to be related to the curvature of (Newtonian) spacetime.

Newtonian gravitation is therefore a theory of the structure of spacetime, the relevant structures being the affine connection, the privileged time coordinate t and the Euclidean spatial metric. In Newtonian physics the (affine) geometry of spacetime is measured by the paths of particles and is unrelated to the geometry of time and space as measured by clocks and rods. Relativity is a lot simpler: there is only one geometry. The geometry of time and space, as measured by clocks and rods, itself governs the motion of particles, as we shall explain below.

1.3.4 Metrics in relativity

In a freely-falling frame special relativity is valid *locally* (whether or not gravity is present), and the spacetime interval (the proper distance or proper time) between neighbouring events, ds, is given by the familiar line element (or metric)

$$ds^2 = c^2dt^2 - (dx^2 + dy^2 + dz^2), \tag{1.7}$$

where t, x, y, z are the time and rectangular spatial coordinates of the freely falling (inertial) observer. In tensor notation this line element can be written

$$ds^2 = \eta_{\alpha\beta}d\xi^\alpha d\xi^\beta, \tag{1.8}$$

where $\eta_{\alpha\beta}$ is the metric tensor and we are using $(\xi^\alpha) = (ct,\, x,\, y,\, z)$ for this special coordinate system at a point. With these coordinates the metric has components $\eta_{\alpha\beta} = \mathrm{diag}(1, -1, -1, -1)$. In this convention ds^2 is positive for a timelike interval, zero for a lightlike interval and negative for a spacelike interval. This is convenient for dealing with the motion of particles. For positive ds^2 then $ds/c = d\tau$, where $d\tau$ is the proper time between the events. For negative ds^2 then $(-ds^2)^{1/2} = dl$ is the proper distance between the two events. An alternative convention often used for the metric is $\eta_{\alpha\beta} = \mathrm{diag}(-1, 1, 1, 1)$. Care is needed in copying formulae from references to adjust these, if necessary, to the convention being employed.

In a general coordinate system, which we shall call (x^μ), we have

$$ds^2 = \eta_{\alpha\beta}\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\partial\xi^\beta}{\partial x^\nu}dx^\mu dx^\nu = g_{\mu\nu}dx^\mu dx^\nu \tag{1.9}$$

(from the chain rule for partial derivatives, $d\xi^\alpha = (\partial\xi^\alpha/\partial x^\mu)dx^\mu$). The quantities $g_{\mu\nu}$ are called the metric coefficients. They are symmetric, that is to say $g_{\mu\nu} = g_{\nu\mu}$, and are, in general, functions of the spacetime coordinates.

In the absence of gravity we can find a *global* coordinate system (ξ^α) in which the metric takes the form (1.8) everywhere. In the presence of gravity we can find such coordinates only in an infinitesimal neighbourhood of each spacetime point.

Problem 1 *Show that neither the Minkowski metric Eq. (1.7) nor the sum of squares is an invariant under the Galilean transformation and hence that there is no interval in Newtonian spacetime (i.e. Newtonian spacetime does not admit a metric).*

The motion of particles is again governed by Eq. (1.6) in relativity, because the arguments leading to it are still valid, except that now both $t \neq \tau$ and $x^0 \neq c\tau$ (since these would not be compatible with the metric Eq. (1.8) or (1.9)).

Now we see that in local freely falling coordinates the metric is obtained from first derivatives of the ξ^μ whereas the affine connection depends on second derivatives. We therefore expect that there will exist a relation between the components of the affine connection and derivatives of components of the metric. This is indeed the case:

$$\Gamma^\mu_{\nu\rho} = \frac{1}{2} g^{\mu\alpha} \left(\partial_\nu g_{\rho\alpha} + \partial_\rho g_{\nu\alpha} - \partial_\alpha g_{\nu\rho} \right), \tag{1.10}$$

where $\partial_\alpha f = \partial f / \partial x^\alpha$ and $(g^{\mu\nu})$ is the inverse matrix to $(g_{\mu\nu})$, so

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

Thus we find the values of $\Gamma^\mu_{\nu\rho}$ from the values of the metric coefficients appropriate to a particular gravitational field.

So in relativity there is just one spacetime geometry defined by the metric and measured by both clocks and rods *and* by particle paths. The motion of particles, is governed by Eq. (1.5), called the geodesic equation (because it is of the same form as the equation for the shortest paths on a curved surface, which are called geodesics). Note that the physical interpretation of the coordinates (x^μ) (what they measure physically) depends on the metric coefficients, so cannot be determined unless the metric is known.

1.3.5 The velocity and momentum 4-vector

In this section we define two important 4-vectors relating to the motion of a particle, namely its velocity and momentum; acceleration will be dealt with later. The 4-velocity vector of a particle with position $(x^\alpha(\tau))$ is given by

$$(u^\mu) = \left(\frac{dx^\mu}{d\tau} \right) = \left(\frac{dx^0}{d\tau}, \frac{dx^1}{d\tau}, \frac{dx^2}{d\tau}, \frac{dx^3}{d\tau} \right),$$

formally as in special relativity. The 4-momentum of a particle of mass m_0 is $p^\mu = m_0 u^\mu$.

We lower indices using the metric tensor $g_{\mu\nu}$, so the covariant components of momentum are $p_\mu = g_{\mu\nu}p^\nu$. Conversely, we raise indices using $g^{\mu\nu}$, so the contravariant components are $p^\mu = g^{\mu\nu}p_\nu$.

Taking the scalar product of the 4-velocity with itself gives a relation that we shall use repeatedly below:

$$u^\mu u_\mu = c^2, \tag{1.11}$$

and, similarly,

$$p^\mu p_\mu = m_0^2 c^2. \tag{1.12}$$

Recall that a scalar produce is independent of the frame of reference in which it is evaluated. In particular, we can choose the local freely falling frame of the particle. In this frame special relativity is valid, so the metric takes the form (1.7) and the components of (u^μ) are $(c, 0, 0, 0)$. Thus in this frame we readily verify Eq. (1.11) and similarly Eq. (1.12).

1.3.6 General vectors and tensors

By definition, if we make a coordinate transformation $x \to x'(x)$, then the contravariant components of a 4-vector v^μ transform according to

$$v'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} v^\nu$$

and the covariant components as

$$v'_\mu = \frac{\partial x^\nu}{\partial x'^\mu} v_\nu,$$

with

$$\frac{\partial x'^\mu}{\partial x^\nu}\frac{\partial x^\nu}{\partial x'^\rho} = \delta^\mu_\rho.$$

Higher rank tensors (having multiple indices) transform by extension of this rule to each index. For example:

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

Problem 2 *Obtain the inverse transformations*

$$v^\mu = \frac{\partial x^\mu}{\partial x'^\nu} v'^\nu \qquad \textit{and} \qquad v_\mu = \frac{\partial x'^\nu}{\partial x^\mu} v'_\nu.$$

Problem 3 *By showing that* $v_\mu v^\mu = v'_\mu v'^\mu$ *verify the invariance of the scalar product.*

1.3.7 Locally measured physical quantities

We often need to calculate the energy or velocity of a particle as measured by a local observer, for example, an observer having a fixed location in some global coordinate system, or a locally freely falling observer. The energy of a particle with respect to a local observer with 4-velocity $u^\mu_{\rm obs}$ is the time component of the 4-momentum of the particle in the observer's frame of reference and is therefore obtained by projecting the 4-momentum on to the velocity 4-vector of the observer. So

$$\mathcal{E} = u^\mu_{\rm obs} p_\mu = m_0 \gamma c^2,$$

where $\gamma = (1 - v^2/c^2)^{-1/2}$ is the usual relativistic gamma factor and v the local velocity of the particle. As in section 1.3.5, we have evaluated the scalar product in the local freely falling frame of the observer where special relativity applies. The scalar product is an invariant, that is, its value is independent of the coordinate system used to evaluate it, so, for example, we can evaluate it in a local inertial frame and also in a global coordinate system and thus relate the global energy to the locally measured energy. We will make frequent use of this property of the scalar product in subsequent chapters.

Physical quantities in two local frames at the same event will be connected by a Lorentz transformation between them even though one or both of the frames may be accelerating. This follows because the instantaneous rates of clocks and lengths of rods are not affected by accelerations and depend only on the relative velocities.

1.3.8 Derivatives in relativity

Given a vector field $v^\mu(x)$, say, it is easy to see that the usual partial derivative $\partial v^\mu/\partial x^\nu$ is, in general, not a physically meaningful quantity: we can give it any value we like by the choice of coordinates x^μ. In particular, a vector field which is constant in one coordinate system, in which its partial derivatives will vanish, may have non-zero derivatives in another coordinate system. The problem arises because the usual definition of a partial derivative compares the values of the quantity in question at two different points and hence, in general, at points where the coordinates can be changed independently. Rates of change having a physical interpretation will be formed from differences in quantities at the same point. Various such derivatives exist, depending on how the quantities are brought to the same point.

The most common, although not the simplest, is the (so called) covariant derivative. This is a generalisation of the behaviour of a tangent vector to a curve. Let $V^\mu = dx^\mu/d\tau$ be the tangent vector to the curve $x^\mu = x^\mu(\tau)$. Eq. (1.6) tells us that

$$\delta V^\mu = -\Gamma^\mu_{\nu\rho} V^\nu \delta x^\rho$$

is the change in the tangent vector on going from x^μ to $x'^\mu = x^\mu + \delta x^\mu$ along a geodesic. We now use this to define the parallel transport of any vector, A^μ say, between two neighbouring points. To form the covariant derivative we take the difference between the vector A^μ at x' and the parallely transported $A^\mu + \delta A^\mu = A^\mu - \Gamma^\mu_{\nu\rho} A^\nu \delta x^\rho$ at x', in the usual limit:

$$\begin{aligned}\nabla_\nu A^\mu &= \lim_{\delta x \to 0} \frac{A^\mu(x+\delta x) - (A^\mu(x) + \delta A^\mu)}{\delta x^\nu} \\ &= \frac{\partial A^\mu}{\partial x^\nu} + \Gamma^\mu_{\nu\rho} A^\rho.\end{aligned}$$

The quantity $\nabla_\nu A^\mu$ is the covariant derivative of A^μ, also denoted by a semi-colon as $A^\mu_{;\nu}$, or sometimes by DA^μ/dx^ν. The covariant derivative therefore measures the derivative corresponding to parallel transport. Note that in this notation the condition for a geodesic becomes that the tangent vector should have zero covariant derivative along the tangent, so $V^\nu \nabla_\nu V^\mu = 0$. This is the total or directional derivative, also written in terms of the affine parameter τ as $DV^\mu/d\tau = 0$.

Problem 4 *The covariant derivative of a scalar is just the usual (partial) derivative (because the value of a scalar is independent of the coordinate system). Hence, for two vector fields, A^μ and B_μ, we have $\nabla_\nu(A^\mu B_\mu) = \partial_\nu(A^\mu B_\mu)$. Use this to show that*

$$\nabla_\nu B_\mu = \frac{\partial B_\mu}{\partial x^\nu} - \Gamma^\rho_{\mu\nu} B_\rho.$$

Since we have seen that in a freely falling frame of reference the components of the affine connection vanish at a point, another definition of the covariant derivative is that it is the ordinary (partial) derivative in a freely falling frame.

A second type of derivative that is sometimes useful is the Lie derivative. To define this we use for the transported vector field at $x'^\mu = x^\mu + V^\mu \delta\tau$ the vector obtained by using the standard transformation law between coordinate systems, namely,

$$\begin{aligned}A^\mu + \delta A^\mu &= \frac{\partial x'^\mu}{\partial x^\nu} A^\nu \\ &= (\delta^\mu_\nu + \frac{\partial V^\mu}{\partial x^\nu}\delta\tau) A^\nu.\end{aligned}$$

Note that the transformation is here regarded as an active one, taking the point x^μ to a different point x'^μ, rather than a passive change of coordinates at a given point. The Lie derivative is then defined as

$$\begin{aligned}\pounds_V A^\mu &= \lim_{\delta\tau \to 0} \frac{A^\mu(x+\delta x) - (A^\mu(x) + \delta A^\mu)}{\delta\tau} \\ &= V^\nu \partial_\nu A^\mu - A^\nu \partial_\nu V^\mu.\end{aligned}$$

Problem 5 *Show that the partial derivatives in the definition of the Lie derivative can be replaced by covariant derivatives, i.e. that*

$$\pounds_V A^\mu = V^\nu \nabla_\nu A^\mu - A^\nu \nabla_\nu V^\mu.$$

Roughly speaking we can say that the Lie derivative is the directional derivative of a vector field along a curve adjusted for the change in the tangent.

Problem 6 *Generalise the tensor transformation law, the covariant derivative and the Lie derivative to a tensor field $T^{\mu\nu}$.*

1.3.9 Acceleration 4-vector

As an example of the covariant derivative we can construct the acceleration 4-vector, which is obtained from the velocity 4-vector $(u^\mu) = (dx^\mu/d\tau)$ as follows:

$$\begin{aligned} a^\mu = \frac{Du^\mu}{d\tau} &= \frac{dx^\alpha}{d\tau}\frac{Du^\mu}{dx^\alpha} \\ &= \frac{dx^\alpha}{d\tau}\left(\frac{\partial u^\mu}{\partial x^\alpha} + \Gamma^\mu_{\beta\alpha}u^\beta\right) \\ &= \frac{du^\mu}{d\tau} + \Gamma^\mu_{\beta\alpha}u^\beta u^\alpha. \end{aligned}$$

The magnitude of the rest frame or proper acceleration a of a particle is given by

$$a^\mu a_\mu = -a^2. \tag{1.13}$$

We shall give an example of the use of the Lie derivative below (problem 8).

1.3.10 Paths of light

In relativity the path of a light ray is also governed by the metric through

$$0 = g_{\mu\nu}dx^\mu dx^\nu$$

and

$$\frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu_{\beta\gamma}\frac{dx^\beta}{d\lambda}\frac{dx^\gamma}{d\lambda} = 0,$$

in which λ is a parameter which varies along the world line of a light ray (and is called an affine parameter, because it maintains for light rays the usual form of the geodesic equation).

1.3.11 Einstein's field equations

The problem of relating the metric coefficients $g_{\mu\nu}(x^\alpha)$, as functions of the coordinates, to the distribution of matter is solved by Einstein's field equations , which are

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu}. \tag{1.14}$$

Here $R_{\mu\nu}$ is the Ricci curvature tensor given by

$$R_{\mu\nu} = \frac{\partial \Gamma^{\gamma}_{\mu\nu}}{\partial x^{\gamma}} - \frac{\partial \Gamma^{\gamma}_{\mu\gamma}}{\partial x^{\nu}} + \Gamma^{\gamma}_{\mu\nu}\Gamma^{\delta}_{\gamma\delta} - \Gamma^{\delta}_{\nu\gamma}\Gamma^{\gamma}_{\mu\delta} \tag{1.15}$$

and $R = g^{\alpha\beta}R_{\alpha\beta}$ is the Ricci scalar. The tensor $T_{\mu\nu}$ is the energy momentum tensor which includes all the sources of the gravitational field excluding the energy and momentum in the gravitational field itself which is accounted for by the non-linearity of the equations. Equations (1.14) provide ten non-linear partial differential equations for the metric coefficients. In chapters 2 and 3 we shall be studying certain symmetrical solutions of the Einstein field equations in a vacuum, for which $T_{\mu\nu} = 0$. In chapter 5 we shall need the full equations (1.14).

Problem 7 *Show that if $T_{\mu\nu} = 0$ then $R = 0$.*

1.3.12 Symmetry and Killing's equation

Despite the complexity of Einstein's equations a surprising number of exact solutions are known. These are obtained by imposing symmetries on the spacetime which restrict the possible form of the metric. We shall meet two cases in the following, namely spherical and axial symmetry. In order to exploit these symmetries we shall need one mathematical result which we derive here.

Let the vector k^{μ} be a direction of symmetry; for example k^{μ} might point along the azimuthal (ϕ -) direction in spherical symmetry, i.e. $(k^{\mu}) = (0,0,0,1)$ in spherical polar coordinates. Then we shall show that

$$k_{\mu;\nu} + k_{\nu;\mu} = 0, \tag{1.16}$$

where the semi-colon denotes the covariant derivative

$$k_{\mu;\nu} = \frac{\partial k_{\mu}}{\partial x^{\nu}} - \Gamma^{\lambda}_{\mu\nu}k_{\lambda}.$$

Eq. (1.16) is called Killing's equation.

Let events P at (x^{μ}) and Q at (x'^{μ}) be separated by a small distance in the direction of symmetry k^{μ}, so

$$x'^{\mu} = x^{\mu} + \varepsilon k^{\mu}. \tag{1.17}$$

In moving from P to Q we do two things: we change the label of the point from x^{μ} to $x^{\mu} + \varepsilon k^{\mu}$ and we make an active transformation of coordinates from x^{μ} to x'^{μ}. The metric coefficients at x'^{μ} are related to those at x^{μ} by

$$g_{\mu\nu}(x^{\lambda}) = \frac{\partial x'^{\alpha}}{\partial x^{\mu}}\frac{\partial x'^{\beta}}{\partial x^{\nu}}g'_{\alpha\beta}(x'^{\lambda}).$$

As we have moved in a direction of symmetry the *function* $g'_{\mu\nu}(x'^{\lambda})$ of the new coordinates has the same form as the *function* $g_{\mu\nu}(x^{\lambda})$ of the original coordinates, so we can substitute

$$g'_{\mu\nu}(x'^{\lambda}) = g_{\mu\nu}(x'^{\lambda}).$$

Expanding $g_{\alpha\beta}(x'^{\lambda})$ as a Taylor series and using Eq. (1.17) to evaluate the partial derivatives we get

$$g_{\mu\nu}(x^{\lambda}) = (\delta^{\alpha}_{\mu} + \varepsilon\partial_{\mu}k^{\alpha})(\delta^{\beta}_{\nu} + \varepsilon\partial_{\nu}k^{\beta})(g_{\alpha\beta}(x^{\lambda}) + \varepsilon k^{\lambda}\partial_{\lambda}g_{\alpha\beta}(x^{\lambda}) + ...).$$

Multiplying out the brackets to first order in ε gives us

$$g_{\mu\nu}(x^{\lambda}) = g_{\mu\nu}(x^{\lambda}) + \varepsilon k^{\lambda}\partial_{\lambda}g_{\mu\nu}(x^{\lambda}) + \varepsilon g_{\mu\beta}(x^{\lambda})\partial_{\nu}k^{\beta} + \varepsilon g_{\alpha\nu}(x^{\lambda})\partial_{\mu}k^{\alpha}.$$

But

$$\partial_{\nu}(g_{\mu\beta}k^{\beta}) = \partial_{\nu}k_{\mu} = g_{\mu\beta}\partial_{\nu}k^{\beta} + k^{\beta}\partial_{\nu}g_{\mu\beta},$$

and we get

$$\begin{aligned} 0 &= \varepsilon(\partial_{\mu}k_{\nu} - k^{\alpha}\partial_{\mu}g_{\alpha\nu}) + \varepsilon(\partial_{\nu}k_{\mu} - k^{\beta}\partial_{\nu}g_{\mu\beta}) + \varepsilon k^{\lambda}\partial_{\lambda}g_{\mu\nu} \\ &= \varepsilon(k_{\mu;\nu} + k_{\nu;\mu}), \end{aligned}$$

where the final equality is obtained using the relation Eq. (1.10). A vector field k^{μ} satisfying this relation is called a Killing vector.

Problem 8 *Show that Killing's equation (1.16) is equivalent to* $\pounds_k g_{\mu\nu} = 0$. *(You will need the result of problem 6.)*