East Tennessee State University

Department of Physics, Astronomy & Geology

Astrophysics ASTR3415

Part 2: Introduction to General Relativity

These notes provide an introduction to the theory of general relativity. First the principles of equivalence and covariance are introduced and discussed. Elementary tensor calculus on riemannian manifolds is then introduced as a means of describing curved spacetimes and arriving at Einstein's field equations.

Martin Hendry

September 2005.

Chapter 1

Introduction: Foundations of GR

General relativity (GR) explains gravitation as a consequence of the curvature of spacetime. In turn spacetime curvature is a consequence of the presence of matter. Spacetime curvature affects the movement of matter, which reciprocally determines the geometric properties and evolution of spacetime. We can sum this up neatly as follows:

> *"Spacetime tells matter how to move,*
>
> *and matter tells spacetime how to curve"*

By the completion of this section, we will see that the "geometry" or curvature properties of spacetime can be described by what is called the **Einstein tensor**, $G_{\mu\nu}$, and the matter/energy properties by the **energy-momentum tensor**, $T_{\mu\nu}$. Einstein's theory of general relativity then states that these two tensors are proportional to each other, i.e.:

$$G_{\mu\nu} = kT_{\mu\nu} \tag{1.1}$$

(where k is a constant, the numerical value of which depends on Newton's gravitational constant.)

Equation (1.1) is in fact a set of 16 equations, known as Einstein's field equations of general relativity, and basically contains everything we need to know in this course! The lectures in this section are all about understanding where equation (1.1) comes from, how it is derived, what the symbols represent and what it means physically. To do that we first need to develop specific mathematical tools (known as tensor analysis) which can describe curved spacetime. We will then derive Einstein's equations, and show that they reduce to familiar Newtonian gravity in the non-relativistic limit. Specific astronomical applications to planetary orbits, stellar structure, neutron stars and black holes and gravitational waves will be covered in Part 3.

GR is, as the name implies, a generalisation of **special relativity** (SR). In SR Einstein attempted to formulate the known laws of physics so that they would be valid in all **inertial frames**. If one accepts that the speed of light is a constant in all such frames, then one has also to accept there is no absolute time. Similarly, the distance between points in 3-D space is no longer invariant, but will be measured differently by different inertial observers. Newtonian gravitation is thus inherently non-relativistic since it describes the gravitational force between two masses as acting instantaneously, and as depending on the distance separating the two masses. Different inertial observers would not agree about either point, and so would not agree about the force of gravity between the two masses. In developing general relativity, Einstein set out to describe gravity in a manner that could be defined consistently by *any* observer, no matter how they were moving relative to any other observer. Einstein realised, however, that this seemed to imply something very profound: that gravity and acceleration are fundamentally equivalent – an idea which he enshrined in the **principles of equivalence**.

1.1 Equivalence principles

The key physical ideas which underpin GR are the weak and strong principles of equivalence.

1.1.1 The weak principle of equivalence

Since the time of Galileo empirical evidence has suggested that all bodies freely falling in the Earth's gravitational field accelerate at the same rate. While we learn about this behaviour (and perhaps conduct experiments to verify it) at high school, its importance is often not strongly emphasised in elementary discussions of gravity and Newton's laws. To understand its signficance we think firstly about what we mean by the concept of **mass**. Within the framework of Newtonian physics mass is defined in two distinct ways:

- The **inertial mass** of a body can be thought of as a measure of the body's resistance to acceleration. According the Newton's 2nd law of motion, a body with inertial mass $m_{\rm I}$ subjected to a force $\vec{F}$ will experience an acceleration, $\vec{a}_{\rm I}$ which satisfies

$$\vec{F}_{\rm I} = m_{\rm I}\,\vec{a}_{\rm I} \tag{1.2}$$

- The **gravitational mass** of a body, on the other hand, is a measure of the gravitational attraction which it exerts on every other body in the Universe. According to Newton's law of universal gravitation, a body with gravitational mass $m_{\rm G}$, a distance r from another body of gravitational mass M will experience a gravitational force due to the second body given by

$$\vec{F}_{\rm G} = G\frac{m_{\rm G}\,M}{r^2}\hat{r} \tag{1.3}$$

  where G is the gravitational constant, and $\hat{r}$ is a unit vector in the direction from $m_{\rm G}$ to M.

There is no *a priori* reason why the inertial and gravitational mass should be related. Yet if we assume that they are in fact *identical*, we immediately have an explanation for the empirical 'fact' that freely falling bodies – regardless of their mass – accelerate at the same rate. We can see this if we apply Newton's 2nd law to equation 1.3; it then follows that the **gravitational acceleration** (which is usually denoted by $\vec{g}$) of the mass m_{G} is independent of m_{G}. This assumption, of the exact equality of the inertial and gravitational mass, is known as the **weak principle of equivalence**. It also implies that "freely falling" bodies behave as if they are in what is known as an **inertial frame** – i.e. they obey Newton's 2nd law of motion.

But how does GR deal with this? After all, we have already noted that Newton's law of Universal gravitation is inherently non-relativistic. Remember GR describes gravity as a consequence of the curvature of spacetime. Expressed in the language of GR, the weak equivalence principle implies that the trajectory (or 'worldline') of a body moving in a gravitational field is a special curve known as a **geodesic** which is independent of the mass of the body and depends only of the curvature of spacetime[1].

1.1.2 The strong principle of equivalence

The strong principle goes further and states that locally (in what we term a **local inertial frame** – a localised region of spacetime in which Newton's 2nd law holds – see later) all physical phenomena are in agreement with special relativity. There are two important and

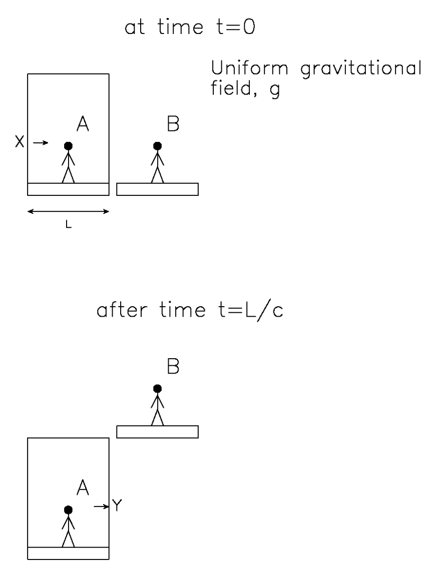
[1]Strictly speaking we should be thinking here of what physicists usually refer to as a 'test particle' – i.e. a body of sufficiently small mass that it has no effect of the curvature of the spacetime around it. If the body were massive enough then it would affect the spacetime curvature, which in turn would change its trajectory – thus implying that the trajectory *does* depend on its mass.

immediate consequences of this principle. The first is that the path of a light ray should be bent by gravitational fields, and secondly, there should be a gravitational redshift.

1.1.3 Bending of light in a gravitational field

Consider a uniform gravitational field, $\vec{g}$. A lift is in free fall in this gravitational field. Of course from the equivalence principle all bodies that do not experience electromagnetic or other non-gravitational forces should just move in straight lines. Special relativity should hold. Suppose the lift has just started to free fall (see Figure 1), and a photon is emitted horizontally from point X on the left hand side of the lift, and after some time hits the other side. There are two observers of these events. A is in the falling lift, and B stands on the platform outside the lift. Free falling A experiences no gravitational field. B on the other hand experiences the gravitational field.

Figure 1



According to A light travels in a straight horizontal line, and so must hit the far side of the lift at point Y at exactly the same height as point X, where it was emitted. The time taken for the photon to travel the width, L, of the lift must be L/c. Both observers should agree approximately on this time. According to B, who sees the lift accelerating to a speed of $g\,L/c$ at the time the photon hits the far side, the point Y will in fact have moved a vertical distance of $\frac{1}{2}gt^2 = \frac{1}{2}g\,(L/c)^2$. Since A's observation must be correct from the equivalence principle, B can only reconcile matters by accepting that the gravitational field has bent the light path.

(In GRII we shall derive this result rigorously for light deflection in the Schwarzschild metric, which corresponds to the exterior spacetime induced by a spherical mass. The deflection of light was one of the classical tests of GR.)

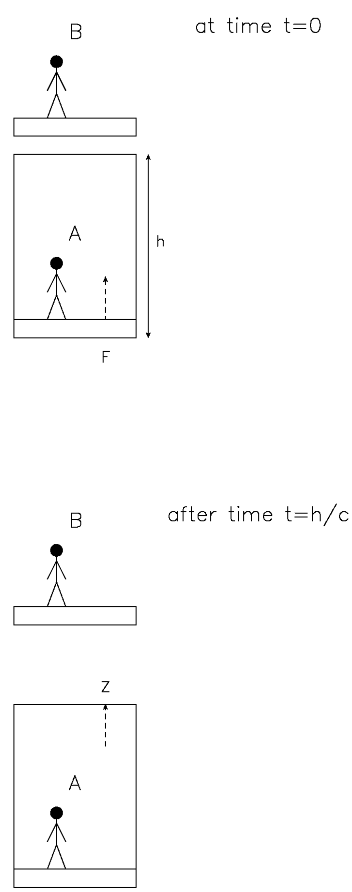
Exercise: Calculate the angular deflection of light at the surface of the Earth. Consider a horizontal path of 1 km length.

1.1.4 Gravitational redshift of spectral lines

With a similar set up as before, now consider a photon emitted upward from the floor of the lift at point F just at the time the lift is allowed to free fall, and that strikes a detector on the ceiling at point Z (see Figure 2). This time our two observers are A, inside the lift, and B, on a platform above the lift. What frequency does A observe when the photon strikes the detector? The photon must have the same frequency as when emitted, because A is locally inertial. If B now conducts the experiment and measures the frequency of the arriving photon, what does B see? According to A, for whom everything is in agreement with special relativity, observer B is receding at speed $v = g\;h/c$ when the photon reaches F. So B would observe the photon to have a redshift of $v/c = g\ h/c^2$. A of course says this is simply a Doppler shift, since B is moving away from the source. B must attribute the shift to the gravitational field – in "climbing" out of the gravitational field, the photon "loses" energy, and so is redshifted.

This redshift can also be expressed in terms of the change in gravitational potential, ϕ, since $gh = -\delta\phi$. Thus $\delta\lambda/\lambda = -\delta\phi/c^2$. Emerging from the gravitational field the photon will be observed to be redshifted. This effect has been observed in the spectral lines of white dwarf stars. On the Earth both this and the bending of light are very weak effects.

Figure 2



Exercise: Calculate the gravitational redshift of a photon moving upward through 1 km vertical height at the Earth's surface.

Exercise: What is the gravitational redshift of a spectral line emitted at the surface of a typical white dwarf star and observed at the Earth?

1.2 Spacetime curvature

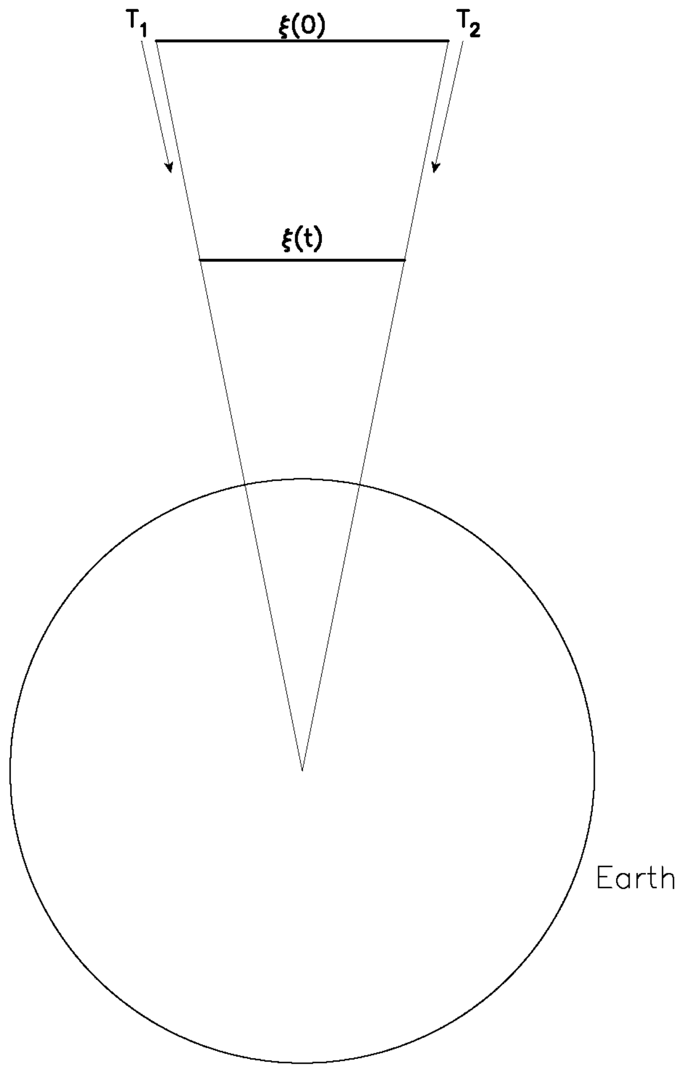
As we have already mentioned, gravitation appears in GR as spacetime curvature. Let us see how this arises.

1.2.1 Locally inertial frames (LIF)

So far we have considered only uniform gravitational fields. Generally gravitational fields are not uniform. Thus a free falling frame is only inertial over a limited spatial and temporal region around a given event. Consider for instance two free test particles that are separated by a small distance and initially at rest with respect to a LIF – see Figure 3. The fact that this is only a locally inertial frame is reflected by the fact that the distance between the two test particles will noticeably change after a certain period of time.

Exercise: Take the situation above the Earth's surface. The initial separation of the two test particles, T_1 and T_2 is ξ_0. Take $\xi_0 = 25$ m. After 7s suppose the separation of T_1 and T_2 is $\xi = \xi_0 + \Delta\xi_0$. Show that $\Delta\xi_0 \approx -10^{-3}$m (take $g = 10$ ms^{-2} and the radius, R, of the Earth to be 6×10^6 m. (Hint: $\Delta\xi/\xi = \Delta r/R$, where Δr is the change in the radial distance of the test particles from the Earth's centre.)

Figure 3



1.2.2 Geodesic deviation

The separation, ξ, between the two free test particles is called the **geodesic deviation**. In general ξ is a vector. (In fact, it is a **four vector**, if we consider time separations as well – see later). It is the acceleration of this geodesic deviation that indicates the presence of a gravitational field, or, as we shall see later, the curvature of spacetime. In the simple example illustrated in Figure 3, we can define our coordinate system so that only the x (i.e. horizontal) component of ξ is non-zero. We denote this component by ξ_x. From similar triangles we have

$$\frac{\xi_x + \Delta\xi_x}{r + \Delta r} = \frac{\xi_x}{r} = k \tag{1.4}$$

where k is a constant. Taking derivatives with respect to time gives

$$\ddot{\xi}_x = k\ddot{r} = -\frac{kGM}{r^2} \tag{1.5}$$

Substituting for $k = \xi_x/r$ yields

$$\ddot{\xi}_x = -\frac{\xi_x}{r}\,\frac{GM}{r^2} = -\frac{GM\xi_x}{r^3} \tag{1.6}$$

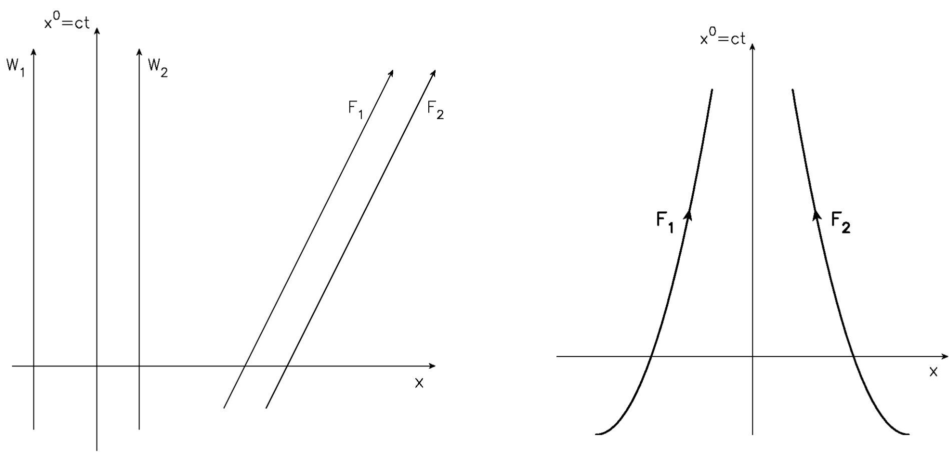
At the Earth's surface, $r \approx R$ so $\ddot{\xi}_x = -GM\xi_x/R^3$. Rewriting in more sensible units we obtain

$$\frac{d^2\xi_x}{d(ct)^2} = -\frac{GM}{R^3c^2}\;\xi_x \tag{1.7}$$

In the flat spacetime of Minkowski free test particles have worldlines that are 'straight'. Thus the acceleration of the geodesic deviation is zero for Minkowski spacetime – see Figure 4a. Worldlines W_1 and W_2 remain parallel, as do worldlines F_1 and F_2.

Figure 4

(a) **(b)**



Test particles in GR have worldlines that are geodesics, but now – because of the presence of matter – the spacetime is **not** flat. Consequently the geodesics are not 'straight lines', and there is an acceleration of the geodesic deviation – see Figure 4b. Worldlines F_1 and F_2 do not remain parallel. In the example illustrated in Figure 3, the test particles are initially at rest with respect to the LIF, but almost imperceptibly they move towards each other as they fall towards Earth. Notice that in our equation

$$\frac{d^2\xi_x}{d(ct)^2} = -\frac{GM}{R^3c^2}\ \xi_x \tag{1.8}$$

the factor

$$\frac{GM}{R^3c^2}$$

has the dimensions m^{-2}. Evaluated at the Earth's surface this quantity has the value $10^{-23}\ \mathrm{m}^{-2}$.

Exercise: Evaluate the factor

$$\frac{GM}{R^3c^2}$$

at the surface of the Earth, the Sun, and a neutron star of one solar mass.

We can understand equation (1.8) in terms of a 2-D analogy. Suppose T_1 and T_2 are on the equator of a sphere of radius a (see Figure 5). Consider geodesics perpendicular to the equator passing through T_1 and T_2 . The arc distance along the geodesics is denoted by s and the separation of the geodesics at s is $\xi(s)$. Evidently this geodesic separation is not constant as we change s. Let us write down the differential equation governing the acceleration of this geodesic separation or deviation. If $\xi(0)$ is the initial deviation, we may write $d\phi = \xi(0)/a$ and so

$$\xi(s) = a\cos\theta\, d\phi = \xi(0)\cos\theta = \xi(0)\cos s/a \tag{1.9}$$

Differentiating $\xi(s)$ twice with respect to s yields

$$\frac{d^2\xi}{ds^2} = -\frac{1}{a^2}\xi \tag{1.10}$$

Compare this with equation (1.8). In some sense the quantity

$$\mathcal{R} = \left\{\frac{GM}{R^3c^2}\right\}^{-\frac{1}{2}}$$

represents the radius of curvature of spacetime at the surface of the Earth.

Exercise: Sketch on Figure 6 the worldline of the Earth, taking the sun as the origin of the coordinate system.

Figure 5

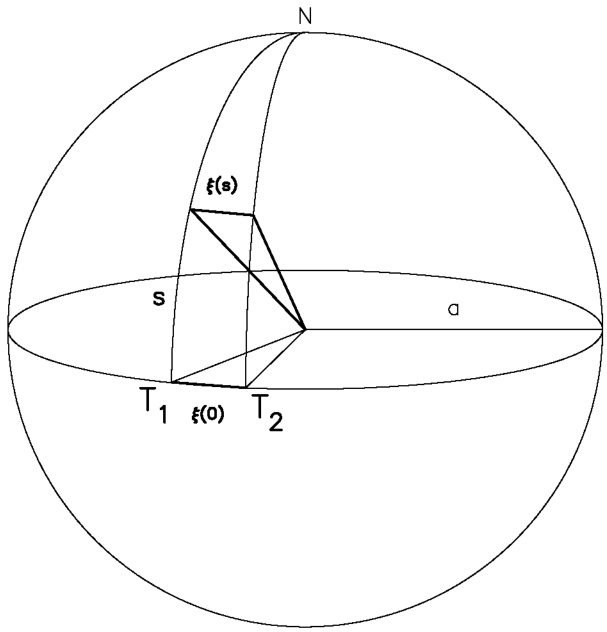
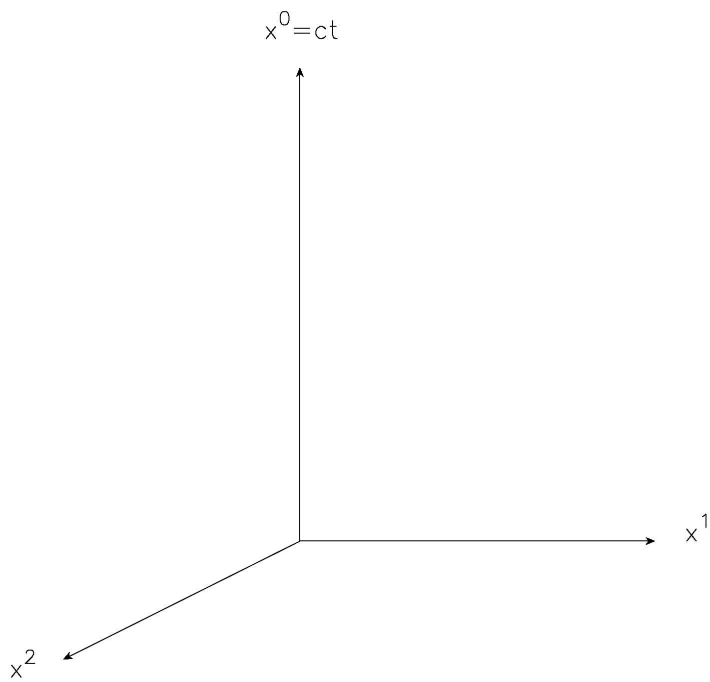


Figure 6



Chapter 2

Covariance

The key idea in SR is that all *inertial frames* are equally valid for the conduct of experiments. This means that the same physical laws should apply in all *Lorentz frames* – i.e. inertial reference frames which differ one from another by a constant velocity shift, so that one can transform from one reference frame to another by applying a *Lorentz transformation*. Expressing this more formally, it means that physical laws should be expressible in a form that makes them **covariant** under what is termed the *Lorentz group* – i.e. the group of all Lorentz transformations[1].

In GR we go one step further: physical laws should remain valid under *all* coordinate transformations. We call this the **principle of general covariance**.

---

[1]We use the term *group* here in its precise mathematical sense: a group is a set and a binary operation which acts on the elements of that set such that certain *group axioms* are satisfied. No formal knowledge of group theory is required for this course, however, and the further discussion of the group properties of Lorentz transformations in Section 3 is non-examinable.

2.1 Covariance in Newton's laws

We can consider Newton's second law of motion to illustrate the notion of covariance.

$$\vec{F} = m\vec{a} \tag{2.1}$$

is a vector equation. Vectors $\vec{F}$ and $\vec{a}$ exist independently of our choice of coordinate system, and so the physical law is independent of the coordinate system. It is a *coordinate free* description.

Alternatively equation (2.1) can be written in component form. Choosing basis vectors $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ we may write

$$\vec{F} = \sum_{i=1}^{3} F^i \vec{e}_i \tag{2.2}$$

and

$$\vec{a} = \sum_{i=1}^{3} a^i \vec{e}_i \tag{2.3}$$

Note that the superscript character i in these equations is *not* an exponent, but is a label. Our reasons for placing this label in an upper position, while the basis vectors $\vec{e}_i$ carry their labels in a lower position, will become clearer as we go along.

Thus equating components, equations (2.2) and (2.3) give

$$F^i = ma^i \tag{2.4}$$

Of course the components of $\vec{F}$ and $\vec{a}$ will have different values if different basis vectors are chosen. (We shall see below that a change in coordinate system will also generate a change in basis vectors). So the *components* of the vector equation are **not** invariant, and will change depending on one's coordinate system, even though the physical law **is** invariant, and holds regardless of one's choice of coordinate system. We say that the vector equation in component form provides a *covariant* description of Newton's law of motion; the prefix 'co' reminds us that the numerical *value* of the *co*mponents depends on our choice of *co*ordinate system.

2.2 Summation convention

Instead of constantly writing

$$\sum_i F^i \vec{e}_i \tag{2.5}$$

it is convenient simply to write

$$F^i \vec{e}_i \tag{2.6}$$

Thus, where one encounters repeated indices **in the configuration upper-lower or lower-upper**, this notation umplies that we should over that index. (Einstein once commented that this 'discovery' was one of his greatest achievements!).

Example 1:

$$\sum_i a^i\, b_i \quad \text{is written as} \quad a^i\, b_i \quad \text{or} \quad a^j\, b_j \tag{2.7}$$

Example 2

$$\sum_j \sum_i A^{ijk}\, B_{ijl} \quad \text{is written as} \quad A^{ijk}\, B_{ijl} \tag{2.8}$$

Note that "i" and "j" are called dummy indices; the choice of letter is unimportant.

2.3 Change of basis

We can also use Newton's second law to illustrate how coordinate bases transform. Let us express the vector $\vec{F}$ in a new basis $\{\vec{e'_1}, \vec{e'_2}, \vec{e'_3}\}$, i.e.

$$\vec{F} = F'^k \vec{e'_k} \tag{2.9}$$

Of course we can also write the original basis vectors $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ in terms of this new basis, i.e.

$$\vec{e}_j = A_j^k \vec{e'_k} \tag{2.10}$$

where $A_j^k \equiv \mathbf{A}$ is a non-singular matrix, since $\{\vec{e'_1}, \vec{e'_2}, \vec{e'_3}\}$ is also a coordinate basis. (Note the use of summation convention here, summing over the index k).

Conversely we may write $\{\vec{e'_1}, \vec{e'_2}, \vec{e'_3}\}$ in terms of basis $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, i.e.

$$\vec{e'_k} = B_k^l \vec{e}_l \tag{2.11}$$

Substituting equation (2.11) into equation (2.10) it follows that

$$\vec{e}_j = A_j^k B_k^l \vec{e}_l \tag{2.12}$$

or

$$A_j^k B_k^l = \delta_j^l \tag{2.13}$$

where

$$\delta_j^l = 1 \text{ when } j = l \text{ and } 0 \text{ when } j \neq l \tag{2.14}$$

δ_j^l is called the kronecker delta. Condition (2.13) we may also write in matrix form as

$$\mathbf{A}\,\mathbf{B} = \mathbf{I} \;\; \mathit{or} \;\; \mathbf{B} = \mathbf{A}^{-1}\,, \quad \mathbf{A} = \mathbf{B}^{-1} \tag{2.15}$$

Clearly

$$\vec{F} = F^i \vec{e}_i = F^i A_i^k \vec{e'_k} \tag{2.16}$$

However we also have $\vec{F} = F'^k \vec{e'_k}$, so necessarily

$$F'^k = A_i^k F^i \tag{2.17}$$

This defines the **transformation law** for the components of a vector. The A_i^k are a set of numbers which tell us precisely how the components of $\vec{F}$ in one coordinate system (with basis vectors $\vec{e}_i$) are related to the components of $\vec{F}$ in another coordinate system (with basis vectors $\vec{e}'_k$). The components of $\vec{a}$ transform in exactly the same way. Thus, if Newton's second law in component form can be written as

$$F^i = ma^i \tag{2.18}$$

in one coordinate system, in the new basis $\{\vec{e'_1}, \vec{e'_2}, \vec{e'_3}\}$ Newton's second law takes precisely the same form as in the original basis, i.e.

$$F'^i = ma'^i \tag{2.19}$$

and for that reason equation (2.18) is called **covariant**, since it takes the same **form** in all coordinate systems, even though the components will, in general, be different in different coordinate systems.

Chapter 3

Some ideas from special relativity

To gain a deeper understanding of GR we need to spend a little time thinking about the mathematics of SR. This will also give us the opportunity to discuss covariance in SR, and introduce some more notation that will be useful in GR.

In SR **events** are identified as points in the **spacetime manifold** (which we define formally in Chapter 4), and four coordinates are required to specify any particular event. These are usually taken to be a time coordinate and three spatial coordinates, which we refer to as (x^0, x^1, x^2, x^3), where x^0 denotes the time coordinate. **Note that the superscript characters here are *not* exponents, but coordinate labels**. It is commonly the case in SR measure time in the same units as space – this is a sensible choice because of the constancy of the speed of light in vacuum in any Lorentz frame. One can identify intervals of time as equivalent to the distance travelled by light in a vacuum in that time interval. Usually the spatial unit is taken to be metres, in which case time is also measured in metres – i.e. the time in which light travels one metre. (A metre is thus equivalent to $(3 \times 10^8)^{-1}$ seconds). Evidently in these units, the speed of light is 1, and we write $c = 1$. (Note that it is also quite common in SR textbooks to measure the time coordinate in terms of the variable ct; in fact in Chapter 1 we expressed the geodesic deviation in terms of this coordinate, in eq. 7).

3.1 Invariant distance in spacetime

Between any two events an invariant "distance" can be defined. Thus in an inertial frame using cartesian coordinates, the "distance" or interval between event $A(x^0_A, x^1_A, x^2_A, x^3_A)$ and event $B(x^0_B, x^1_B, x^2_B, x^3_B)$ is given by

$$s^2 = -(x^0_A - x^0_B)^2 + (x^1_A - x^1_B)^2 + (x^2_A - x^2_B)^2 + (x^3_A - x^3_B)^2 \tag{3.1}$$

Another inertial observer, using coordinates $\{x'^0, x'^1, x'^2, x'^3\}$, would obtain exactly the same value when evaluating

$$s'^2 = -(x'^0_A - x'^0_B)^2 + (x'^1_A - x'^1_B)^2 + (x'^2_A - x'^2_B)^2 + (x'^3_A - x'^3_B)^2 \tag{3.2}$$

For this reason the quantity s^2 in equation (3.1) is called an **invariant**.

Rather than two finitely separated events, consider two infinitesimally separated events whose coordinates in inertial frame S are (x^0, x^1, x^2, x^3) and $(x^0+\Delta x^0, x^1+\Delta x^1, x^2+\Delta x^2, x^3+\Delta x^3)$ respectively. The interval between the two events is given by

$$\Delta s^2 = -(\Delta x^0)^2 + (\Delta x^1)^2 + (\Delta x^2)^2 + (\Delta x^3)^2 \tag{3.3}$$

3.2 The metric in special relativity

We can write equation (3.3) in the form

$$\Delta s^2 = \eta_{\alpha\beta} \Delta x^\alpha \Delta x^\beta \tag{3.4}$$

where $\eta_{00} = -1$, $\eta_{11} = 1$, $\eta_{22} = 1$, $\eta_{33} = 1$ and all other components are zero. $\eta_{\alpha\beta}$ is the SR counterpart of a geometrical object which plays a central role in GR: the metric tensor, $g_{\alpha\beta}$. Sometimes we write the metric tensor in matrix form, i.e.

$$\mathbf{N} = \begin{pmatrix} \eta_{00} & \eta_{01} & \eta_{02} & \eta_{03} \\ \eta_{10} & \eta_{11} & \eta_{12} & \eta_{13} \\ \eta_{20} & \eta_{21} & \eta_{22} & \eta_{23} \\ \eta_{30} & \eta_{31} & \eta_{32} & \eta_{33} \end{pmatrix} \tag{3.5}$$

In equation (3.4), $\eta_{\alpha\beta}$ is diagonal and constant because:

- we only consider spatially cartesian coordinates, and inertial frames and
- in SR spacetime is **flat**.

Note that some authors define the metric tensor of SR (and in turn of GR, as we will see later) with a different **signature** – associating the minus sign in equations (3.1) – (3.3) with the terms involving not the *time* coordinate but the *spatial* coordinates. Both sign conventions for the metric are equally valid, so long as internal consistency is maintained. What is crucial is that the sign of the time and spatial coordinate terms is *different*; it is this difference which makes the geometry of 'flat' spacetime fundamentally different from the geometry of 'flat' Euclidean space.

We define $\eta^{\alpha\beta}$ to be the inverse of $\eta_{\alpha\beta}$, i.e.

$$\eta^{\alpha\lambda}\eta_{\lambda\beta} = \delta^{\alpha}_{\beta} \tag{3.6}$$

where δ^{α}_{β} is the Kronecker delta defined in equation (2.14).

3.3 Spacelike, timelike and null intervals

Δs^2 has the dimensions of length squared.

$$\begin{aligned} \Delta s^2 &> 0 \text{ the interval is spacelike} \\ \Delta s^2 &< 0 \text{ the interval is timelike} \\ \Delta s^2 &= 0 \text{ the interval is null} \end{aligned} \tag{3.7}$$

- When the interval is **spacelike** a Lorentz frame can be found in which the two events are simultaneous
- When the interval is **timelike** a Lorentz frame can be found in which the two events have the same spatial coordinates
- When the interval is **null** each event lies on the light cone of the other.

3.4 Proper time

If an interval is timelike, $\Delta\tau$ denotes the **proper time** between the two events and is given by $\Delta\tau^2 = -\Delta s^2$. The set of all events in the past and future of a particle is its **worldline**. The proper time is the time recorded by a clock which follows the same worldline as the particle – i.e. we can think of it as being 'carried along' by the particle, just as we carry along our wristwatch, for example – and hence is at rest with respect to the particle.

From equation (3.3) it follows that

$$\Delta\tau = \{(\Delta x^0)^2 - (\Delta x^1)^2 - (\Delta x^2)^2 - (\Delta x^3)^2\}^{1/2} \tag{3.8}$$

Also, since

$$\Delta x^1/\Delta x^0 = u^1$$

etc., where u^1, u^2 and u^3 are the components of the 3-velocity, u, of the particle, we may write

$$\Delta\tau = \left(1 - |u|^2\right)^{1/2} \Delta x^0 = \gamma(u)^{-1}\Delta x^0 \tag{3.9}$$

where $|u|^2 = (u^1)^2 + (u^2)^2 + (u^3)^2$ (with apologies for the very ugly notation here).

Thus between any two events on its worldline the particle will record a time of

$$\tau = \int_{E_1}^{E_2} d\tau \tag{3.10}$$

If we evaluate this time in any inertial frame this may be written

$$\tau = \int_{E_1}^{E_2} \left(1 - |u|^2\right)^{1/2} dt = \int_{E_1}^{E_2} \gamma(u)^{-1} dt \tag{3.11}$$

3.5 The Lorentz group

Consider inertial frames S, with coordinate system $\{x^0,\ x^1,\ x^2,\ x^3\}$, and S', with coordinate system $\{x'^0, x'^1, x'^2, x'^3\}$. ($S'$ travels at uniform velocity u, say, w.r.t. S, but their axes are not necessarily aligned.) We may write in matrix notation

$$\mathbf{x}' = \mathbf{\Lambda x} + \mathbf{c} \tag{3.12}$$

where $\mathbf{c}$ is a constant and $\mathbf{\Lambda}$ is a 4×4 Lorentz matrix. Noting that Λ^μ_ν is the element in the μ row and ν column of the matrix $\mathbf{\Lambda}$, in index notation equation (3.12) becomes

$$x'^\mu = \Lambda^\mu_\alpha x^\alpha + c^\mu \tag{3.13}$$

If the origins are chosen to coincide when $x^0 = x'^0 = 0$ then $c^\mu = 0$. We may write

$$\Delta x'^\mu = \frac{\partial x'^\mu}{\partial x^\alpha}\Delta x^\alpha \tag{3.14}$$

From equation (3.13) it is now obvious that

$$\Lambda^\mu_\alpha = \frac{\partial x'^\mu}{\partial x^\alpha} \tag{3.15}$$

so we could write equation (3.14) in the alternative form

$$\Delta x'^{\mu} = \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \Delta x^{\alpha} \tag{3.16}$$

To ensure the invariance of Δs^2, $\mathbf{\Lambda}$ has to satisfy

$$\mathbf{\Lambda}^{\mathrm{T}} \mathbf{N}' \mathbf{\Lambda} = \mathbf{N} \tag{3.17}$$

Matrices which satisfy equation (3.17) are called **Lorentz matrices**. In component form this condition may be written as

$$\eta'_{\mu\nu} = \frac{\partial x^{\alpha}}{\partial x'^{\mu}} \frac{\partial x^{\beta}}{\partial x'^{\nu}} \eta_{\alpha\beta} \tag{3.18}$$

It also follows that

$$\eta'^{\mu\nu} = \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\nu}}{\partial x^{\beta}} \eta^{\alpha\beta} \tag{3.19}$$

Exercise: Verify equations (3.18) and (3.19).

If S' moves along the x^1 axis of S, the Lorentz matrix takes the form

$$\begin{pmatrix} \gamma(u) & -u\gamma(u) & 0 & 0 \\ -u\gamma(u) & \gamma(u) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.20}$$

The set of transformations given by equation (3.12) forms a group. Similarly the set of Lorentz matrices forms a group. The latter may easily be demonstrated by showing that if $\mathbf{\Lambda_1}$ and $\mathbf{\Lambda_2}$ satisfy equation (3.17) then so does $\mathbf{\Lambda_1 \Lambda_2}$. The inverse of $\mathbf{\Lambda}$ also satisfies equation (3.17) as does the unit matrix, $\mathbf{I}$.

Exercise: Show that the Lorentz matrices form a group.

Exercise: Show that the transformations in equation (3.12) form a group, of which the Lorentz group is a subgroup. (This larger group is called the Poincaré group.)

3.6 Four vectors

The four velocity of a particle is defined to be

$$v^\mu = \frac{dx^\mu}{d\tau} \tag{3.21}$$

Strictly speaking equation (3.21) defines only the components in some coordinate system of the four velocity, $\vec{v}$. $\vec{v}$ exists independently of all coordinate systems. There is a natural set of basis vectors associated with any given coordinate system (we shall see this later), and for the usual cartesian coordinates $\{x^0,\ x^1,\ x^2,\ x^3\}$ in inertial frame, S, these basis vectors, $\{\vec{e}_0, \vec{e}_1, \vec{e}_2, \vec{e}_3\}$ correspond to unit displacements along the x^0, x^1, x^2 and x^3 axes. Thus we may write

$$\vec{v} = v^\mu \vec{e}_\mu \tag{3.22}$$

The components of the four velocity must transform under a Lorentz transformation in the same way as Δx^μ, since $\Delta\tau$ is an invariant. Writing out the components in a given Lorentz frame, we have

$$v^0 = \frac{dx^0}{d\tau} = \left(1 - (v)^2\right)^{-1/2} = \gamma(v) \tag{3.23}$$

$$v^1 = \frac{dx^1}{d\tau} = \frac{dx^1}{dx^0}\frac{dx^0}{d\tau} = \gamma(v)v^x \tag{3.24}$$

$$v^2 = \frac{dx^2}{d\tau} = \frac{dx^2}{dx^0}\frac{dx^0}{d\tau} = \gamma(v)v^y \tag{3.25}$$

$$v^3 = \frac{dx^3}{d\tau} = \frac{dx^3}{dx^0}\frac{dx^0}{d\tau} = \gamma(v)v^z \tag{3.26}$$

where v^x, v^y and v^z are the components of the three velocity, $\mathbf{v}$. Sometimes one sees written $v^\mu = \gamma(1, \mathbf{v})$, although this is a bit of a 'mixed metaphor' and rather ugly notation.

Notice that

$$v'^\mu = \frac{dx'^\mu}{d\tau} = \Lambda^\mu_\nu v^\nu = \frac{\partial x'^\mu}{\partial x^\nu} v^\nu \tag{3.27}$$

Exercise: In frame S a particle has four velocity $\gamma(u)(1, u, 0, 0)$. If frame S' moves at velocity, v, along x-axis of S, what are the components of the four velocity in frame S'? Comment on your answer.

The magnitude of the four velocity, $(\vec{v}.\vec{v})^{1/2}$ is defined via

$$\vec{v}.\vec{v} = \eta_{\alpha\beta} v^{\alpha} v^{\beta} \tag{3.28}$$

$\vec{v}.\vec{v}$ is often written v^2, which can be confusing. Notice that $\vec{v}.\vec{v}$ as given by equation (3.28) is an invariant – i.e. one would obtain exactly the same value when evaluated in any inertial frame. It is easy to verify that

$$\vec{v}.\vec{v} = \eta_{\alpha\beta} v^{\alpha} v^{\beta} = \gamma^2 \left(-1 + (v)^2\right) = -1 \tag{3.29}$$

Exercise: Show that

$$\eta_{\alpha\beta} v^{\alpha} v^{\beta} = {\eta'}_{\alpha\beta} {v'}^{\alpha} {v'}^{\beta}$$

3.6.1 Covariant Components

As a precursor to our discussion of *one-forms* in Section 4, we can introduce what are called the *covariant* components, v_μ, defined by

$$v_\mu = \eta_{\mu\alpha} v^{\alpha} \tag{3.30}$$

Thus $v_0 = -v^0$, $v_i = v^i$, where i takes the values from 1 to 3. Hence we may write

$$\vec{v}.\vec{v} = \eta_{\alpha\beta} v^{\alpha} v^{\beta} = v_\mu v^\mu \tag{3.31}$$

3.6.2 The four momentum

So far we have dealt with the four velocity. We can simply define the four momentum of a particle as

$$\vec{p} = m\vec{v} \tag{3.32}$$

where m is the rest mass of the particle. In component notation equation (3.32) becomes

$$p^\mu = mv^\mu \tag{3.33}$$

and is evidently valid in all inertial frames.

3.6.3 The four acceleration

We define the four acceleration as

$$\vec{a} = \frac{d\vec{v}}{d\tau} = \frac{d}{d\tau}v^\mu \vec{e_\mu} \tag{3.34}$$

In an inertial and cartesian frame in which

$$\frac{d\vec{e}_\mu}{d\tau} = 0$$

we may write

$$\vec{a} = \frac{dv^\mu}{d\tau}\vec{e_\mu} \tag{3.35}$$

In this Lorentz frame we may write $\vec{a}$ in component form as

$$a^\mu = \frac{dv^\mu}{d\tau} = \frac{d^2x^\mu}{d\tau^2} \tag{3.36}$$

The (components of the) four acceleration may also be defined as

$$a^\mu = \frac{dv^\mu}{d\tau} \tag{3.37}$$

and must transform in the same way as v^μ.

p^μ and a^μ transform in exactly the same way as v^μ. $\vec{p}$ and $\vec{a}$ provide examples of four vectors. In the same way as before, we can define the *covariant* components of these four vectors as

$$p_\mu = \eta_{\mu\alpha} p^\alpha \tag{3.38}$$

$$a_\mu = \eta_{\mu\alpha} a^\alpha \tag{3.39}$$

Exercise: Show that $p_\mu p^\mu = -m^2$.

The relativistic version of Newton's laws now has to be of the form

$$\vec{f} = m\frac{d\vec{v}}{d\tau} \tag{3.40}$$

where $\vec{f}$ is a four force; i.e. it has to be a four vector. In a Lorentz frame, equation (3.40) in component form becomes

$$f^\mu = m\frac{d^2x^\mu}{d\tau^2} \tag{3.41}$$

Conservation of four momentum will read

$$\sum_{i=1}^{N} \vec{p_{(i)}} = \vec{P}_{\rm tot} \tag{3.42}$$

where $\vec{P}_{\rm tot}$ is a constant four vector, and the sum is over N particles. In component form equation (3.42) becomes

$$\sum_{i=1}^{N} p^\mu_{(i)} = P^\mu_{\rm tot} \quad \text{for } \mu = 0,1,2,3 \tag{3.43}$$

This evidently could also be written in terms of covariant components. (Note that in equation (3.43) the subscript '(i)' is simply a label to denote the $i^{\rm th}$ particle and should not be confused with the covariant components, p_μ). Equations (3.40) and (3.42) are explicitly independent of all coordinate systems. Equations (3.43), on the other hand, which are in component form, will be valid in all frames, although the components p^μ_a will be different according to in which frame they are evaluated.

Equation (3.41), with a^μ as $d^2x^\mu/d\tau^2$ is valid in all inertial frames, but not true in arbitrary coordinate systems. Equation (3.41) is covariant under Lorentz transformations, but **not** under more general transformations.

3.7 Generalising to tensors

So far we have only discussed how four vectors transform, but we will now begin to extend our discussion to more general geometrical objects known as **tensors**. We shall usually just talk about the components of a tensor, e.g. $F_{\mu\nu}$, and their transformation properties, and sometimes simply refer to these components as the tensor itself. Mathematically this is not entirely correct, but is expedient and should not lead to confusion in the context of this course.

Thus we can simply *define* a tensor in terms of its transformation properties. (We will develop this in more detail in Chapter 4).

A simple example of a tensor would be $C^{\mu\nu} = a^\mu b^\nu$. Under a Lorentz transformation this tensor would transform as

$$C'^{\mu\nu} = \frac{\partial x'^\mu}{\partial x^\alpha}\frac{\partial x'^\nu}{\partial x^\beta}C^{\alpha\beta} \tag{3.44}$$

An equation of the form

$$C^{\mu\nu} = D^{\mu\nu} \tag{3.45}$$

where both **C** and **D** are tensors, will necessarily be valid in all Lorentz frames. Such equations are called tensor equations, or covariant equations.

The basis vectors $\vec{e}_\mu$ corresponding to coordinates $\{x^0, x^1, x^2, x^3\}$ are related to basis vectors

$\vec{e'}_\mu$ corresponding to coordinates $\{x'^0, x'^1, x'^2, x'^3\}$, via the transformation law

$$\vec{e}_\mu = \frac{\partial x'^\mu}{\partial x^\mu}\vec{e'}_\mu \tag{3.46}$$

This makes sense, as we want four vectors, $\vec{a}$, etc to be independent of coordinates, and evidently

$$\vec{a} \;=\; a^\mu \vec{e}_\mu = \; a^\mu \frac{\partial x'^\mu}{\partial x^\mu}\vec{e'}_\mu = a'^\mu \vec{e'}_\mu \tag{3.47}$$

There is a close relationship between the basis vector $\vec{e}_\mu$ and the operator $\partial/\partial x^\mu$. We shall see this when we consider curved spaces. Note that the operator has the same 'transformation' properties as $\vec{e}_\mu$, i.e.

$$\frac{\partial}{\partial x^\mu} = \frac{\partial x'^\mu}{\partial x^\mu}\frac{\partial}{\partial x'^\mu} \tag{3.48}$$

which is a relationship we use in elementary calculus, when changing variables – e.g. to perform a multi-dimensional integral. These ideas of variable transformations from calculus will play a central role in our development of transformation laws for coordinate systems in curved spacetimes.

Chapter 4

Manifolds

To deal with curved spacetime – and to allow us to define geometrical objects such as vectors and tensors in a completely general way – we have to introduce the notion of a *manifold*. A manifold is essentially a continuous space which is locally flat. More generally we can regard a manifold as any set which can be continuously parametrised: the number of independent parameters is the *dimension* of the manifold, and the parameters themselves are the *coordinates* of the manifold. A *differentiable manifold* is one which is both continuous and differentiable. This means that we can define a scalar function (or *scalar field*) – ϕ, say – at each point of the manifold, and that ϕ is differentiable. We describe this more explicitly, in terms of coordinates, in the next section.

Examples of differentiable manifolds are:

1. the two dimensional plane, $\mathbf{E^2}$
2. the two dimensional sphere, $\mathbf{S^2}$
3. a two dimensional torus (e.g. surface of a doughnut)

In this course we will be concerned with a particular class of differentiable manifolds known as *Riemannian* manifolds. A Riemannian manifold is a differentiable manifold on which a *distance*, or *metric*, has been defined.

Evidently $\mathbf{E^2}$ has a natural distance defined on it. If we take a Cartesian coordinate system $\{x, y\}$ the distance, dl, between two neighbouring points, P, with coordinates (x, y), and Q, with coordinates $(x + dx, y + dy)$, is given by

$$dl^2 = dx^2 + dy^2 \tag{4.1}$$

dl^2 is often called the *line element*. Note that the distance between P and Q is defined in a coordinate-free way – i.e. it exists independently of one's choice of coordinate system, although in equation (4.1) a Cartesian coordinate system has been chosen to represent it. If we use instead a polar coordinate system, $\{r, \theta\}$, where

$$r = (x^2 + y^2)^{\frac{1}{2}} \tag{4.2}$$

and

$$\theta = \mathrm{atan}\frac{y}{x} \tag{4.3}$$

then the distance between P and Q may be written as

$$dl^2 = dr^2 + r^2 d\theta^2 \tag{4.4}$$

The Minkowski spacetime of special relativity is another example of a Riemannian manifold, with metric defined via equation (3.4).

If one imagines a curved manifold embedded in a higher dimensional Euclidean space (e.g. a 2-D sphere embedded in $\mathbf{E^3}$) then there is a natural metric which we can adopt, which is just the Euclidean distance function of the higher-dimensional space in which the manifold is embedded. Thus examples (ii) and (iii) above are usually seen as embedded in $\mathbf{E^3}$, in which case it is straightforward to derive the expression for the line element.

Exercise: Show that the line element for the surface of a sphere of radius, R, embedded in 3-D Euclidean space is given in spherical polar coordinates by

$$dl^2 = R^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{4.5}$$

4.1 Differentiable manifolds

Consider a point, P, in a Riemannian manifold, to which a set of coordinates has been assigned.

> **Mathematical Aside:** The formal mathematics of defining coordinates for P need not concern us in this lecture course, but the interested reader can find useful discussions in any introductory textbook on differential geometry. Loosely speaking, it involves covering the points of the manifold by a collection of *open sets*, $\mathbf{U^i}$, each of which is mapped onto $\mathbf{R^n}$ by a one-to-one mapping, ϕ^i. The pair $(\mathbf{U^i}, \phi^i)$ is called a *chart*, and the collection of charts an *atlas*. One can think of each chart as defining a different coordinate system.

Suppose point P belongs to $\mathbf{U}$ and $\mathbf{U}'$. From chart $(\mathbf{U}, \phi)$, P has coordinates $\{x^1, x^2, ..., x^n\}$ and from $(\mathbf{U}', \phi')$ it has coordinates $\{x'^1, x'^2, ..., x'^n\}$. There will be a functional relationship between the two sets of coordinates; i.e. we can write

$$x'^j = f^j(x^1, x^2, ..., x^n) \tag{4.6}$$

for each $j = 1, 2, ..., n$. In somewhat loose notation we can also, for expedience, write this as

$$x'^j = x'^j(x^1, x^2, ..., x^n) \tag{4.7}$$

If the all partial derivatives of $f^j(x^1, x^2, ..., x^n)$ exist for all orders then we call the manifold differentiable. With this notion of a differentiable manifold, a great deal of geometric structure can be defined, including **functions**, **tangent vectors**, **one-forms** and **tensors**.

4.2 Functions on a manifold

One can define a function, f, on a manifold, $\mathbf{M}$. At any point, P, of the manifold the function takes a real value

$$f : \mathbf{M} \rightarrow \mathbf{R} \tag{4.8}$$

In a particular coordinate representation, P has coordinates $\{x^1, x^2, ..., x^n\}$. We may then write simply

$$f_P = f(x^1, x^2, ..., x^n) \tag{4.9}$$

In another, primed, coordinate coordinate system P has coordinates $\{x'^1, x'^2, ..., x'^n\}$. Thus we may write

$$\begin{aligned} f_P &= f(x^1, x^2, ..., x^n) \\ &= f\left(x^1(x'^1, x'^2, ..., x'^n), x^2(x'^1, x'^2, ..., x'^n), ..., x^n(x'^1, x'^2, ..., x'^n)\right) \\ &= f'(x'^1, x'^2, ..., x'^n) \end{aligned} \tag{4.10}$$

f is called a *scalar* function; this means that its numerical value at each point of the manifold is the same real number, no matter which coordinate representation is used.

4.3 Vectors and one-forms

The intuitive picture of a vector which we have learned in elementary maths and physics courses is based on the simple idea of an arrow representing a *displacement* between two points in space. In Chapter 2 we introduced the idea that a vector, $\vec{a}$, exists independently of our choice of coordinate system, but the *components* of $\vec{a}$ take different values in different coordinate systems, and we can define a transformation law for the components of the vector. Consider, for example, the displacement vector, $\vec{\Delta x}$, with components Δx^{μ} and $\Delta x'^{\mu}$ in an unprimed and primed coordinate system respectively. Recall from equation (3.14) that

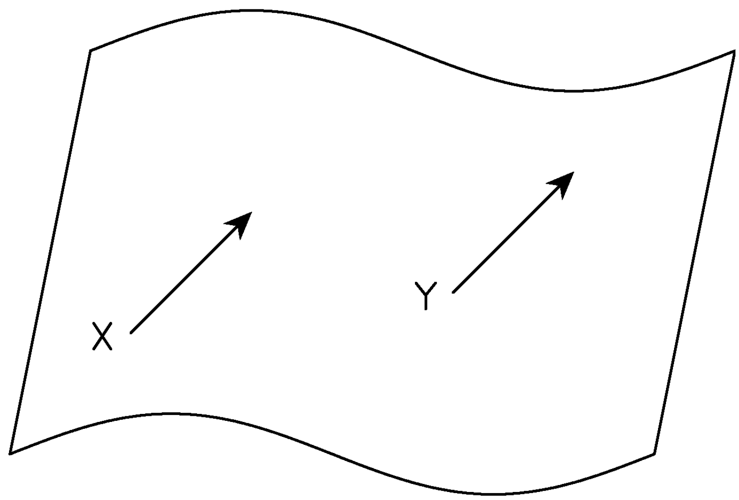
$$\Delta x'^{\mu} = \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \Delta x^{\alpha} \tag{4.11}$$

Consider now two displacement vectors, $\Delta\vec{x}$ and $\Delta\vec{y}$. How can we decide if $\Delta\vec{x}$ and $\Delta\vec{y}$ are equal when – as shown in Figure 7 – they are defined at different points on our manifold? For vectors in $\mathbf{E^3}$ with Cartesian coordinates, for example, we can simply 'translate' $\Delta\vec{y}$ to X and compare the components of $\Delta\vec{y}$ with those of $\Delta\vec{x}$. We can do this for *any* point, Y, and in *any* coordinate system, however, **only** if the transformation law for vectors is the same at every point of the manifold. This is indeed the case for Cartesian coordinates in $\mathbf{E^3}$, but will **not** be true for a general curved manifold. This is because for a curved manifold the coefficients of the transformation law in equation (4.11) are in general functions of position, i.e.

$$\frac{\partial x'^{\mu}}{\partial x^{\alpha}} = \frac{\partial x'^{\mu}(x^1, x^2, ..., x^n)}{\partial x^{\alpha}} \tag{4.12}$$

In other words, the transformation law between the primed and unprimed coordinate systems is in general *different* at different points of the manifold. Thus, it is *not* enough to define the components of a vector; we also need to specify the point of the manifold at which the vector (and its components) are defined.

Figure 7



The fact that the transformation law coefficients of equation (4.11) are in general functions of position also means that we have no 'universal' set of coordinate basis vectors on a curved manifold, as is the case for Euclidean space. (In fact, this is precisely why it is not enough simply to define the components of a vector – with respect to some fiducial set of basis vectors – since these vectors will *not* in general form a basis at every point of the manifold). There is, however, a means of defining a natural set of basis vectors for each point of the manifold which allows us to develop a more general picture of what we mean by a vector – and one which is equally valid in a curved spacetime.

4.3.1 Tangent vectors

Suppose we have a scalar function, ϕ, defined at a point, P, of a Riemannian manifold, where P has coordinates $\{x^1, x^2, ..., x^n\}$ in some coordinate system. Since our manifold is differentiable we can evaluate the derivative of ϕ with respect to each of the coordinates, x^i, for $i = 1, ..., n$. In fact, since ϕ is completely arbitrary, we can think of the derivatives as a set of n 'operators', denoted by

$$\frac{\partial}{\partial x^i}$$

These operators act on any scalar function, ϕ, and yield the rate of change of the function with respect to the x^i.

We can now define a **tangent vector** at point, P, as a linear operator of the form

$$a^\mu \frac{\partial}{\partial x^\mu} \equiv a^1 \frac{\partial}{\partial x^1} + a^2 \frac{\partial}{\partial x^2} + ... + a^n \frac{\partial}{\partial x^n} \tag{4.13}$$

(Note the use of the summation convention). This tangent vector operates on any function, ϕ, and essentially gives the rate of change of the function – or the *directional derivative* – in a direction which is defined by the numbers $(a^1, a^2, ..., a^n)$. We can define the addition of two tangent vectors in the obvious way

$$a^\mu \frac{\partial}{\partial x^\mu} + b^\mu \frac{\partial}{\partial x^\mu} = (a^\mu + b^\mu) \frac{\partial}{\partial x^\mu} \tag{4.14}$$

> **Mathematical Aside:** With this straightforward definition of addition, a little formal mathematics easily shows that the set of all tangent vectors form what is called a *vector space*

Thus, the operator,

$$a^\mu \frac{\partial}{\partial x^\mu}$$

behaves like a vector, the components of which are $(a^1, a^2, ..., a^n)$. We therefore write

$$\vec{a} = a^\mu \frac{\partial}{\partial x^\mu} \tag{4.15}$$

The n operators $\frac{\partial}{\partial x^\mu}$ can be thought of as forming a set of basis vectors, $\{\vec{e_\mu}\}$, spanning the vector space of tangent vectors at P.

What exactly do these basis vectors represent? We can find a simple geometrical picture for the $\vec{e_\mu}$ by first crystallising the notion of a *curve*, C, defined on our manifold. Our intuitive notion of a curve is simply of a connected series of points on the manifold; in the mathematical literature, however, we call this a *path*, and the term *curve* is instead reserved for the particular case of a path which has been *parametrised*.

Thus, a curve is a function which maps an interval of the real line into the manifold. Putting this more simply, a curve is a path with a real number (s, say) associated with each point of the path; we call s the parameter of the curve. Note also that once we choose a coordinate system each point on the curve has coordinates, $\{x^\mu\}$, which may also be expressed as functions of the parameter, s, i.e.

$$x^\mu = x^\mu(s) \quad \mu = 1, ..., n \tag{4.16}$$

Once we specify our coordinate system, we can consider a particular set of curves which use the *coordinates themselves* as their parameter. For example, point P with coordinates $\{x^1, x^2, ..., x^n\}$ lies on the n curves which we obtain by allowing only the value of x^i to vary along the i^{th} curve ($i = 1, ..., n$) and fixing all other coordinate values to be equal to their values at P. (To visualise a simple example, think of circles of equal latitude and longitude on the 2-sphere manifold). The basis vector, $\vec{e_i} \equiv \frac{\partial}{\partial x^i}$ can be thought of simply as the *tangent* to the i^{th} curve. This geometrical picture is illustrated in Figure 8, again for the straightforward example of the 2-sphere. Note that the basis vectors $\vec{e_\phi}$ and $\vec{e_\theta}$ are *different* at points X and Y of the manifold.

And what of a more general curve in the manifold? Here we simply connect the notion, introduced above, of a tangent vector as a directional derivative to our straightforward geometrical picture of a tangent to a curve. Figure 9 shows a curve, with parameter s, and with tangent vectors drawn at points with different parameter values. Suppose the coordinates of the points on the curve are $\{x^\mu(s)\}$, for $\mu = 1, ..., n$. Then the components, T^μ, of the tangent vector with respect to the basis $\{\vec{e_\mu}\} \equiv \{\frac{\partial}{\partial x^\mu}\}$ are simply given by

$$T^\mu = \frac{dx^\mu}{ds} \tag{4.17}$$

To sum up, we can represent vectors as tangent vectors of curves in our manifold. Once we have specified our coordinate system, we can write down the components of a vector defined at any point of the manifold with respect to the natural basis generated by the derivative operators $\{\frac{\partial}{\partial x^\mu}\}$ at that point. A vector **field** can be defined by assigning a tangent vector at every point of the manifold, so that the components, $\{a^\mu\}$ now become functions of the coordinates, i.e.

$$\vec{a} = a^\mu(x^1, x^2, ..., x^n)\,\vec{e}_\mu \tag{4.18}$$

Figure 8

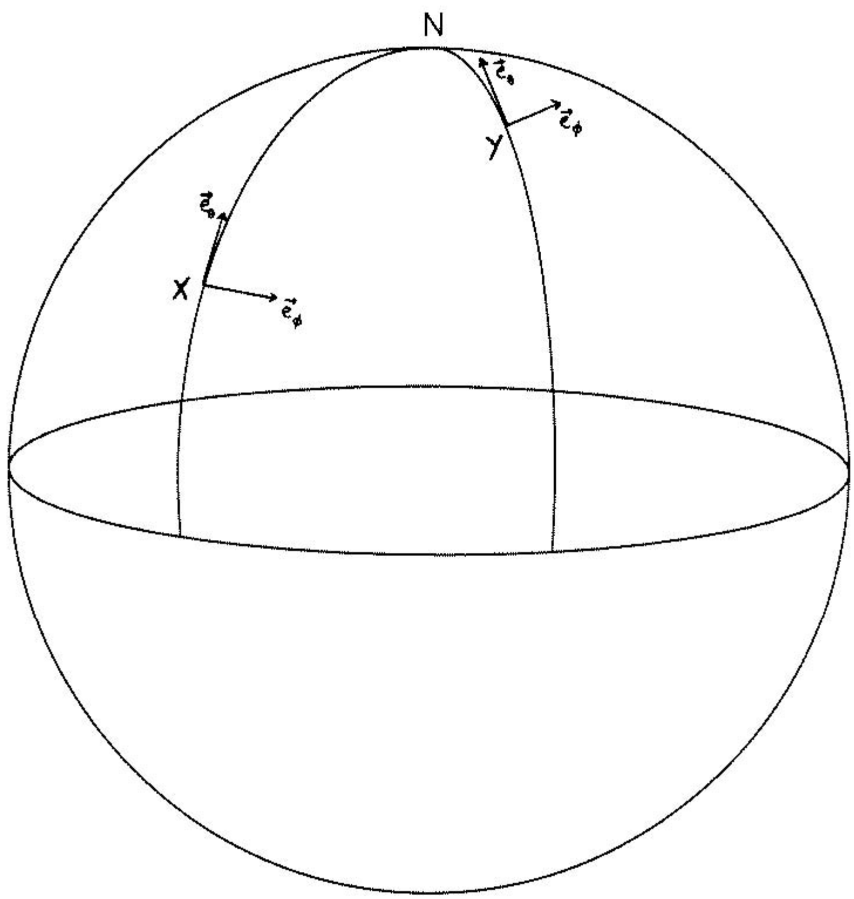
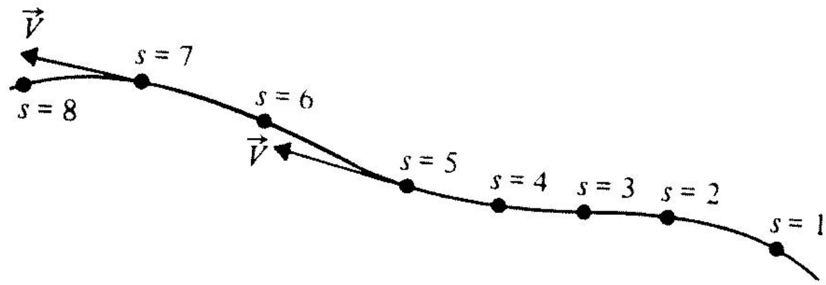


Figure 9



4.3.2 Transformation law for vectors

Suppose we change to a new coordinate system $\{x'^1, x'^2, ..., x'^n\}$. Our basis vectors are now

$$\vec{e'_\mu} \equiv \frac{\partial}{\partial x'^\mu} \tag{4.19}$$

How do the components, $\{a^1, a^2, ..., a^n\}$, transform in our new coordinate system? We have already derived this transformation law for displacements in equation (3.14). To see how the law arises within the framework of our tangent vector description, let the vector $\vec{a}$ operate on an arbitrary scalar function, ϕ. Then

$$\vec{a}(\phi) = a^\nu \frac{\partial \phi}{\partial x^\nu} \tag{4.20}$$

By the chain rule for differentiation we may write this as

$$\vec{a}(\phi) = a^\nu \frac{\partial x'^\mu}{\partial x^\nu} \frac{\partial \phi}{\partial x'^\mu} \tag{4.21}$$

However, if we write $\vec{a}$ directly in terms of coordinate basis $\{\vec{e'_\mu}\} = \{\frac{\partial}{\partial x'^\mu}\}$, we have

$$\vec{a}(\phi) = a'^\mu \frac{\partial \phi}{\partial x'^\mu} \tag{4.22}$$

Comparing equation (4.21) with (4.22) it is evident that

$$a'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} a^\nu \tag{4.23}$$

Thus the components of the tangent vector transform according to equation (4.23). We call this equation the transformation law for a **contravariant vector**, and say that the components of $\vec{a}$ transform **contravariantly**. (The term 'contravariant' is used to distinguish these vectors from another type of geometrical object – covariant vectors or 'covectors' – which we will meet in the next subsection. The more modern name for covariant vectors, however, is 'one-forms', and we will generally adopt that name in order to avoid this source of ambiguity). We denote the components of a contravariant vector as superscripts.

As we remarked above, equation (4.23) is the same transformation law as we introduced in equation (3.14) to describe the transformation of a small displacement, $\Delta\vec{x}$. To fix these ideas within our tangent vector framework, consider two neighbouring points P and Q of the manifold with coordinates $\{x^\mu\}$ and $\{x^\mu + dx^\mu\}$ respectively. Here the dx^μ are considered to be infinitesimal. In another, primed, coordinate system P and Q have coordinates $\{x'^\mu\}$ and $\{x'^\mu + dx'^\mu\}$. Since

$$x'^\mu = x'^\mu(x^1, x^2, ..., x^n) \tag{4.24}$$

it follows that

$$\begin{aligned} x'^\mu + dx'^\mu &= x'^\mu(x^1 + dx^1, x^2 + dx^2, ..., x^n + dx^n) \\ &= x'^\mu(x^1, x^2, ..., x^n) + \frac{\partial x'^\mu}{\partial x^\nu} dx^\nu \end{aligned} \tag{4.25}$$

It then follows that

$$dx'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} dx^\nu \tag{4.26}$$

Thus, the infinitesimals $\{dx^\mu\}$ transform as contravariant components. One can refer to $\{dx^1, dx^2, ..., dx^n\}$ as being a contravariant vector.

Equations (4.23) and (4.26) are two examples of the prototype transformation law for *any* contravariant vector. Any quantity, A^μ, with n components which can be evaluated in any coordinate system, and which transform in the same way as dx^μ, according to the transformation law of equation (4.26), is called a contravariant vector. (Strictly, in the coordinate-free approach one would talk of $\vec{A} = A^\mu \vec{e}_\mu$ being the vector, and A^μ its components). Thus for any contravariant vector

$$A'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} A^\nu \tag{4.27}$$

4.3.3 One-Forms

What is the relationship between the basis vectors $\vec{e'_\mu}$ and $\vec{e_\mu}$ in the primed and unprimed coordinate systems? From equation (4.19) we have

$$\vec{e'_\mu} = \frac{\partial x^\nu}{\partial x'^\mu} \vec{e_\nu} \tag{4.28}$$

(which is equivalent to equation (3.46) of Chapter 3 above).

Thus we see that the basis vectors do *not* transform in the same way as the components of a contravariant vector. This should not be too surprising, since the transformation of a basis and the transformation of components are different things: the former is the expression of *new* vectors in terms of *old* vectors; the latter is the expression of the *same* vector in terms of a new basis.

In fact, the form of the transformation in equation (4.28) is the same as the transformation law for another type of geometrical object, which we call a **covariant vector**, **covector**, or (in more modern literature) a **one-form**. Any quantity, A_μ, with n components which can be evaluated in any coordinate system, is said to be a one-form if the components transform according to the equation

$$A'_\mu = \frac{\partial x^\nu}{\partial x'^\mu} A_\nu \tag{4.29}$$

(The old-fashioned name for a one-form, a covariant vector, arose because the components transform in the *same* way as (i.e. 'co') basis vectors, while the components of a contravariant vector transform in the *opposite* way (i.e. 'contra') to basis vectors).

One can simply regard equation (4.29) as *defining* a one-form. Many modern textbooks on differential geometry, however, begin by defining a one-form as a linear mapping which acts on a vector to give a real number. (Starting from this definition one can then arrive at equation 4.29).

One-forms are usually denoted by a tilde above a symbol, just as vectors are denoted by an arrow above a symbol. Thus $\tilde{p}(\vec{a})$ is a real number. If $\tilde{p}$, $\tilde{q}$, $\tilde{r}$ and $\tilde{s}$ are one-forms, then we define their addition and scalar multiplication properties via the relations

$$\tilde{s} = \tilde{p} + \tilde{q} \equiv \tilde{s}(\vec{a}) = \tilde{p}(\vec{a}) + \tilde{q}(\vec{a}) \tag{4.30}$$

$$\tilde{r} = \alpha\tilde{p} \equiv \tilde{r}(\vec{a}) = \alpha\tilde{p}(\vec{a}) \tag{4.31}$$

for any vector, $\vec{a}$. With these rules the set of all one-forms is a *vector space*, which we call the *dual* space of of the vector space of contravariant vectors. There is a close relationship between these two vector spaces. For example, there is a natural basis of one-forms which we denote $\tilde{\omega}^\alpha \equiv \tilde{dx}^\alpha$, which are related to the basis *vectors*, $\vec{e}_\beta$, by the equation

$$\tilde{\omega}^\alpha(\vec{e}_\beta) = \delta^\alpha_\beta \tag{4.32}$$

With this basis we can write any one-form in terms of components

$$\tilde{p} = p_\alpha \tilde{\omega}^\alpha \tag{4.33}$$

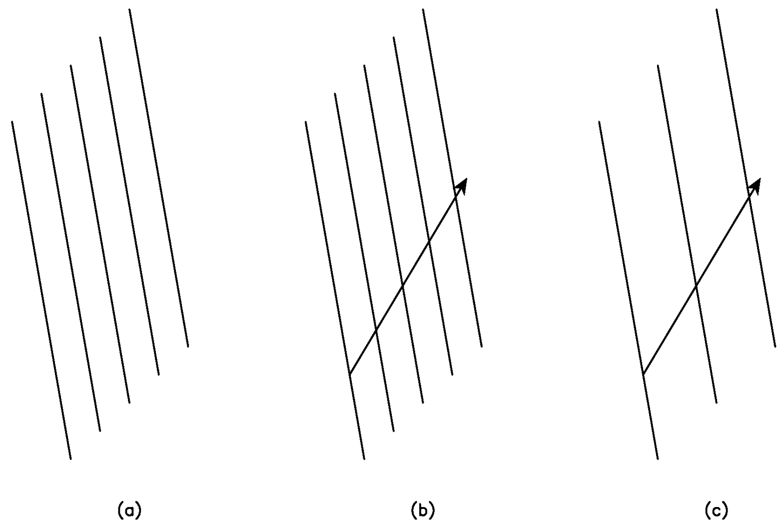
where $p_\alpha = \tilde{p}(\vec{e}_\alpha)$.

4.3.4 Picture of a One-Form

If we need a picture to represent a vector we usually think of an arrow. It is helpful to have a picture of a one-form as well. First of all, it is not an arrow. Its picture must reflect the fact that it maps vectors into real numbers. The picture generally used by mathematicians is shown in Figure (10). Here the one-form consists of a series of surfaces in the neighbourhood of a point in the manifold [panel (a)]. The 'magnitude' of the one form is given by the spacing between the surfaces: the smaller the spacing, the larger the magnitude of the one-form. Thus the one-form in panel (b) has a larger magnitude than the one-form in panel (c). In this picture, the real number produced when a one-form acts on a vector is the number of surfaces that the vector crosses; hence, the closer the spacing the larger the number. The one-form doesn't define a unique direction, since it is not a vector. Rather it defines a way of 'slicing' the manifold.

In order to justify this picture we shall look at a particular one-form: the *gradient*.

Figure 10



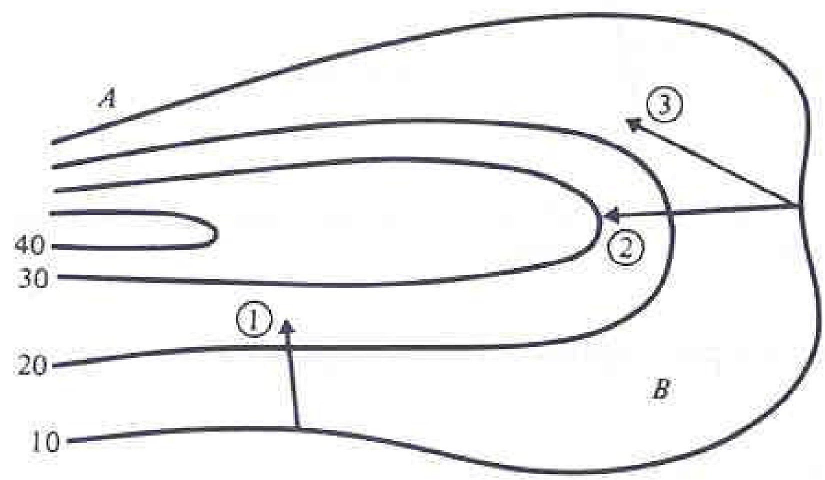
As in Section 4.3.1, consider a scalar field, ϕ, defined everywhere on a Riemannian manifold. At point P, with coordinates $\{x^1, x^2, ..., x^n\}$, form the derivatives $\{\frac{\partial\phi}{\partial x^1}, \frac{\partial\phi}{\partial x^2}, ..., \frac{\partial\phi}{\partial x^n}\}$. Suppose we now change to a new, primed, coordinate system, in which the point P has coordinates $\{x'^1, x'^2, ..., x'^n\}$. From the chain rule for differentiation we have, for each μ,

$$\frac{\partial\phi}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial\phi}{\partial x^\nu} \tag{4.34}$$

Comparing equation (4.34) with equation (4.29) we see that the gradient of ϕ fits our definition of a one-form, since its components transform according to the correct transformation law. It is usually denoted by $\tilde{d}\phi$. (In elementary courses on calculus and geometry the gradient is usually introduced as a *vector*, i.e. with a defined direction. We will see in the next section why this is justifiable, at least for Euclidean space).

The gradient enables us to justify our picture of a one-form, through the following simple example. Figure 11 is part of a topographical map, showing contours of equal elevation. If h is the elevation, then the gradient, $\tilde{d}h$, is clearly largest in an area like A, where the contour lines are closest together, and smallest near B, where the contour lines are spaced far apart. Moreover, suppose one wanted to know how much elevation a walk between two points would involve. One would lay out on the map a vector between the starting and finishing points, and the number of contours which the line crossed would give the change in elevation. For example, vector (1) crosses 1.5 contours, while vector (2) crosses 2 contours. Vector (3) starts from the same point as vector (2), but goes in a different direction and winds up only about 0.75 contours higher.

Figure 11



4.4 Tensors

Having defined what me mean by vectors and one-forms, in terms of how their components transform under a general coordinate transformation, we can now extend our definition to the more general class of geometrical object which we call *tensors*.

A tensor of type (l, m), defined on an n dimensional manifold, is a linear operator which maps l one-forms and m (contravariant) vectors into a real number (i.e. scalar). Such a tensor has a total of n^{l+m} components.

The transformation law for a general (l,m) tensor follows from its linearity, and from the transformation laws for a vector and one-form, in order that the scalar quantity obtained when the tensor operates on l one-forms and m vectors is independent of one's choice of coordinate system. We can write this general transformation law as follows

$$A'^{u_1\,u_2\,\ldots\,u_l}_{r_1\,r_2\,\ldots\,r_m} = \frac{\partial x'^{u_1}}{\partial x^{t_1}} \cdots \frac{\partial x'^{u_l}}{\partial x^{t_l}} \frac{\partial x^{q_1}}{\partial x'^{r_1}} \cdots \frac{\partial x^{q_m}}{\partial x'^{r_m}} A^{t_1\,t_2\,\ldots\,t_l}_{q_1\,q_2\,\ldots\,q_m} \tag{4.35}$$

This somewhat intimidating equation appears much more straightforward for some specific cases. First note that a contravariant vector is in fact a $(1,0)$ tensor (since it operates on a one-form to give a scalar). To see this, substitute $l=1$, $m=0$ into equation (4.35) and recover equation (4.27). Similarly a one-form is a $(0,1)$ tensor (and more trivially a *scalar* is a $(0,0)$ tensor).

A $(2,0)$ tensor, say T^{ij}, is called a *contravariant* tensor of rank 2 and transforms according to the transformation law

$$T'^{ij} = \frac{\partial x'^i}{\partial x^k}\frac{\partial x'^j}{\partial x^l}T^{kl} \tag{4.36}$$

A $(0,2)$ tensor, say B_{ij}, is called a *covariant* tensor of rank 2, and transforms according to the law

$$B'_{ij} = \frac{\partial x^k}{\partial x'^i}\frac{\partial x^l}{\partial x'^j}B_{kl} \tag{4.37}$$

An important example of a $(0,2)$ tensor is the *metric tensor*, $g_{\alpha\beta}$, which we will discuss in more detail in the next Chapter. We already met the metric tensor, $\eta_{\alpha\beta}$, in Chapter 3, in its simplified form for the Minkowski metric of special relativity. Recall from equations (3.18) and (3.19) that we gave the transformation law for Lorentz matrices which ensured the invariance of the interval, Δs^2, in different Lorentz frames. We can now see that the form of equations (3.18) and (3.19) is consistent with $\eta_{\alpha\beta}$ and $\eta^{\alpha\beta}$ being a second rank covariant and contravariant tensor respectively. Note also that our expression, in equation (3.28), for the magnitude (i.e. scalar product) of the four velocity now makes more sense: we expect a second rank covariant tensor ($\eta_{\alpha\beta}$) operating on two contravariant vectors (v^α, v^β) to give a real number.

A tensor which has both upper and lower indices, which means that it has both contravariant and covariant terms in its transformation law, is known as a *mixed tensor*. The simplest example (after the trivial case of a $(0,0)$ tensor) is a $(1,1)$ tensor, D^i_j, say. Its transformation law is

$$D'^i_j = \frac{\partial x'^i}{\partial x^k} \frac{\partial x^l}{\partial x'^j} D'^k_l \tag{4.38}$$

An important example of a $(1,1)$ tensor is the Kronecker delta, δ^i_j, which we met first in Chapter 2.

Exercise: Show that δ^i_j transforms as a $(1,1)$ tensor.

One way to construct a $(1,1)$ tensor is to take the *product* (formally the *outer product*) of a vector and one-form, i.e.

$$D^i_j = A^i B_j \tag{4.39}$$

(Clearly we may generalise this procedure to mixed tensors of higher rank).

4.4.1 Contraction of tensors

We can also take another kind of product in equation (4.39), known as the **inner product**, or **contraction** of a vector and one-form; i.e. we form the quantity $A^i B_i$ (where, as usual, the summation convention is implied). This quantity can easily be shown to be an invariant or scalar in the sense that

$$A'^j B'_j = A^i B_i \tag{4.40}$$

Exercise: Verify equation (4.40)

We can generalise the operation of contraction to the case of any two tensors, and over an arbitrary number of indices, provided that an equal number of upper and lower indices are selected. In general, contraction over k indices will produce from a tensor of type (l, m) a new tensor of type $(l - k, m - k)$. For example, the contraction of the two tensors G^{ijk}_{lm} and R^s_{tu} over the indices i and t, j and u and l and s will give the $(1, 1)$ tensor $G^{ijk}_{lm} R^l_{ij}$, where now only the indices k and m are free indices.

Chapter 5

Spacetime and the metric

5.1 The spacetime metric

Spacetime is a 4 dimensional manifold. The points of this manifold are called **events**. We can also define a *distance*, or interval, between neighbouring events – i.e. spacetime is a *Riemannian* manifold. If the interval between these events in **timelike**, we define the distance between them as the **proper time** recorded by a particle on whose worldline the events lie.

Suppose now that the separation between the events is **spacelike**. If the separation between the events is sufficiently small, one can choose as one's coordinate system a local inertial frame in which the 'distance' between the events is simply be given by (with $c = 1$).

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 \tag{5.1}$$

However, because spacetime is curved, no coordinate system can be constructed in which the distance between any two arbitrary events can always be expressed by equation (5.1). Only a **local** Lorentz frame can be found. The line element given by equation (5.1) cannot be valid throughout spacetime if the spacetime is curved.

(In exactly the same way, no coordinate system can be found in which the line element for the sphere embedded in 3D Euclidean space reduces everywhere to the Cartesian form $dl^2 = dx^2 + dy^2$)

Suppose a coordinate system has been set up in spacetime. Each event, P, is provided with 4 coordinate values, say $\{x^0, x^1, x^2, x^3\}$. These coordinates can be quite general, and are not necessarily assumed to represent time and spatial coordinates. A neighbouring event, Q, has coordinates $\{x^0 + dx^0, x^1 + dx^1, x^2 + dx^2, x^3 + dx^3\}$. We shall write the (invariant) distance between these events as

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \tag{5.2}$$

$g_{\mu\nu}$ is called the **metric tensor** (stricly speaking $\mathbf{g}$, in coordinate-free notation, is the metric tensor and $g_{\mu\nu}$ is just the metric tensor components, but this distinction is rarely made in practice).

To say that spacetime is *locally Minkowskian* (sometimes loosely stated as 'locally flat') means that there exists a coordinate transformation that reduces equation (5.2) to the form of equation (5.1) in the neighbourhood of event P. The assumption that spacetime locally reduces to Minkowski form is tantamount to assuming that it is possible to transform to a coordinate system such that $g_{\mu\nu} = \eta_{\mu\nu}$ at event P. (We shall see that in fact we can also find a transformation such that $g_{\mu\nu,\alpha} = 0$ at P. Such a coordinate system is called a geodesic coordinate system at P.)

If ds^2 is not **positive definite**, it is called a *pseudo-Riemannian* space. (A metric is positive definite if ds^2 is always greater or equal to zero, and equal to zero only when $dx^\mu = 0$).

We can choose $g_{\mu\nu}$ to be symmetric in μ and ν, which means that it has 10 independent components. Remember that in general $g_{\mu\nu}$ are functions of the coordinates (x^0, x^1, x^2, x^3).

5.2 Transformation law for the metric

At any event, P, the value of $g_{\mu\nu}$ will depend on the coordinates used. It is easy to see that the metric must, in fact, transform as a $(0, 2)$ tensor, since ds^2 is invariant, and $dx^\mu dx^\nu$ transforms as a $(2, 0)$ tensor. Thus we have the transformation law

$$g'_{\mu\nu} = \frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x^\beta}{\partial x'^\nu} g_{\alpha\beta} \tag{5.3}$$

which generalises, and justifies, equation (3.18).

5.3 Role of the metric tensor in GR

The metric tensor, $g_{\mu\nu}$, describes the geometric properties of spacetime. GR, the theory of spacetime, must describe how the intrinsic properties of spacetime are affected and determined by the presence and movement of gravitating matter.

How will $g_{\mu\nu}$ enter the field equations of GR? There are several clues. We saw in Chapter 1 that the equivalence of gravitational and inertial mass could be explained by insisting that test particles move along geodesics in spacetime. We further argued that the acceleration of the deviation of test particles (geodesic deviation, ξ) was determined by the 'curvature' of spacetime. Hence one should expect second order derivatives of $g_{\mu\nu}$ to play a key role since these are necessary to express the curvature. So if the field equations are to be covariant, we must find a tensorial quantity involving second order derivatives of $g_{\mu\nu}$ that describes the geometry of spacetime and that can be related to the matter and energy content of the Universe (which we will describe by another tensor, the *energy-momentum* tensor). There should evidently be further constraints on the theory.

- In the non-relativistic limit Newtonian gravity should be regained.

- The theory should also be consistent with SR in the limit of low matter density.

These constraints mean that the theory should be consistent with energy and momentum conservation.

5.4 Contravariant components of the metric tensor

The components g_{ij} transform as covariant components. We can also define contravariant components of the metric tensor g^{jk} by requiring

$$g_{ik}g^{jk} = \delta_i^j \tag{5.4}$$

g^{jk} defined in this way must be unique, since g_{ij} is nonsingular. g^{jk} transforms as a $(2,0)$ tensor, since g_{ik} transforms as a $(0,2)$ tensor and δ_i^j as a $(1,1)$ tensor.

Exercise: The line element in 2-D Euclidean space is given by

$$dl^2 = dx^2 + dy^2$$

Show that

$$g^{xx} = g^{yy} = 1, \quad g^{xy} = g^{yx} = 0$$

Write down the line element in terms of the coordinates $\{u, v\}$ where $u = ax + by$ and $v = cx + dy$, where a, b, c, d are constants. Using the fact that g_{ij} and g^{ij} respectively transform covariantly and contravariantly, determine them in the coordinate system $\{u, v\}$.

Exercise: Write down the contravariant and covariant components of the metric tensor for $\mathbf{E^2}$ in polar coordinates, i.e. $\{r, \theta\}$, defined by $x = r\cos\theta$, $y = r\sin\theta$.

5.5 Raising and lowering indices and contraction

Given any contravariant vector A^i it is possible to define, via the metric tensor, an associated one-form, which we denote as A_i and which is defined by

$$A_i = g_{ik} A^k \tag{5.5}$$

This operation is often called *lowering the index*.

Similarly by using g^{ij} we can raise the index of a covariant quantity B_i to obtain a contravariant quantity B^i, viz.

$$B^i = g^{ij} B_j \tag{5.6}$$

An example of raising the index would be to obtain the components of the **vector gradient**, $\vec{d\phi}$, from the one-form gradient, $\tilde{d\phi}$, defined in equation (4.34). Thus

$$(\vec{d\phi})^i = g^{ij} (\tilde{d\phi})_j \tag{5.7}$$

In elementary courses on vector calculus, the gradient is usually introduced as a vector: equation (5.6) establishes the relationship between the gradient as a one-form (as we have introduced it in this course) and as a vector. Note that from equation (5.6) it follows immediately that for Euclidean space with a Cartesian basis (for which the metric tensor is equal to the identity matrix) the components of a contravariant vector and its associated one-form are, in fact, identical. Thus, we see that in this particular case no distinction need be made between the one-form and vector description of the gradient.

The role of the metric in connecting one-forms and vectors is crucial. We can see this qualitatively by considering again the topographic map of Figure 11. Our intuitutive idea of the vector gradient involves an arrow pointing in the direction in which the contours of the map are changing most rapidly. In order to define this, we need a measure of how many contours are crossed *per unit length* in a particular arrow direction. Thus, we see that the vector gradient is larger in magnitude at position (1) than at positions (2) or (3). However, the crucial point in this picture is the phrase *per unit length* – i.e. in order to define a vector gradient we first need to define the notion of what we mean by *length*: in other words we need a *metric*.

The process of raising or lowering indices can be carried out with tensors of any rank and type. For example

$$D^{ijk..}_{lm} = g_{lp}g_{mq}D^{ijkpq} \tag{5.8}$$

Some care must be taken in positioning the indices. The dots have been placed here to indicate the indices over which contraction has taken place, although in general we shall omit the dots and just write D^{ijk}_{lm}. Note that $D^{..ijk}_{lm}$ defined by

$$D^{..ijk}_{lm} = g_{lp}g_{mq}D^{pqijk} \tag{5.9}$$

is not the same as $D^{ijk..}_{lm}$ unless D^{ijkpq} possesses some symmetry.

Exercise: $T^{\mu\nu}$ is symmetric. Show that $T_{\mu}{}^{\nu} = T^{\nu}{}_{\mu}$.

The *magnitude* of a vector A^i is $g_{ij}A^iA^j$, which is of course invariant, since g_{ij} is a $(0,2)$ tensor and A^i and A^j are both $(1,0)$ tensors. Notice

$$g_{ij}A^iA^j = A_jA^j = g^{ij}A_iA_j \tag{5.10}$$

$g_{ij}A^iB^j$ may be regarded as the scalar product of two vectors.

Chapter 6

Covariant differentiation

Any dynamical physical theory must deal in time varying quantities, and if this theory is also to be relativistic, spatially varying quantities too. Since GR is a covariant theory, we are confronted with the problem of constructing quantities that represent rates of change, but which can be defined in any coordinate system. In other words, we need to define a *derivative* which transforms covariantly under a general coordinate transformation: we call this the **covariant derivative**.

For any scalar function, say ϕ, defined on the manifold, the partial derivative

$$\phi_{,\nu} \equiv \frac{\partial \phi}{\partial x^\nu}$$

transforms as a $(0,1)$ tensor, i.e.

$$\phi'_{,\nu} \equiv \frac{\partial \phi'(x')}{\partial x'^\nu} = \frac{\partial x^\mu}{\partial x'^\nu}\frac{\partial \phi(x)}{\partial x^\mu} \equiv \frac{\partial x^\mu}{\partial x'^\nu}\phi_{,\mu} \tag{6.1}$$

Mathematical Aside: the notation in equation (6.1) may appear a little awkward. A scalar function simply maps each point, P, of the manifold to a real number. If the coordinates of P are $\{x^1, x^2, ..., x^n\}$ in an unprimed coordinate system and $\{x'^1, x'^2, ..., x'^n\}$ in another, primed, coordinate system, then the *form* of the scalar function may be different in the two coordinate systems, although $\phi(x^1, x^2, ..., x^n) = \phi'(x'^1, x'^2, ..., x'^n)$. (We saw this in equation (4.10) of Chapter 4 above). Consider, for example, $\phi(x, y) = x^2 + y^2$, defined on $\mathbf{E^2}$. Then $\phi(x, y) = \phi'(r, \theta) = r^2$.

If we take the second derivatives of ϕ, however, the quantities $\phi_{\ ,\nu\mu}$ do *not* have the transformation properties of a tensor. Indeed in general if T^{ijk}_{lmn} is a tensor, $T^{ijk}_{lmn,p}$ will usually not be. Consider for instance $A^i_{\ ,j}$.

$$\begin{aligned} A'^i_{,j} &= \frac{\partial A'^i(x')}{\partial x'^j} \\ &= \frac{\partial}{\partial x'^j}(\frac{\partial x'^i}{\partial x^k} A^k) \\ &= \frac{\partial x^l}{\partial x'^j}\frac{\partial}{\partial x^l}(\frac{\partial x'^i}{\partial x^k} A^k) \\ &= \frac{\partial x'^i}{\partial x^k}\frac{\partial x^l}{\partial x'^j} A^k_{,l} + \frac{\partial x^l}{\partial x'^j}\frac{\partial^2 x'^i}{\partial x^l \partial x^k} A^k \end{aligned} \tag{6.2}$$

Evidently the presence of the second term of equation (6.2),

$$\frac{\partial x^l}{\partial x'^j}\frac{\partial^2 x'^i}{\partial x^l \partial x^k} A^k$$

is the reason why $A^i_{\ ,j}$ does not transform as a tensor. Under linear transformations $A^i_{\ ,j}$ will behave like a tensor, since for linear transformations

$$\frac{\partial^2 x'^i}{\partial x^l \partial x^k} = 0$$

However, under nonlinear transformations this will not be the case.

The question is how can we design a derivative-like quantity that *does* transform as a tensor?

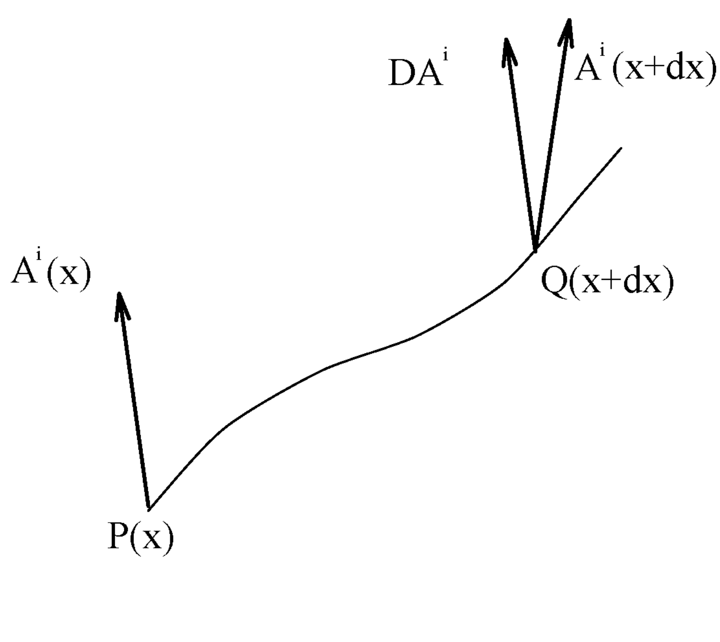
6.1 Parallel transport

The root of the problem is that in obtaining the derivative of A^i with respect to some coordinate, x, say, we compute the difference between the components $A^i(x)$ at P, and $A^i(x + dx)$ at neighbouring point Q.

> **Mathematical note:** we have simplified our notation here a little for clarity. Suppose the coordinate with respect to which we are differentiating is $x = x^1$. In fact we compute the difference between the components $A^i(x^1, x^2, ..., x^n)$ at P and $A^i(x^1 + dx^1, x^2, ..., x^n)$ at Q – i.e. with all other coordinates held fixed. For simplicity we write these expressions simply as $A^i(x)$ and $A^i(x + dx)$ respectively.

Now this difference, $A^i(x+dx) - A^i(x)$, is **not** a vector (i.e. a $(1, 0)$ tensor), since $A^i(x+dx)$ transforms as a vector component at Q and $A^i(x)$ as a vector component at P. To overcome this problem we need to introduce the important notion of **parallel transport**.

Figure 12



The idea (see Figure 12) is to displace or transport the vector from P to Q. We require this displaced quantity to transform as a vector at Q. How can this be done? We wish to construct a vector, $D\vec{A}$, at Q which is somehow related to $\vec{A}$ at P. We shall think of this as a **transport**, or displacement, of $\vec{A}$ from P to Q. We write the displaced vector $D\vec{A}$ at Q in component form as $DA^i(x+dx)$. From the linearity of tensors the difference, $A^i(x+dx) - DA^i(x+dx)$ will then transform as a vector at Q.

For this displacement to fit in with the notion of 'parallel' which we carry over from our everyday experience of flat space we should like the process or mapping to be *linear*. Hence, if $\vec{A}$ and $\vec{B}$ are transported into $D\vec{A}$ and $D\vec{B}$ respectively, then $a\vec{A}+b\vec{B}$ should be transported into $aD\vec{A}+bD\vec{B}$. We would also like the mapping to be unique – i.e. no two different vectors at P are mapped into the same vector at Q. Later, when we consider the particular case of Riemannian manifolds, we shall impose some further conditions on this transport – e.g. that it preserves the magnitude of vectors and angles between vectors – but for the time being we only demand that it is linear and unique.

If we write

$$DA^i(x+dx) = A^i(x) + \delta A^i(x) \tag{6.3}$$

then linearity requires $\delta A^i(x)$ to be linearly dependent on the components A^i. We should also expect it to depend linearly on the coordinate displacement, dx^k, between P and Q. From these considerations, we can in fact guess the form of $\delta A^i(x)$ to be

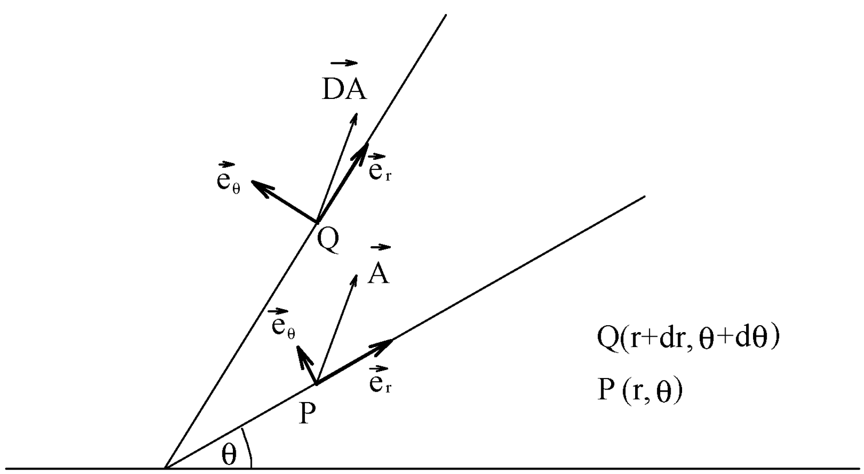
$$\delta A^i(x) = -\Gamma^i_{jk} A^j dx^k \tag{6.4}$$

(The minus sign here is just to accord with future definition of what are called the *affine connections* or *Christoffel symbols*. Usually the term Christoffel symbol is reserved for the particular case of a Riemannian space, where – as we shall see – they can be expressed in terms of the metric tensor).

6.2 Parallel transport in E^2

To justify the form of equation (6.4) consider parallel transport in $\mathbf{E^2}$. In $\mathbf{E^2}$ the displaced vector $\vec{DA}$ is, in fact, *equal* to $\vec{A}$ (see Figure 13).

Figure 13



However, even in $\mathbf{E^2}$, the components of $\vec{D}A$ at Q, in an arbitrary coordinate system, will not necessarily be the same as those of $\vec{A}$ at P. In a Cartesian coordinate system we expect that $DA^i = A^i$, but this will not be true for all coordinate systems. Consider, for example, polar coordinates $\{r, \theta\}$. We can define basis vectors $\{\vec{e}_\theta, \vec{e}_r\}$ by

$$\vec{e}_\theta = \frac{\partial \vec{r}}{\partial \theta}, \quad \vec{e}_r = \frac{\partial \vec{r}}{\partial r} \tag{6.5}$$

where

$$\vec{r} = r\cos\theta\ \vec{i} + r\sin\theta\ \vec{j} \tag{6.6}$$

and $\{\vec{i}, \vec{j}\}$ are the Cartesian basis vectors for $\mathbf{E^2}$. It is then easy to show that

$$\vec{e}_\theta = -r\sin\theta\ \vec{i} + r\cos\theta\ \vec{j} \tag{6.7}$$

$$\vec{e}_r = \cos\theta\ \vec{i} + \sin\theta\ \vec{j} \tag{6.8}$$

Now since $\vec{DA} = \vec{A}$, we must have

$$(DA^i)\ \vec{e}_i(x + dx)\ =\ A^i\ \vec{e}_i(x) \tag{6.9}$$

where $\{\vec{e}_i(x)\}$ are basis vectors at P and $\{\vec{e}_i(x + dx)\}$ are basis vectors at Q (in general different from those at P). Writing

$$\vec{e}_i(x + dx)\ =\ \vec{e}_i(x)\ +\ \frac{\partial\vec{e}_i}{\partial x^k}dx^k \tag{6.10}$$

and

$$DA^i\ =\ A^i(x) + \delta A^i(x) \tag{6.11}$$

and substituting into equation (6.9) yields, to first order,

$$\delta A^i(x)\ \vec{e}_i\ +\ \frac{\partial\vec{e}_i}{\partial x^k}A^i dx^k\ =\ 0 \tag{6.12}$$

Evidently the vector

$$\frac{\partial\vec{e}_i}{\partial x^k}$$

can itself be expressed in terms of the basis vectors at P, and we write (with foresight)

$$\frac{\partial\vec{e}_i}{\partial x^k}\ =\ \Gamma^j_{ik}\ \vec{e}_j \tag{6.13}$$

Substituting equation (6.13) into equation (6.10) we obtain

$$(\delta A^j(x) + \Gamma^j_{ik}A^i dx^k)\ \vec{e}_j\ =\ 0 \tag{6.14}$$

This is a vector equation, and necessarily each component must be zero. Thus we obtain

$$\delta A^j(x) = -\Gamma^j_{ik}\ A^i dx^k \tag{6.15}$$

Reorganising the indices we can write this precisely in the form of equation (6.4).

Consider, as an example, the case of polar coordinates. We have

$$\frac{\partial \vec{e}_\theta}{\partial \theta} = -r\cos\theta\ \vec{i} - r\sin\theta\ \vec{j} = -r\vec{e}_r \Rightarrow \Gamma^\theta_{\theta\theta} = 0, \quad \Gamma^r_{\theta\theta} = -r \tag{6.16}$$

$$\frac{\partial \vec{e}_\theta}{\partial r} = -\sin\theta\ \vec{i} + \cos\theta\ \vec{j} = \frac{\vec{e}_\theta}{r} \Rightarrow \Gamma^\theta_{\theta r} = \frac{1}{r}, \quad \Gamma^r_{\theta r} = 0 \tag{6.17}$$

$$\frac{\partial \vec{e}_r}{\partial \theta} = -\sin\theta\ \vec{i} + \cos\theta\ \vec{j} = \frac{\vec{e}_\theta}{r} \Rightarrow \Gamma^\theta_{r\theta} = \frac{1}{r}, \quad \Gamma^r_{r\theta} = 0 \tag{6.18}$$

$$\frac{\partial \vec{e}_r}{\partial r} = 0 \Rightarrow \Gamma^\theta_{rr} = 0, \quad \Gamma^r_{rr} = 0 \tag{6.19}$$

Notice that here $\Gamma^i_{jk} = \Gamma^i_{kj}$, i.e. the Christoffel symbols are symmetric in their lower indices.

Of course the Christoffel symbols, Γ^i_{jk}, will be different when evaluated in different coordinate systems. In the above example we evaluated the Christoffel symbols in polar coordinates; for Cartesian coordinates, on the other hand, they will all be zero. However we want the components DA^i to transform as vector components at Q. This obviously imposes certain constraints on the transformation properties of the Christoffel symbols. In particular Γ^i_{jk} do **not** transform as tensor components. The transformation law for the Christoffel symbols is given by

$$\Gamma'^i_{jk} = \frac{\partial x'^i}{\partial x^r}\frac{\partial x^s}{\partial x'^j}\frac{\partial x^t}{\partial x'^k}\Gamma^r_{st} + \frac{\partial x'^i}{\partial x^l}\frac{\partial^2 x^l}{\partial x'^j \partial x'^k} \tag{6.20}$$

6.3 Parallel transport for general manifolds

The preceding section showed how we may define parallel transport from an arbitrary point P to neighbouring point Q in $\mathbf{E^2}$. In fact we defined parallel transport in the obvious way, by requiring that the parallel transported vector, $\vec{D}A$, at Q is identically equal to the original vector, $\vec{A}$, at P. Equation (6.12) then expressed the difference between the components, A^i, at P and the parallel transported components, DA^i, at Q, with respect to an arbitrary basis, in terms of the **connection symbols**, Γ^i_{jk}.

As a consequence of how we defined parallel transport, the connection symbols ensure that the DA^i transform as contravariant components at Q when the A^i transform as contravariant components at P. Connection symbols which have this property define what is known as an **affine connection**, and a general manifold on which we have defined an affine connection is called an **affine manifold**.

6.4 Parallel transport of a scalar

We can also define the parallel transport of a scalar from P, with coordinates x, to Q, with coordinates $x + dx$. Thus if ϕ is a scalar function on the manifold, we define

$$D\phi(x + dx) = \phi(x) \tag{6.21}$$

With this trivial definition we have

$$\phi(x + dx) - D\phi(x + dx) = \phi(x + dx) - \phi(x) \simeq \frac{\partial \phi}{\partial x^k} dx^k \tag{6.22}$$

This quantity is evidently a scalar. No matter what coordinate system it is evaluated in it is always the same – it is simply the difference between the value of the scalar function at P and at Q. It follows that the quantity $\frac{\partial \phi}{\partial x^k}$ must therefore transform as a $(0, 1)$ tensor, since

dx^k transforms as a $(1,0)$ tensor and the LHS of equation (6.22) is an invariant. (In fact we already saw that $\frac{\partial\phi}{\partial x^k}$ transforms as a $(0,1)$ tensor in equation (6.1) above, from the chain rule for partial differentiation).

6.5 Parallel transport of a tensor

The notion of parallel transport can easily be extended to tensors of arbitrary rank, although the algebra becomes somewhat messier for higher rank tensors. However, once we have defined parallel transport of a contravariant vector through equations (6.11) to (6.15), and a scalar through equation (6.21), parallel transport of an arbitrary tensor can only take one form.

6.5.1 Parallel transport of $(0,1)$ tensor

To see how this generalisation works let us first consider the covariant components, B_i, of a $(0,1)$ tensor – i.e. a one-form. In a similar manner to our treatment of vectors, let us write the parallel transport of an arbitrary one form in terms of its covariant components B_i, viz.

$$DB_i(x+dx) = B_i(x) + \delta B_i(x) \tag{6.23}$$

Now since B_iA^i is a scalar, for arbitrary A^i, it must be displaced as a scalar. Thus, from equation (6.21)

$$D(B_iA^i) = (B_iA^i) + \delta(B_iA^i) = B_iA^i \tag{6.24}$$

Thus

$$\delta(B_iA^i) = 0 \tag{6.25}$$

from which it follows that

$$B_i\delta A^i + \delta B_i A^i = 0 \tag{6.26}$$

From equations (6.15) and (6.26) we obtain

$$\delta B_i = \Gamma^j_{ik} B_j dx^k \tag{6.27}$$

from which an expression for $DB_i(x + dx)$ then follows.

6.5.2 Parallel transport of a $(2, 0)$ tensor

Consider now the parallel transport of a $(2, 0)$ tensor, whose components are T^{ij} in some coordinate frame. In a similar manner to our treatment of vectors and one-forms, we write

$$DT^{ij}(x + dx) = T^{ij}(x) + \delta T^{ij}(x) \tag{6.28}$$

Now take two arbitrary $(0, 1)$ tensors with (covariant) components B_i and C_i. Since $B_i C_j T^{ij}$ is a scalar, we must have

$$D(B_i C_j T^{ij}) = B_i C_j T^{ij} + \delta(B_i C_j T^{ij}) = B_i C_j T^{ij} \tag{6.29}$$

Thus it follows that

$$B_i C_j \delta T^{ij} + B_i \delta C_j T^{ij} + \delta B_i C_j T^{ij} = 0 \tag{6.30}$$

Now from equation (6.27),

$$\delta B_i = \Gamma^n_{ik} B_n dx^k$$

and

$$\delta C_j = \Gamma^m_{jk} C_m dx^k$$

Substituting into equation (6.30) we obtain

$$B_i C_j \delta T^{ij} + B_i \Gamma^m_{jk} C_m dx^k T^{ij} + \Gamma^n_{ik} B_n dx^k C_j T^{ij} = 0 \tag{6.31}$$

Rearranging and changing the dummy indices we get

$$B_i C_j (\delta T^{ij} + \Gamma^j_{mk} T^{im} dx^k + \Gamma^i_{nk} T^{nj} dx^k) = 0 \tag{6.32}$$

Since B_i and C_j were arbitrary $(0, 1)$ tensors, the term in the bracket must be zero. Thus

$$\delta T^{ij} = -\Gamma^j_{mk} T^{im} dx^k - \Gamma^i_{nk} T^{nj} dx^k \tag{6.33}$$

Exactly the same reasoning can be applied to a tensor of arbitrary rank. Thus for the tensor D^{ij}_{kl} we have

$$\delta D^{ij}_{kl} = -\Gamma^i_{mp} D^{mj}_{kl} dx^p \; - \Gamma^j_{mp} D^{im}_{kl} dx^p \; + \Gamma^m_{kp} D^{ij}_{ml} dx^p \; + \Gamma^m_{lp} D^{ij}_{km} dx^p \tag{6.34}$$

Exercise: Write down an expression for δP^{klm}_i

6.6 Covariant derivative

Having defined the parallel transport of scalars, vectors one-forms and general tensors, we now have a means to define a derivative-like quantity which transforms like a tensor – i.e. the **covariant derivative**.

6.6.1 Covariant differentiation of a scalar

We have already noted above in equation (6.1) that $\partial\phi/\partial x^k$ transforms as a $(0, 1)$ tensor. (We also obtained the same result in equation (6.22), after applying a parallel displacement to ϕ). Thus we define the covariant derivative of ϕ simply to be equal to the partial derivative of ϕ.

6.6.2 Covariant differentiation of a one-form

Consider first a $(0, 1)$ tensor, **B** defined over the manifold. In some coordinate system and coordinate basis the components may be written B_i. Consider two points P and Q with coordinates x^k and $x^k + dx^k$ respectively. If we have defined parallel transport in terms of an affine connection, then the quantity $B_i(x + dx) - DB_i(x + dx)$ also transforms as a $(0, 1)$ tensor. Substituting from equation (6.23) we obtain

$$B_i(x + dx) - DB_i(x + dx) = B_i(x + dx) - B_i(x) - \delta B_i(x) \tag{6.35}$$

We can rewrite the first two terms of the right hand side as

$$B_i(x + dx) - B_i(x) = \frac{\partial B_i}{\partial x^k} dx^k = B_{i,k} dx^k \tag{6.36}$$

Thus

$$B_i(x + dx) - DB_i(x + dx) = (B_{i,k} - \Gamma^j_{ik} B_j) dx^k \tag{6.37}$$

This implies that $B_{i,k} - \Gamma^j_{ik} B_j$ transforms as a $(0, 2)$ tensor, since the LHS is a $(0, 1)$ tensor and dx^k is $(1, 0)$ tensor. We call $B_{i,k} - \Gamma^j_{ik} B_j$ the **covariant derivative** of the B_i, and denote covariant differentiation by a semi-colon, i.e.

$$B_{i;k} = B_{i,k} - \Gamma^j_{ik} B_j \tag{6.38}$$

As an example of covariant differentiation, consider the $(0, 1)$ tensor

$$\frac{\partial \phi}{\partial x^i} = \phi_{\ ,i}$$

We can now define the covariant derivative of $\phi_{\ ,i}$ to be

$$\phi_{,i;k} = \phi_{,ik} - \Gamma^j_{ik} \phi_{,j} \tag{6.39}$$

It is easy to show that $\phi_{,ik} - \phi_{,ki}$ transforms as a $(0, 2)$ tensor. If we choose the affine connection to be symmetric – i.e. $\Gamma^j_{ik} = \Gamma^j_{ki}$ – then

$$\phi_{,i;k} - \phi_{,k\ ;i} = \phi_{,ik} - \phi_{,ki} \tag{6.40}$$

A space for which this is the case is called **torsion free**.

6.6.3 Covariant differentiation of a vector

For a $(1,0)$ tensor, the covariant derivative can be defined in an analogous way to that for a one-form. Suppose that a vector field $\vec{A}$ is defined on the manifold. Choosing coordinates and the corresponding coordinate basis, we have

$$A^i(x+dx) - DA^i(x+dx) = A^i + A^i_{,k}dx^k - (A^i - \Gamma^i_{jk}A^j dx^k) = (A^i_{,k} + \Gamma^i_{jk}A^j)dx^k \quad (6.41)$$

The LHS of equation (6.41) transforms as a contravariant vector, as does dx^k. Thus $A^i_{,k} + \Gamma^i_{jk}A^j$ must transform as a $(1,1)$ tensor.

We write

$$A^i_{;k} = A^i_{,k} + \Gamma^i_{jk}A^j \quad (6.42)$$

and refer to $A^i_{;k}$ as the covariant derivative of A^i with respect to x^k.

6.6.4 Covariant differentiation of tensor

Since we have defined parallel transport for a tensor of arbitrary rank, it is straightforward to define covariant differentiation of such a tensor. For example, the covariant derivative of a $(3,3)$ tensor is given by

$$T^{ijk}_{lmn;p}dx^p = T^{ijk}_{lmn}(x+dx) - DT^{ijk}_{lmn}(x+dx) = T^{ijk}_{lmn,p}dx^p - \delta T^{ijk}_{lmn} \quad (6.43)$$

The last term on the right can easily (if lengthily!) be expressed in terms of the affine connections, so that equation (6.43) can be reduced to

$$T^{ijk}_{lmn;p} = T^{ijk}_{lmn,p} + \Gamma^i_{rp}T^{rjk}_{lmn} + \Gamma^j_{rp}T^{irk}_{lmn} + \Gamma^k_{rp}T^{ijr}_{lmn} - \Gamma^r_{lp}T^{ijk}_{rmn} - \Gamma^r_{mp}T^{ijk}_{lrn} - \Gamma^r_{np}T^{ijk}_{lmr} \quad (6.44)$$

6.7 Christoffel symbols

So far we have defined the Christoffel symbols without any reference to the metric tensor, g_{ij}. We simply assumed that some parallel transport of vectors and scalars was defined on the manifold, and showed that this notion of parallel transport could naturally be extended to tensors of arbitrary rank. However, if a metric is defined on the manifold there is a very natural definition of the Christoffel symbols, or affine connections, in terms of g_{ij}. The definition is as follows.

Suppose we require that the magnitude of a parallel displaced vector is equal to the magnitude as the original vector. More generally, suppose that the *scalar product* of two arbitrary vectors, A^i and B^j, is invariant under parallel transport. Then the Christoffel symbols are immediately and uniquely defined in terms of the metric tensor, g_{ij}. Thus if we require that

$$g_{ij}(x+dx)DA^i(x+dx)DB^j(x+dx) = g_{ij}(x)A^i(x)B^j(x) \tag{6.45}$$

for arbitrary A^i and B^j, then it follows (if we assume that the connections are symmetric in their lower indices) that

$$\Gamma^i_{jk} = \frac{1}{2}g^{il}(g_{lj,k} + g_{lk,j} - g_{jk,l}) \tag{6.46}$$

Exercise: Prove equation (6.46).

6.8 Geodesic coordinates

In a Riemannian manifold we can always find a coordinate system in which at a given point not only does g_{ij} reduce to diagonal form (i.e. space is locally Minkowskian), but the first derivates of g_{ij} are also zero, i.e. $g_{ij,k} = 0$. In such a coordinate system, equation (6.46) implies that the Christoffel symbols are all identically zero. We call such a coordinate system a **geodesic coordinate system**.

Exercise: Show that the Christoffel symbols, $\Gamma'^{\,i}_{jk}$, are zero at the point P in the coordinate system defined by

$$x'^{\mu} = x^{\mu} + \frac{1}{2}\Gamma^{\mu}_{\alpha\beta}\,(x^{\alpha} - x^{\alpha}_{P})(x^{\beta} - x^{\beta}_{P})$$

where $\Gamma^{\mu}_{\alpha\beta}$ are the Christoffel symbols evaluated at P in the unprimed coordinate system.

Chapter 7

Geodesics

Earlier we stated that material particles not acted on by forces other than gravitational forces have worldlines that are geodesics. Similarly photons also follow geodesics. Yet we haven't yet given a complete definition of a geodesic.

One can adopt one of two approaches. Either one can define a geodesic as an extremal path between two events, in the sense that the proper time along the path joining the two events is an extremum.

We shall take a different approach here, and define the geodesic in terms of parallel transport. First, recall from Chapter 4, equation (4.16), that we defined a *curve* as a parametrised path, i.e.

$$x^\mu = x^\mu(\eta) \tag{7.1}$$

where η is a parameter and μ runs over the four indices. We can define a tangent vector to this path with components

$$T^\mu = \frac{dx^\mu}{d\eta} \tag{7.2}$$

We now define a *geodesic as a curve along which the tangent vector to the curve is parallel-transported*. We do not insist that the length of the transported tangent vector should have the same length as the tangent vector at the new point (note that this would in any case only have a meaning for Riemannian spaces, for which a metric is defined). We demand only that it be parallel. (The magnitude of a tangent vector along a curve will depend on the parameterisation of the curve. A change in parametrisation gives a different curve, but the same set of points. A change in parametrisation will change the magnitude of the tangent vector).

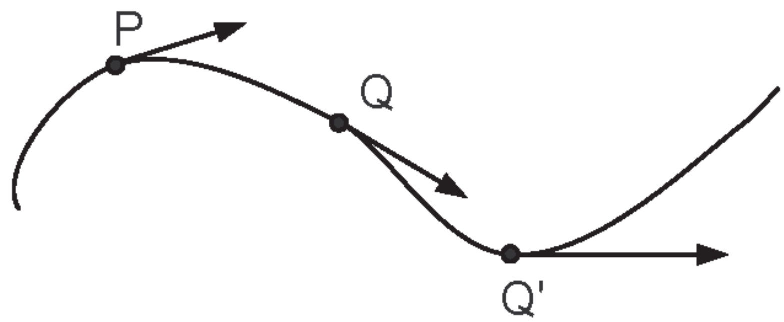
When we parallel transport a tangent vector at an arbitrary point P to a neighbouring point Q along the geodesic curve, then it is in the same direction as the tangent vector at Q. (See Figure 14).

Thus if the curve $x^\mu = x^\mu(\eta)$ is a geodesic, when the tangent vector, $\vec{T}_P$, is parallel transported from point P (with parameter value η_0, say) to point Q (with parameter value η) and tangent vector $\vec{T}_Q$, it follows that

$$\vec{T}_P \rightarrow f(\eta)\,\vec{T}_Q \tag{7.3}$$

The scalar function, $f(\eta)$, depends on the parameterisation of the curve.

Figure 14



Consider now a point, Q', in the neighbourhood of Q with parameter $\eta + d\eta$. Suppose we parallel transport $\vec{T}_P$ directly to Q'. This will yield

$$\vec{T}_P \;\rightarrow\; f(\eta + d\eta)\, \vec{T}_{Q'} \tag{7.4}$$

which must be the same as we obtain by first parallel transporting $\vec{T}_P$ to Q and then parallel transporting $f(\eta)\vec{T}_Q$ to Q'. Consider the second step, from Q to Q'. Writing the parallel-transported vector at Q' in component form, and using equations (6.11) and (6.15), we obtain

$$f(\eta + d\eta)\, T^{\mu}(\eta + d\eta) \;=\; f(\eta)T^{\mu}(\eta) - f(\eta)\Gamma^{\mu}_{\alpha\beta}\, T^{\alpha}(\eta)T^{\beta}(\eta)d\eta \tag{7.5}$$

where here we have written the term corresponding to dx^k in equation (6.15) as $T^{\beta}(\eta)d\eta$.

Expanding to first order in $d\eta$ we obtain the **geodesic equation**

$$f(\eta)\frac{dT^{\mu}}{d\eta} + T^{\mu}\frac{df}{d\eta} + f(\eta)\Gamma^{\mu}_{\alpha\beta}\, T^{\alpha}(\eta)T^{\beta}(\eta) \;=\; 0 \tag{7.6}$$

In this equation the choice of parameter η was arbitrary. The equation can be reduced to a somewhat simpler form by choosing a new parameter. A tedious, though fairly straightforward, calculation shows that a change in parameter to λ defined by

$$\lambda \;=\; -C_0 \int f(\eta)^{-1} d\eta \;+\; \lambda_0 \tag{7.7}$$

where C_0 and λ_0 are constants, gives a new geodesic equation

$$\frac{ds^\mu}{d\lambda} + \Gamma^\mu_{\alpha\beta} s^\alpha s^\beta = 0 \tag{7.8}$$

where

$$s^\mu = \frac{dx^\mu}{d\lambda}$$

λ is called an **affine parameter**. The above equation can evidently be written as

$$\frac{d^2 x^\mu}{d\lambda^2} + \Gamma^\mu_{\alpha\beta} \frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = 0 \tag{7.9}$$

Notice that any linear transformation of an affine parameter also gives affine parameter.

7.1 Geodesics of material particles

Consider the worldline of a material particle. In a given coordinate system the the worldline may be written with the proper time, τ, along the worldline as the parameter, i.e.

$$x^\mu = x^\mu(\tau) \tag{7.10}$$

The four velocity of the particle is given by

$$v^\mu = \frac{dx^\mu}{d\tau} \tag{7.11}$$

and is the tangent vector to the worldline. In this case the function, f, in equation (7.3) is simply a constant and equal to either $+1$ or -1. We can take it to be $+1$, and it then is clear from equation (7.7) that τ is an affine parameter. Thus, we obtain the geodesic equation (7.9) with τ as parameter, i.e.

$$\frac{dv^\mu}{d\tau} + \Gamma^\mu_{\alpha\beta} v^\alpha v^\beta = 0 \tag{7.12}$$

or alternatively

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta} \frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} = 0 \tag{7.13}$$

One can show (see Examples Sheet 5) that

$$g_{\alpha\beta}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} \tag{7.14}$$

is constant along a geodesic, i.e.

$$\frac{d}{d\tau}(g_{\alpha\beta}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau}) = 0 \tag{7.15}$$

7.2 Geodesics of photons

For photons, the proper time τ cannot be used to parametrize the worldlines, since $d\tau$ is zero.

If we use an arbitrary affine parameter λ the null geodesics will be described by

$$\frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = 0 \tag{7.16}$$

and since it is a null geodesic, necessarily

$$g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = 0 \tag{7.17}$$

7.3 Examples

Example 1: Consider the almost trivial example

$$ds^2 = -dt^2 + dx^2 = -d\tau^2 \tag{7.18}$$

For a material particle we have the geodesic equations

$$\frac{d^2x}{d\tau^2} = 0$$

$$\frac{d^2t}{d\tau^2} = 0$$

which have solution

$$x = A\tau + B$$

$$t = C\tau + D$$

where A, B, C and D are constants of integration. Evidently the velocity of the particle is given by A/C. Furthermore, since

$$-(\frac{dx}{d\tau})^2 + (\frac{dt}{d\tau})^2 = 1$$

we see that A and C are not independent, but satisfy $C^2 - A^2 = 1$.

For a photon we cannot use the proper time τ. Introducing an arbitrary affine parameter, λ, we have

$$\frac{d^2x}{d\lambda^2} = 0$$

$$\frac{d^2t}{d\lambda^2} = 0$$

giving

$$x = A\lambda + B, t = C\lambda + D$$

However we know that

$$(\frac{dx}{d\lambda})^2 - (\frac{dt}{d\lambda})^2 = 0 \tag{7.19}$$

which implies that $A^2 = C^2$. Thus null geodesics have the form

$$x = A\lambda + B, \quad t = A\lambda + D$$

or

$$x = -A\lambda + B, \quad t = A\lambda + D$$

Example 2: Consider now the metric

$$dl^2 = dr^2 + r^2 d\theta^2 \tag{7.20}$$

which obviously corresponds to 2-d Euclidean space in polar coordinates.

The distance, l, along a curve is an affine parameter since

$$g_{ij}\frac{dx^i}{dl}\frac{dx^j}{dl} = (\frac{dr}{dl})^2 + r^2\,(\frac{d\theta}{dl})^2 = 1 \tag{7.21}$$

Thus the geodesic equations are

$$\frac{d^2r}{dl^2} + \Gamma^r_{\theta\theta}\frac{d\theta}{dl}\frac{d\theta}{dl} + \Gamma^r_{\theta r}\frac{d\theta}{dl}\frac{dr}{dl} + \Gamma^r_{r\theta}\frac{dr}{dl}\frac{d\theta}{dl} + \Gamma^r_{rr}\frac{dr}{dl}\frac{dr}{dl} = 0 \tag{7.22}$$

and

$$\frac{d^2\theta}{dl^2} + \Gamma^\theta_{\theta\theta}\frac{d\theta}{dl}\frac{d\theta}{dl} + \Gamma^\theta_{\theta r}\frac{d\theta}{dl}\frac{dr}{dl} + \Gamma^\theta_{r\theta}\frac{dr}{dl}\frac{d\theta}{dl} + \Gamma^\theta_{rr}\frac{dr}{dl}\frac{dr}{dl} = 0 \tag{7.23}$$

In equations (6.16) - (6.19) we derived the values of the Christoffel symbols in polar coordinates. Substituting these values, the geodesic equations simplify to

$$\frac{d^2r}{dl^2} - r(\frac{d\theta}{dl})^2 = 0 \tag{7.24}$$

and

$$\frac{d^2\theta}{dl^2} + \frac{2}{r}(\frac{d\theta}{dl})^2 = 0 \tag{7.25}$$

respectively.

One obvious set of solutions to these equations is $r = l$, $\theta =$ constant. In fact any straight line will satisfy these equations.

7.4 Geodesics as extremal paths

We chose to define a geodesic by using the notion of parallel transport. This did not in fact depend on the notion of a metric, but only of an affine space. As we have already mentioned, in a metric space one could also define a geodesic as an extremal path in the sense that, along the geodesic between two events E_1, and E_2

$$\delta \int_{E_1}^{E_2} d\tau = 0 \tag{7.26}$$

We shall not prove it here, but equation (7.26) also yields the geodesic equation (7.9).

Chapter 8

The energy momentum tensor

We have now developed the necessary mathematical machinery to describe the curvature of spacetime in a manner which can be expressed *tensorially*, i.e. independently of one's choice of coordinate system. In Chapter 9 we will define the *Riemann Christoffel tensor*, which provides this geometrical description of curvature. Before we do that, however, in this Chapter we first consider the *source* of this spacetime curvature – the gravitating matter content of spacetime.

The energy momentum tensor (also known as the stress energy tensor) describes the presence and motion of gravitating matter. In this Chapter we will discuss the energy momentum tensor for the particular case of a 'perfect fluid', which is a mathematical idealisation but one which is a good approximate description of the gravitating matter in many astrophysical situations. We will derive the form of the energy momentum tensor for a perfect fluid in special relativity; this form is easily generalised to GR. Finally we will prove an important result about the *divergence* of the energy momentum tensor.

8.1 Perfect fluids

Many Newtonian gravitational problems can be considered simply as the interaction of a small number of point-like massive particles – e.g. the 2-body interaction between the Earth and the Moon can be accurately modelled in this way. Even in Newtonian theory, however, there are many contexts (e.g. the motion of stars in the Galaxy) where the number of gravitating 'particles' is too large to follow their individual trajectories. Instead we treat the system as a smooth continuum, or **fluid**, and describe its behaviour in terms of the locally averaged properties (e.g. the *density*, *velocity* or *temperature*) of the particles in each **fluid element** – by which we mean a small region of the fluid surrounding some point in the continuum within which the behaviour of the particles is fairly homogeneous.

This fluid description is also useful for many-particle systems in special relativity, although we must be careful about defining quantities such as density and velocity which are frame-dependent – i.e. we need to find a *covariant* description of the fluid (which, we will see, is why we require a *tensor* to describe the gravitating matter).

The simplest type of relativistic fluid is known as 'dust'. To a physicist, a fluid element of dust means a collection of particles which are all at rest with respect to some Lorentz frame. Many textbooks (including Schutz) refer to this Lorentz frame as the *momentarily comoving rest frame* (MCRF) of the fluid element. This name helps to reinforce the point that the fluid element as a whole may possess a bulk motion with respect to the rest of the fluid, and indeed this relative motion may not be uniform – i.e. the fluid element may be accelerating. At any moment, however, the instantaneous velocity of the fluid element allows us to define its MCRF, although the MCRF of neighbouring elements will in general be different at that instant, and the MCRF of the fluid element will also in general be different at different times.

If the fluid element is dust, however, then at any instant in the MCRF of the fluid element the individual particles possess no random motions of their own.

Generally, however, the particles within a fluid element *will* have random motions, and these will give rise to **pressure** in the fluid (c.f. motions of the molecules in an ideal gas). A fluid element may also be able to exchange energy with its neighbours via **heat conduction**, and there may be **viscous forces** present between neighbouring fluid elements. When viscous forces exist they are directed parallel to the interface between neighbouring fluid elements, and result in a **shearing** of the fluid.

A relativistic fluid element is said to be a **perfect fluid** if, in its MCRF, the fluid element has no heat conduction or viscous forces. It follows from this definition that dust is the special case of a pressure-free perfect fluid.

8.2 Definition of the energy momentum tensor

We can define the energy momentum tensor, $\mathbf{T}$, in terms of its components in some coordinate system, $\{x^1, x^2, ..., x^n\}$, for each fluid element. Thus we define $T^{\alpha\beta}$ for a fluid element to be equal to the **flux of the α component of four momentum of all gravitating matter across a surface of constant x^β**. [1]

Thus, the change, Δp^α, in the α component of the four momentum due to the flux through a surface element, ΔS_ν, at constant x^ν, is given by

$$\Delta p^\alpha = T^{\alpha\nu} \Delta S_\nu \tag{8.1}$$

[1] By 'gravitating matter' we mean here all material particles, plus (from the equivalence of matter and energy) any electromagnetic fields and particle fields which may be present, although in both this course and Gravitation II we will generally consider only material particles in the form of ordinary baryonic matter

(Note the use of the summation convention).

Mathematical aside: although we will not discuss the details of why $\mathbf{T}$ transforms as a $(2,0)$ tensor, note that this type of tensor is suggested by the definition of $\mathbf{T}$ given above. We require $\mathbf{T}$ to operate on two one-forms in order to produce a scalar quantity, the flux: the first one-form selects the component, α, of the four momentum; the second one-form selects the coordinate surface, x^β = constant, across which the flux is measured. This surface is defined by a one-form. To see this note that in general a surface will be defined by an equation of the form $\phi(x^1, x^2, ..., x^n)$ = constant, for some scalar function, ϕ. (x^β = constant is simply a special case of this). The surface may, therefore, equivalently be defined by the gradient of ϕ – i.e. the one-form, $\tilde{d}\phi$. This is why we write the surface element, ΔS_ν, with a subscript, since it transforms as a one-form.

The term 'surface element' merits some further explanation for the particular case of Minkowski spacetime. Since we are considering a 4-dimensional manifold, the components of the surface element, ΔS_ν, are in fact 3-dimensional *volume* elements. Suppose we have three linearly independent small displacement vectors, $\vec{\Delta}x_{(1)}$, $\vec{\Delta}x_{(2)}$ and $\vec{\Delta}x_{(3)}$. (Here the subscript (i) is merely a label denoting which displacement vector one is referring to). Then ΔS_ν is defined to be

$$\Delta S_\nu = \epsilon_{\nu\alpha\beta\gamma} \Delta x^\alpha_{(1)} \Delta x^\beta_{(2)} \Delta x^\gamma_{(3)}$$

where $\epsilon_{\nu\alpha\beta\gamma}$ is known as the **alternating symbol** and is a $(0,4)$ tensor which takes the value 1 if $\{\nu\alpha\beta\gamma\}$ is an **even** permutation of $\{0123\}$, and -1 if it is an **odd** permutation, and zero if any two (or more) subscripts are equal. We will have no further need to discuss the alternating symbol in this course, but we introduce it here for completeness, since it helps to explain why ΔS_ν transforms as a $(0,1)$ tensor.

Suppose we choose $\vec{\Delta}x_{(1)}$, $\vec{\Delta}x_{(2)}$ and $\vec{\Delta}x_{(3)}$ to be parallel to the x^1, x^2 and x^3 axes, and write

$$\{x^\mu_{(1)}\} \quad = \quad \{0, \Delta x, 0, 0\}$$

$$\begin{aligned}
\{x^\mu_{(2)}\} &= \{0, 0, \Delta y, 0\} \\
\{x^\mu_{(3)}\} &= \{0, 0, 0, \Delta z\}
\end{aligned}$$

then it follows from the properties of the alternating symbol that

$$\{\Delta S_\nu\} = \{\Delta x \Delta y \Delta z, 0, 0, 0\}$$

Similarly, if we take

$$\begin{aligned}
\{x^\mu_{(1)}\} &= \{\Delta t, 0, 0, 0\} \\
\{x^\mu_{(2)}\} &= \{0, \Delta x, 0, 0\} \\
\{x^\mu_{(3)}\} &= \{0, 0, \Delta y, 0\}
\end{aligned}$$

then in this case

$$\{\Delta S_\nu\} = \{0, 0, 0, \Delta t \Delta x \Delta y\}$$

from which the emergent pattern should be clear.

8.3 Components of the energy momentum tensor

Having defined **T** using equation (8.1) we now examine the individual components of **T** and consider what they each represent.

Recall from Chapter 3 that the x^0 component of the four momentum of a particle is

$$p^0 = m\gamma(v)$$

where m is the rest mass of the particle. Thus, the x^0 component of the four momentum, as measured in a particular Lorentz frame, is the total energy of the particle. (Remember that in relativistic physics energy and momentum are equivalent, so that all four components of the four momentum together specify the total energy and momentum of a particle, with the relative amount of each dependent the Lorentz frame in which they are measured). The total energy is the sum of the rest mass energy (i.e. mc^2 with $c = 1$) and kinetic energy, thus

$$\text{Total energy} \quad = \quad m + m(\gamma - 1)$$

where the second term reduces to the familiar Newtonian expression for kinetic energy in the limit of $v << 1$

Consider first the components of $\mathbf{T}$ in the MCRF of a fluid element. We begin with the '(0, 0)' component. Taking small displacement vectors parallel to the spatial axes we obtain $\{\Delta S_\nu\} = \{\Delta x \Delta y \Delta z, 0, 0, 0\}$. (See the mathematical aside above for details). Thus, from equation (8.1) it follows that

$$\Delta p^0 = T^{00} \Delta x \Delta y \Delta z \tag{8.2}$$

Hence T^{00} is the **energy density** of the fluid element. Similarly

$$\Delta p^i = T^{i0} \Delta x \Delta y \Delta z \quad (i = 1, 2, 3) \tag{8.3}$$

from which we see that T^{i0} is the i^{th} component of the **momentum density** of the fluid element.

Consider now T^{0i}, for $i = 1, 2, 3$. Taking T^{01} first, by a similar argument to that given above we choose small displacement vectors parallel to the x^0, x^2 and x^3 axes, to obtain $\{\Delta S_\nu\} = \{0, \Delta y \Delta z \Delta t, 0, 0\}$, from which it follows that

$$\Delta p^0 = T^{01} \Delta y \Delta z \Delta t \tag{8.4}$$

But $\Delta y \Delta z$ is the area of the face of our volume element perpendicular to the x^1 axis. Thus, T^{01} denotes the **energy flux** (i.e. energy per unit area per unit time) from the fluid element in the x^1 direction. Clearly the obvious corresponding interpretations follow for T^{02} and T^{03}.

Finally, consider T^{ij}, for $i, j = 1, 2, 3$. Taking T^{i1} as our illustration, we again choose small displacement vectors parallel to the x^0, x^2 and x^3 axes, to obtain the same components of ΔS_ν as above, so that

$$\Delta p^i = T^{i1} \Delta y \Delta z \Delta t \quad (i = 1, 2, 3) \tag{8.5}$$

Thus T^{i1} is the flux of the i^{th} component of the momentum in the x^1 direction, with the corresponding meaning for T^{i2} and T^{i3}.

The transfer of momentum across the interface between neighbouring fluid elements indicates that the elements exert a force on each other. If the forces are directed *perpendicular* to the interface between neighbouring fluid elements, then we see that such forces are represented by the *diagonal* T^{ij} components, i.e. with $i = j$. If, on the other hand, the forces are directed *parallel* to the interface between neighbouring fluid elements, then they are represented by the *off-diagonal* T^{ij} components, i.e. with $i \neq j$.

8.4 Symmetry of the energy momentum tensor

An important property of the energy momentum tensor which we will state here, but not prove, is that **T** is **symmetric** – i.e. in any coordinate system $T^{\alpha\beta} = T^{\beta\alpha}$, $(\alpha, \beta = 0, 1, 2, 3)$. To see a proof that the energy momentum tensor is symmetric, in the MCRF of a fluid element and hence in any frame, refer to e.g. Schutz, p. 102 et. seq.

8.5 Energy momentum tensor for a perfect fluid

8.5.1 Components of T in the MCRF for dust

In this case the energy momentum tensor takes a very simple form. Since the particles in the fluid element are at rest, there is no momentum transfer. (For a general fluid, even if the particles are at rest there can be a flux of energy and momentum through heat conduction, but not for dust, which is a perfect fluid). Also there is no momentum flux, which means that $T^{ij} = 0$, $(i, j = 1, 2, 3)$. In fact the only non-zero component is $T^{00} = \rho$, the energy density of the fluid element.

8.5.2 Components of T in the MCRF for a general perfect fluid

This case is only slightly less straightforward than that of dust. Again the T^{00} component is equal to the energy density, ρ. Since there is no bulk motion of the fluid element and there is no heat conduction for a perfect fluid, the energy flux $T^{0i} = 0$ for $i = 1, 2, 3$. Moreover, from the symmetry of **T** we also have that the momentum density, $T^{i0} = 0$, for $i = 1, 2, 3$. For the spatial components, $T^{ij} = 0$ if $i \neq j$, since these terms correspond to viscous forces parallel to the interface between fluid elements and these forces are zero for a perfect fluid. Thus T^{ij} is a diagonal matrix. But T^{ij} must be diagonal in all reference frames – e.g. under all possible rotations. This is possible only if T^{ij} is a scalar multiple of the identity matrix, i.e. $T^{11} = T^{22} = T^{33}$.

Thus, T^{ii} is the flux of the i^{th} component of momentum in the x^i direction, perpendicular to the fluid element interface. Equivalently, it is the *force per unit area*, perpendicular to the interface. This is just the **pressure**, P, exerted by the random motions of the particles in the fluid element. Hence we can write $\mathbf{T}$ as

$$\mathbf{T} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P \end{pmatrix} \tag{8.6}$$

8.5.3 Components of T in a general Lorentz frame

Consider now the components of $\mathbf{T}$ in an arbitrary Lorentz frame. We again begin with the case of dust. If we suppose that all particles in the fluid element have equal rest mass (energy), m, then we can write the energy density as

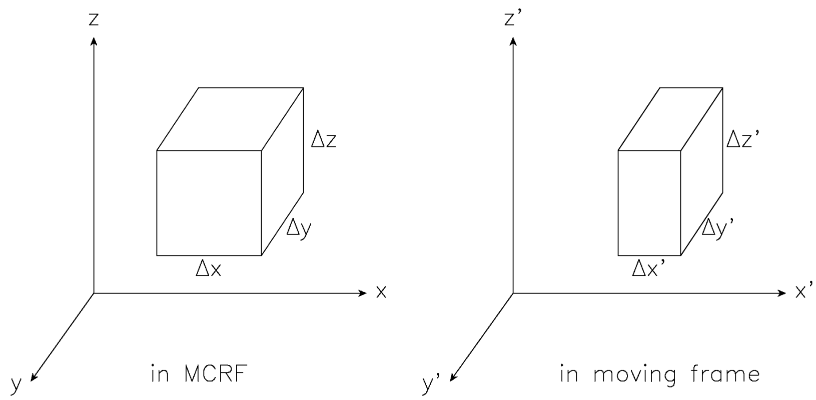
$$\rho = mn \tag{8.7}$$

where n is the *number density* of particles in the fluid element.[2] Suppose , in the MCRF, the fluid element is a cuboid with spatial dimensions Δx, Δy and Δz parallel to the x^1, x^2 and x^3 axes (see Figure 15). Then the number of particles in the fluid element in the MCRF is simply

$$N = n\Delta x\Delta y\Delta z \tag{8.8}$$

[2]Even if the particles have different rest masses, we can still write an equation of the same form as (206), replacing m with a weighted average over the fluid particles

Figure 15



Now suppose that we view the fluid element in a primed Lorentz frame which is moving parallel to the x^1 axis with velocity, v. Assuming fluid particles are neither created nor destroyed, then N will be constant but Δx will be Lorentz contracted. (Δy and Δz are unchanged). Thus, the number density measured in the primed frame will be

$$n' = n\gamma(v) \tag{8.9}$$

but the *energy* of the particles measured in the primed frame will also be increased by a factor of γ. Thus the energy density measured in the primed frame will be

$$\rho' = T'^{00} = \frac{\rho}{1-(v)^2} \tag{8.10}$$

It is straightforward, though tedious, to show that the same expression holds when the velocity, v, is not parallel to any of the coordinate axes and has arbitrary 3-velocity components $\{v^1, v^2, v^3\}$.

A similar argument then shows that the energy *flux* across the x'^j surface in the primed frame is given by

$$T'^{0j} = \frac{\rho v^j}{1-(v)^2} \tag{8.11}$$

and the i^{th} component of momentum flux across the x'^j surface by

$$T'^{ij} = \frac{\rho v^i v^j}{1-(v)^2} \tag{8.12}$$

We can collect together all of these terms into the general expression

$$T'^{\alpha\beta} \;=\; \rho u'^{\alpha} u'^{\beta} \tag{8.13}$$

where $\vec{u}' = \{u'^{\alpha}\}$ is the *four* velocity of the fluid element in the primed frame. (Compare with equations (3.23) - (3.26) to see this).

For a perfect fluid, with non-zero pressure, equation (8.13) further generalises to

$$T^{\alpha\beta} \;=\; (p+P)u^{\alpha}u^{\beta} \;+\; P\eta^{\alpha\beta} \tag{8.14}$$

(where for clarity we have dropped the primes). Here $\eta^{\alpha\beta}$ are the contravariant components of the metric tensor for Minkowski spacetime (see Chapter 3). It is easy to see that when $v = 0$, equation (8.14) reduces to equation (8.6).

8.6 Conservation of Energy and Momentum

Conservation of momentum and energy requires that

$$T^{\mu\nu}_{\;\;,\nu} = 0 \tag{8.15}$$

To see this, consider again the 3-d spatial volume of Figure 15, with dimensions Δx, Δy and Δz, and consider the **energy** passing into and out of the volume in time $\Delta t = \Delta x^0$.

The energy through the 2-d surface ΔA_x, perpendicular to the x^1 axis, in time Δt is given by

$$\Delta E_{\text{in}} = \Delta p^0_{\text{in}} = T^{01}(t,x,y,z)\Delta A_x \Delta t \tag{8.16}$$

The energy out of the 2-d surface $\Delta A_{x+\Delta x}$, perpendicular to the $x^1 \equiv x$ axis, in time Δt is similarly given by

$$\Delta E_{\text{out}} = \Delta p^0_{\text{out}} = T^{01}(t, x+\Delta x, y, z)\Delta A_x \Delta t \tag{8.17}$$

Thus the net energy into the volume in the x^1 direction is

$$\Delta E(x) = T^{01}(t, x, y, z)\Delta A_x \Delta t - T^{01}(t, x+\Delta x, y, z)\Delta A_x \Delta t \tag{8.18}$$

which can be approximated as

$$\Delta E(x) = -\frac{\partial T^{01}}{\partial x^1}\Delta x \Delta A_x \Delta t = -\frac{\partial T^{01}}{\partial x^1}\Delta V \Delta t \tag{8.19}$$

Similarly the net energy into the volume element ΔV in time Δt in the $x^2 \equiv y$ direction and $x^3 \equiv z$ direction respectively is given by

$$\Delta E(y) = -\frac{\partial T^{02}}{\partial x^2}\Delta V \Delta t \tag{8.20}$$

and

$$\Delta E(z) = -\frac{\partial T^{03}}{\partial x^3}\Delta V \Delta t \tag{8.21}$$

The total increase in energy inside the volume ΔV in time Δt is thus given by

$$\Delta E = -(\frac{\partial T^{01}}{\partial x^1} + \frac{\partial T^{02}}{\partial x^2} + \frac{\partial T^{03}}{\partial x^3})\Delta V \Delta t \tag{8.22}$$

On the other hand the increase in energy within ΔV is also simply

$$\Delta E = T^{00}(t+\Delta t, x, y, z)\Delta V - T^{00}(t, x, y, z)\Delta V \tag{8.23}$$

which is approximately

$$\frac{\partial T^{00}}{\partial x^0}\Delta V \Delta t \tag{8.24}$$

Equating (8.22) with (8.24) yields

$$\frac{\partial T^{00}}{\partial x^0} + \frac{\partial T^{01}}{\partial x^1} + \frac{\partial T^{02}}{\partial x^2} + \frac{\partial T^{03}}{\partial x^3} = 0 \tag{8.25}$$

This establishes equation (8.15) for the case where $\mu = 0$. A similar procedure, considering the change in momentum flux across ΔV, will yield the remaining equations for $\mu = 1, 2, 3$. Thus, the **divergence** of the energy momentum tensor is equal to zero.

8.7 Extending to GR

In Chapter 1 we introduced the strong principle of equivalence which stated that, in a local inertial frame, all physical phenomena are in agreement with special relativity. In the light of our study of tensors in this course, we can write down an immediate consequence of the strong principle of equivalence as follows

Any physical law which can be expressed as a tensor equation in SR has exactly the same form in a local inertial frame of a curved spacetime

This statement holds since, in the LIF, physics – and hence the form of physical laws – is indistinguishable from the physics of SR. This is a very important result because it allows us to generalise the form of physical laws which are valid in SR to the case of GR, with semi-colons (denoting covariant derivatives) replacing commas (denoting partial derivatives) where appropriate.

How is this extension justified? From the principle of covariance a tensorial description of physical laws must be equally valid in any reference frame. Thus, if a tensor equation holds in one frame it must hold in any frame. In particular, a tensor equation derived in a LIF (i.e. assuming SR) remains valid in an arbitrary reference frame (i.e. assuming GR).

Hence, the energy momentum tensor for a perfect fluid in GR takes the form

$$T^{\mu\nu} = (\rho + P)u^{\mu}u^{\nu} + Pg^{\mu\nu} \tag{8.26}$$

where $g^{\mu\nu}$ denotes the contravariant metric tensor for a general curved spacetime (which of course reduces locally to $\eta^{\mu\nu}$).

We can extend to GR in this way the result of equation (8.15), on the conservation of energy and momentum. Thus, for a fluid element in a general curved spacetime

$$T^{\mu\nu}_{;\nu} = 0 \tag{8.27}$$

If this were *not* the case – i.e. if there existed some point, P, at which $T^{\mu\nu}_{;\nu} \neq 0$ – then we could construct a LIF at P (e.g. by changing to geodesic coordinates) in which all Christoffel symbols are zero. In this new frame covariant derivatives reduce to partial derivatives, implying that $T^{\mu\nu}_{,\nu} \neq 0$, which contradicts equation (8.15).

The general technique of using the principles of covariance and equivalence to extend the validity of tensor equations from SR to GR, usually by evaluating their components in the LIF where Christoffel symbols vanish, is a very powerful one and is commonly met in the relativity literature. It is sometimes referred to informally as the 'comma goes to semi colon rule'.

Chapter 9

The Riemann Christoffel tensor

GR explains gravitational effects as consequences of the curvature of spacetime, which arises from the presence and motion of gravitating matter. In the previous Chapter we outlined how the matter content of spacetime could be described by the energy momentum tensor, $\mathbf{T}$. Thus our remaining objective is to relate $\mathbf{T}$ to the curvature of spacetime.

The curvature of a metric space is described by the **Riemann Christoffel tensor** (often also referred to simply as the **Riemann tensor**). One would expect any tensorial quantity that describes the curvature at least to depend on the second order derivatives of the metric tensor, as for a suitable choice of coordinates (geodesic coordinates) we saw previously that the first order derivatives can always be made zero. Thus we should expect the Riemann Christoffel tensor to depend on $g_{\mu\nu}$, $g_{\mu\nu,\alpha}$ and $g_{\mu\nu,\alpha,\beta}$. (N.B. in future we shall simply write $g_{\mu\nu,\alpha\beta}$ for $g_{\mu\nu,\alpha,\beta}$).

The form of the Riemann Christoffel tensor can be derived in several different ways, although we shall discuss here only one derivation – which follows on from our preliminary discussion of spacetime curvature in Chapter 1.

9.1 Riemann Christoffel tensor from the geodesic deviation

We consider two free falling test particles above the Earth and measure their separation in a frame free-falling with the particles (i.e. a LIF). In the free falling frame the particles will accelerate towards one another, albeit almost imperceptibly. Their worldlines are geodesics in spacetime, and the geodesic deviation is the separation of these geodesics. The **acceleration of the geodesic deviation** tells us that spacetime is curved.

Consider then two test particles (labelled 1 and 2) moving along nearby geodesics (see Figure 16). Let $\xi^\mu(\tau)$ denote the (infinitesimal) separation of the particles at proper time τ, so that

$$x_2^\mu(\tau) = x_1^\mu(\tau) + \xi^\mu(\tau) \tag{9.1}$$

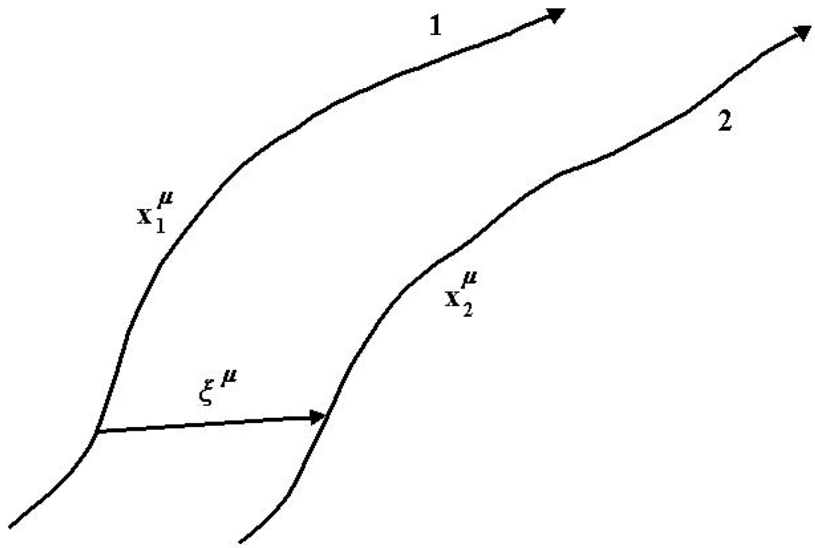
Now the worldlines of the two particles are described by the Geodesic equation, viz:

$$\frac{d^2 x_1^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta}(x_1)\frac{dx_1^\alpha}{d\tau}\frac{dx_1^\beta}{d\tau} = 0 \tag{9.2}$$

and

$$\frac{d^2 x_2^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta}(x_2)\frac{dx_2^\alpha}{d\tau}\frac{dx_2^\beta}{d\tau} = 0 \tag{9.3}$$

Figure 16



Note that we may write

$$\Gamma^\mu_{\alpha\beta}(x_2) = \Gamma^\mu_{\alpha\beta}(x_1 + \xi) = \Gamma^\mu_{\alpha\beta}(x_1) + \Gamma^\mu_{\alpha\beta,\gamma}\xi^\gamma \tag{9.4}$$

Subtracting equation (9.3) from equation (9.2) and keeping only up to first order terms in ξ yields the following equation for the acceleration of ξ^μ (dropping the subscript 1)

$$\frac{d^2\xi^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta}v^\alpha\frac{d\xi^\beta}{d\tau} + \Gamma^\mu_{\alpha\beta}v^\beta\frac{d\xi^\alpha}{d\tau} + \Gamma^\mu_{\alpha\beta,\gamma}\xi^\gamma v^\alpha v^\beta = 0 \tag{9.5}$$

Equation (9.5) can be put into covariant form by introducing the covariant operator $D/D\tau$.

9.1.1 Definition of $D/D\tau$

Consider an arbitrary vector field, A^μ, and a worldline defined by $x^\alpha = x^\alpha(\tau)$. We define

$$\frac{DA^\mu}{D\tau} = \frac{dA^\mu}{d\tau} + \Gamma^\mu_{\alpha\beta}A^\alpha\frac{dx^\beta}{d\tau} \tag{9.6}$$

$D/D\tau$ is essentially the covariant derivative along the worldline. To see how equation (9.6) is obtained, we displace $A(\tau)$ through parallel transport from the event at proper time τ on the worldline of a particle to the event at $\tau + d\tau$. Thus

$$\frac{DA^\mu}{D\tau} = \lim_{d\tau \to 0} \frac{A^\mu(\tau + d\tau) - DA^\mu(\tau)}{d\tau} \tag{9.7}$$

Since

$$DA^\mu = A^\mu - \Gamma^\mu_{\alpha\beta} A^\alpha dx^\beta \tag{9.8}$$

substitution into equation (9.7) yields

$$\frac{DA^\mu}{D\tau} = \frac{dA^\mu}{d\tau} + \Gamma^\mu_{\alpha\beta} A^\alpha \frac{dx^\beta}{d\tau} \tag{9.9}$$

Writing

$$v^\beta = \frac{dx^\beta}{d\tau} \tag{9.10}$$

we have

$$\frac{DA^\mu}{D\tau} = \frac{dA^\mu}{d\tau} + \Gamma^\mu_{\alpha\beta}\ A^\alpha v^\beta \tag{9.11}$$

or in operator form

$$\frac{D}{D\tau} = \frac{d}{d\tau} + \Gamma^\mu_{\alpha\beta} v^\beta \tag{9.12}$$

Exercise: Show that the geodesic equation may be written as

$$\frac{Dv^\mu}{D\tau} = 0$$

9.1.2 Covariant differential equation for the Riemann Christoffel tensor

From equation (9.12) it follows that

$$\frac{D^2\xi^\mu}{D\tau^2} = \frac{D}{D\tau}(\frac{D\xi^\mu}{D\tau}) = \frac{d}{d\tau}(\frac{D\xi^\mu}{D\tau}) + \Gamma^\mu_{\sigma\delta}\frac{D\xi^\sigma}{D\tau} v^\delta \tag{9.13}$$

Substituting for $D\xi^\mu/D\tau$ in equation (9.13) we obtain

$$\frac{D^2\xi^\mu}{D\tau^2} = \frac{d}{d\tau}(\frac{d\xi^\mu}{d\tau} + \Gamma^\mu_{\alpha\beta}\xi^\alpha v^\beta) + \Gamma^\mu_{\sigma\delta}(\frac{d\xi^\sigma}{d\tau} + \Gamma^\sigma_{\alpha\beta}\ \xi^\alpha v^\beta)v^\delta \tag{9.14}$$

Now

$$\frac{d}{d\tau}(\Gamma^\mu_{\alpha\beta}\xi^\alpha v^\beta) = \Gamma^\mu_{\alpha\beta,\gamma}\frac{dx^\gamma}{d\tau}\xi^\alpha v^\beta + \Gamma^\mu_{\alpha\beta}\frac{d\xi^\alpha}{d\tau}v^\beta + \Gamma^\mu_{\alpha\beta}\xi^\alpha\frac{dv^\beta}{d\tau} \tag{9.15}$$

Since each particle's worldline is a geodesic we have

$$\frac{dv^\beta}{d\tau} = \frac{d^2x^\beta}{d\tau^2} = -\Gamma^\beta_{\sigma\delta}v^\sigma v^\delta \tag{9.16}$$

where we have written

$$v^\beta = \frac{dx^\beta}{d\tau}$$

Substituting equations (9.15) and (9.16) into (9.14) and reordering we obtain

$$\begin{aligned}\frac{D^2\xi^\mu}{D\tau^2} = \frac{d^2\xi^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta,\gamma}v^\gamma\xi^\alpha v^\beta + \Gamma^\mu_{\alpha\beta}\frac{d\xi^\alpha}{d\tau}v^\beta + \Gamma^\mu_{\sigma\delta}\frac{d\xi^\sigma}{d\tau}v^\delta \\ +(\Gamma^\mu_{\beta\delta}\Gamma^\beta_{\alpha\sigma} - \Gamma^\mu_{\alpha\beta}\Gamma^\beta_{\alpha\delta})v^\sigma v^\delta\xi^\alpha\end{aligned} \tag{9.17}$$

However from equation (9.5)

$$\frac{d^2\xi^\mu}{d\tau^2} = -(\Gamma^\mu_{\alpha\beta}v^\alpha\frac{d\xi^\beta}{d\tau} + \Gamma^\mu_{\alpha\beta}v^\beta\frac{d\xi^\alpha}{d\tau} + \Gamma^\mu_{\alpha\beta,\gamma}\xi^\gamma v^\alpha v^\beta) \tag{9.18}$$

and substitution into equation (9.17) yields

$$\frac{D^2\xi^\mu}{D\tau^2} = -R^\mu{}_{\alpha\beta\gamma}v^\alpha\xi^\beta v^\gamma \tag{9.19}$$

where

$$R^\mu{}_{\alpha\beta\gamma} = \Gamma^\sigma_{\alpha\beta}\Gamma^\mu_{\sigma\gamma} - \Gamma^\sigma_{\alpha\gamma}\Gamma^\mu_{\sigma\beta} + \Gamma^\mu_{\alpha\beta,\gamma} - \Gamma^\mu_{\alpha\gamma,\beta} \tag{9.20}$$

The $(1,3)$ tensor, $\mathbf{R}$, is the Riemann Christoffel tensor. Notice that if the spacetime is flat then

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

i.e. all components of the Riemann Christoffel tensor are identically zero.

9.1.3 Fully covariant form of the Riemann Christoffel tensor and its symmetries

We may obtain the components of the $(0,4)$ tensor corresponding to $R^{\mu}{}_{\alpha\beta\gamma}$ by using the covariant components of the metric tensor to lower the index, viz.

$$R_{\alpha\beta\gamma\delta} = g_{\alpha\sigma} R^{\sigma}{}_{\beta\gamma\delta} \tag{9.21}$$

$R_{\alpha\beta\gamma\delta}$ has the following important symmetries, which reduce the number of independent components in $4d$ to 20.

$$R_{\alpha\beta\gamma\delta} = -R_{\alpha\beta\delta\gamma} \tag{9.22}$$

$$R_{\beta\alpha\gamma\delta} = -R_{\alpha\beta\gamma\delta} \tag{9.23}$$

$$R_{\alpha\beta\gamma\delta} + R_{\alpha\delta\beta\gamma} + R_{\alpha\gamma\delta\beta} = 0 \tag{9.24}$$

In equation (9.24) we have simply taken a cyclic permutation on the last three indices.

It is fairly straightforward, though tedious, to derive the above symmetry relations if we work in a geodesic coordinate system, in which it follows from equation (6.46) that

$$R_{\alpha\beta\gamma\delta} = \frac{1}{2}(g_{\alpha\delta,\beta\gamma} - g_{\alpha\gamma,\beta\delta} + g_{\beta\gamma,\alpha\delta} - g_{\beta\delta,\alpha\gamma}) \tag{9.25}$$

Since the symmetries hold in geodesic coordinates, we argue that they hence must generally be valid. Note also that equation (9.25) shows the Riemann Christoffel tensor to depend on second derivatives of the metric tensor, as predicted.

9.2 The Bianchi identities

The Bianchi identities, for the derivatives of the Riemann Christoffel tensor, state that

$$R_{\alpha\beta\gamma\delta;\lambda} + R_{\alpha\beta\lambda\gamma;\delta} + R_{\alpha\beta\delta\lambda;\gamma} = 0 \tag{9.26}$$

Note that the three terms cyclically permute the final three indices. We will not discuss in any detail the proof of the Bianchi identities in this course, although the proof is again quite straightforward if one chooses first a geodesic coordinate system and then generalises to any frame by appealing to the tensorial nature of equation (9.26). For completeness, a summary of the proof is appended at the end of this section.

9.3 Riemann tensor via parallel displacement

Another geometrically instructive method for deriving the form of the Riemann Christoffel tensor involves parallel transporting a vector around a closed curve in spacetime. If the spacetime is flat, then the vector is mapped back onto itself when the loop is completed – i.e. when we return to the starting point (see Figure 17a). In a curved spacetime, however, the parallel transported vector will not in general be mapped onto itself when we return to the starting point (see Figure 17b). We can express the net change in the components of the vector, after transport around the closed loop, in terms of the Riemann Christoffel tensor. We will not discuss this approach any further here, but the interested reader is referred to e.g.

Schutz, Chap. 6, for more details. Obviously, one obtains the same form for the Riemann Christoffel tensor when one derives it by considering parallel displacement in this way.

Figure 17a

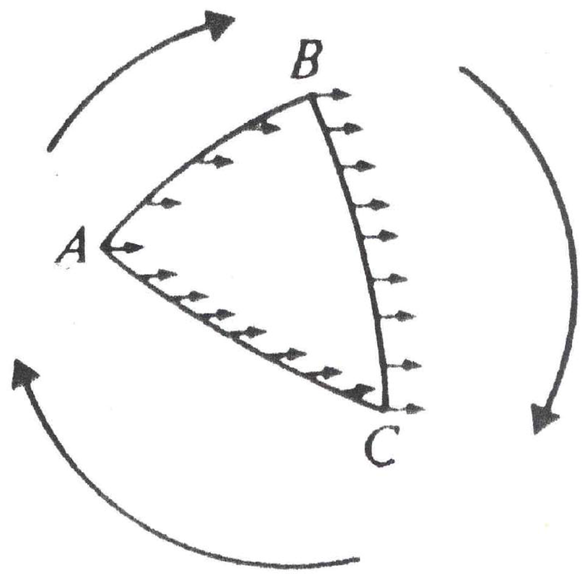
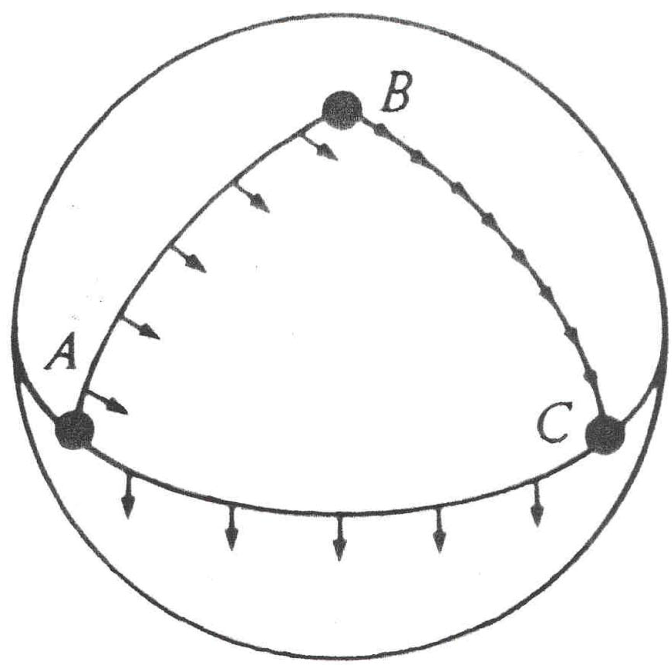


Figure 17b



9.4 Proof of the Bianchi identities

The left hand side of equation (9.26) is a $(0,5)$ tensor. Hence if we can show that its components are zero at an arbitrary point on the manifold in a given coordinate system, then they must be zero in all coordinate systems. We choose geodesic coordinates at arbitrary point, P. From the definition of the Riemann Christoffel tensor

$$R^{\mu}{}_{\beta\gamma\delta} = \Gamma^{\sigma}_{\beta\gamma}\Gamma^{\mu}_{\sigma\delta} - \Gamma^{\sigma}_{\beta\delta}\Gamma^{\mu}_{\sigma\gamma} + \Gamma^{\mu}_{\beta\gamma,\delta} - \Gamma^{\mu}_{\beta\delta,\gamma}$$

Now, since $g_{\alpha\mu;\lambda} = 0$ (see tutorial 5, Qs. 5,6), we can write

$$R_{\alpha\beta\gamma\delta;\lambda} = (g_{\alpha\mu}R^{\mu}{}_{\beta\gamma\delta})_{;\lambda} = g_{\alpha\mu}R^{\mu}{}_{\beta\gamma\delta;\lambda}$$

Now all the Christoffel symbols are zero at P in geodesic coordinates, so from the above two equations (changing covariant back to partial derivatives, which are interchangeable at P in geodesic coordinates)

$$R_{\alpha\beta\gamma\delta;\lambda} = g_{\alpha\mu}(\Gamma^{\mu}_{\beta\gamma,\delta\lambda} - \Gamma^{\mu}_{\beta\delta,\gamma\lambda})$$

Substituting for the Christoffel symbols using

$$\Gamma^{\sigma}_{\alpha\beta} = g^{\sigma\nu}(g_{\nu\alpha,\beta} + g_{\nu\beta,\alpha} - g_{\alpha\beta,\nu})$$

and also using the fact that in geodesic coordinates $g_{\alpha\beta,\gamma} = g^{\alpha\beta}_{\ \ ,\gamma} = 0$ one obtains

$$R_{\alpha\beta\gamma\delta;\lambda} = \frac{1}{2}(g_{\alpha\gamma,\beta\delta\lambda} + g_{\alpha\delta,\beta\gamma\lambda} - g_{\beta\gamma,\alpha\delta\lambda} - g_{\beta\delta,\alpha\gamma\lambda})$$

Writing out the corresponding expressions for the other two terms on the left hand side of equation (9.26) and adding gives the required result, i.e.

$$R_{\alpha\beta\gamma\delta;\lambda} + R_{\alpha\beta\lambda\gamma;\delta} + R_{\alpha\beta\delta\lambda;\gamma} = 0$$

in geodesic coordinates at P. Hence the Bianchi identities must hold in all coordinate systems because of their tensorial nature.

Chapter 10

Einstein's equations

The Riemann Christoffel tensor, $R^{\mu}{}_{\alpha\beta\gamma}$, describes the curvature of spacetime. Einstein's equations relate this curvature to the matter and energy content of the Universe, as described by the energy momentum tensor, $\mathbf{T}$. But in Chapter 8 we introduced $\mathbf{T}$ as a $(2,0)$ tensor. Thus, Einstein's equations involve various *contractions* of the Riemann Christoffel tensor.

10.1 Ricci tensor

We can contract the Riemann Christoffel tensor to form a $(0,2)$ tensor, which we call the **Ricci tensor** defined by

$$R_{\alpha\gamma} = R^{\mu}_{\alpha\mu\gamma} \tag{10.1}$$

i.e. contracting on the *second* of the lower indices. (N.B. some authors choose to define $R_{\alpha\gamma}$ as minus this value). From equation (9.21) it follows that we can also write the components of the Ricci tensor as

$$R_{\alpha\gamma} = g^{\sigma\delta} R_{\sigma\alpha\delta\gamma} \tag{10.2}$$

It is easy to show (see Tutorial 6) that $R_{\alpha\beta} = R_{\beta\alpha}$, i.e. the Ricci tensor is symmetric.

10.2 Curvature Scalar, R, and contravariant Ricci tensor

By further contracting the Ricci tensor with the contravariant components of the metric, one obtains the **curvature scalar**, viz:

$$R = g^{\alpha\beta} R_{\alpha\beta} \tag{10.3}$$

One may also use the metric to raise the indices of the Ricci tensor, and thus express it in contravariant form, viz:

$$R^{\mu\nu} = g^{\mu\alpha} g^{\nu\beta} R_{\alpha\beta} \tag{10.4}$$

$R^{\mu\nu}$ is also symmetric.

10.3 The Einstein tensor

Using the contravariant form of the Ricci tensor, we define the **Einstein tensor**, **G**, viz:

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R \tag{10.5}$$

where R is the curvature scalar. Note that since $R^{\mu\nu}$ is symmetric, so too is $G^{\mu\nu}$.

The Einstein tensor is of crucial physical significance in general relativity, since it can be shown from the Bianchi identities that

$$G^{\mu\nu}_{\ \ ;\nu} = 0 \tag{10.6}$$

10.4 Proof of $G^{\mu\nu}_{\ \ ;\nu} = 0$

Consider the Bianchi identities

$$R_{\alpha\beta\mu\nu;\lambda} + R_{\alpha\beta\lambda\mu;\nu} + R_{\alpha\beta\nu\lambda;\mu} = 0 \tag{10.7}$$

From the symmetry properties of the Riemann Christoffel tensor, equation (9.22), we have

$$R_{\alpha\beta\lambda\mu;\nu} = -R_{\alpha\beta\mu\lambda;\nu} \tag{10.8}$$

so that the Bianchi identities can be re-written as

$$R_{\alpha\beta\mu\nu;\lambda} - R_{\alpha\beta\mu\lambda;\nu} + R_{\alpha\beta\nu\lambda;\mu} = 0 \tag{10.9}$$

Contracting equation (10.9) with $g^{\alpha\mu}$ we obtain

$$g^{\alpha\mu}R_{\alpha\beta\mu\nu;\lambda} - g^{\alpha\mu}R_{\alpha\beta\mu\lambda;\nu} + g^{\alpha\mu}R_{\alpha\beta\nu\lambda;\mu} = 0 \tag{10.10}$$

Using the product rule for covariant differentiation, and the fact that $g^{\alpha\mu}_{\ \ ;\beta} = 0$, equation (10.10) can be written as

$$(g^{\alpha\mu}R_{\alpha\beta\mu\nu})_{;\lambda} - (g^{\alpha\mu}R_{\alpha\beta\mu\lambda})_{;\nu} + (g^{\alpha\mu}R_{\alpha\beta\nu\lambda})_{;\mu} = 0 \tag{10.11}$$

This simplifies to

$$R^{\mu}{}_{\beta\mu\nu;\lambda} - R^{\mu}{}_{\beta\mu\lambda;\nu} + R^{\mu}{}_{\beta\nu\lambda;\mu} = 0 \tag{10.12}$$

i.e.

$$R_{\beta\nu;\lambda} - R_{\beta\lambda;\nu} + R^{\mu}{}_{\beta\nu\lambda;\mu} = 0 \tag{10.13}$$

We can further contract this equation by multiplying by $g^{\beta\nu}$ to obtain

$$g^{\beta\nu}R_{\beta\nu;\lambda} - g^{\beta\nu}R_{\beta\lambda;\nu} + g^{\beta\nu}R^{\mu}{}_{\beta\nu\lambda;\mu} \tag{10.14}$$

Again using the product rule and the fact that $g^{\alpha\mu}_{\ \ ;\beta} = 0$, this further simplifies to (after contracting the first two terms and writing the third in terms of the fully covariant form of the Riemann Christoffel tensor)

$$R_{;\lambda} - R^{\nu}_{\lambda;\nu} + g^{\beta\nu}g^{\alpha\mu}R_{\alpha\beta\nu\lambda;\mu} = 0 \tag{10.15}$$

Using the fact that

$$R_{\alpha\beta\nu\lambda} = -R_{\beta\alpha\nu\lambda} \tag{10.16}$$

the third term on the left hand side of equation (10.15) can therefore be re-written as

$$g^{\beta\nu}g^{\alpha\mu}R_{\alpha\beta\nu\lambda;\mu} = -(g^{\beta\nu}g^{\alpha\mu}R_{\beta\alpha\nu\lambda})_{;\mu} = -(g^{\alpha\mu}R_{\alpha\lambda})_{;\mu} = -R^{\mu}{}_{\lambda;\mu} \equiv -R^{\nu}{}_{\lambda;\nu} \tag{10.17}$$

Thus, from equation (10.15) we have

$$R_{;\lambda} - 2R^{\nu}{}_{\lambda;\nu} = 0 \tag{10.18}$$

Multiplying each term by $-\frac{1}{2}g^{\mu\lambda}$ we obtain

$$g^{\mu\lambda}R^{\nu}{}_{\lambda;\nu} - \frac{1}{2}g^{\mu\lambda}R_{;\lambda} \tag{10.19}$$

Once again using the product rule and the fact that the first covariant derivatives of the metric vanish this yields

$$(g^{\mu\lambda}R^{\nu}{}_{\lambda})_{;\nu} - (\frac{1}{2}g^{\mu\nu}R)_{;\nu} = 0 \tag{10.20}$$

which of course is equivalent to

$$G^{\mu\nu}_{;\nu} = 0 \tag{10.21}$$

10.5 Einstein's equations

We saw in Chapter 8 that conservation of energy and momentum implied that

$$T^{\mu\nu}_{;\nu} = 0 \tag{10.22}$$

We now see that from the Bianchi identities

$$G^{\mu\nu}_{;\nu} = 0 \tag{10.23}$$

Thus we have automatically that

$$T^{\mu\nu}_{;\nu} = G^{\mu\nu}_{;\nu} \tag{10.24}$$

These are Einstein's equations, and Einstein took as their solution

$$G^{\mu\nu} = kT^{\mu\nu} \tag{10.25}$$

i.e. the Einstein tensor is equal to the energy momentum tensor, multiplied by a constant, k.

Thus, the geometry of spacetime (described by the Einstein tensor) and the mass-energy content of spacetime (described by the energy momentum tensor) are fundamentally interlinked.

Note that there are 10 independent equations incorporated in equation (10.25), since $G^{\mu\nu}$ and $T^{\mu\nu}$ are symmetric.

We determine the constant k by requiring that we should recover the laws of Newtonian gravity and dynamics in the limit of a weak gravitational field and non-relativistic motion.

10.6 Non-relativistic weak field limit and Newtonian gravity

If the deviation from flatness in our curved spacetime is small we can find a coordinate system for which

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} \tag{10.26}$$

where $\eta_{\alpha\beta}$ is the Lorentz metric $(-1, 1, 1, 1)$ and $h_{\alpha\beta} << 1$ for α, $\beta = 0, 1, 2, 3$. If we also write

$$g^{\alpha\gamma} = \eta^{\alpha\gamma} + h^{\alpha\gamma} \tag{10.27}$$

then

$$\delta^\gamma_\beta = g_{\alpha\beta} g^{\alpha\gamma} = (\eta_{\alpha\beta} + h_{\alpha\beta})(\eta^{\alpha\gamma} + h^{\alpha\gamma}) \tag{10.28}$$

from which it follows that, to first order in $h_{\alpha\beta}$

$$h^{\nu\gamma} = -\eta^{\alpha\gamma}\eta^{\nu\beta}h_{\alpha\beta} \tag{10.29}$$

Consider now the motion of a 'free-falling' test particle. The wordline of the particle is a geodesic and so satisfies

$$\frac{d^2x^\mu}{d\tau^2} = -\Gamma^\mu_{\sigma\delta}v^\sigma v^\delta \tag{10.30}$$

For non-relativistic particles the proper time τ may be replaced by t. Furthermore $v^i << 1$ for $i = 1,\, 2,\, 3$ and $v^0 = 1$. Thus the space components of equation (10.30) reduce to

$$\frac{d^2x^i}{dt^2} = -\Gamma^i_{00} \tag{10.31}$$

Evaluating Γ^i_{00} using equation (6.46) we have

$$\begin{aligned} \Gamma^i_{00} &= \frac{1}{2}g^{i\sigma}(g_{0\sigma,0} + g_{\sigma i,0} - g_{00,\sigma}) \\ &= \frac{1}{2}(\eta^{i\sigma} + h^{i\sigma})(h_{0\sigma,0} + h_{\sigma i,0} - h_{00,\sigma}) \end{aligned} \tag{10.32}$$

Keeping only first order terms in $h_{\alpha\beta}$, and noting equation (10.29) we have

$$\Gamma^i_{00} = \frac{1}{2}(h_{0i,0} + h_{ii,0} - h_{00,i}) \tag{10.33}$$

Assuming time derivatives are small compared with space derivatives this implies[1]

$$\Gamma^i_{00} = -\frac{1}{2}h_{00,i} \tag{10.34}$$

Thus the equation of motion of the particle is simply

$$\frac{d^2x^i}{dt^2} = \frac{1}{2}h_{00,i} \tag{10.35}$$

[1] If time derivatives are *not* small then one generates gravitational waves – see G&R II for more details

In Newtonian theory the equation of motion is

$$\frac{d^2x^i}{dt^2} = -\frac{\partial\phi}{\partial x^i} \tag{10.36}$$

where ϕ is the Newtonian gravitational potential. This agrees with the weak field GR limit of equation (10.36) if we put

$$h_{00,i} = -2\frac{\partial\phi}{\partial x^i} \tag{10.37}$$

Equation (10.37) simply shows how the metric (in GR) can be chosen to ensure that geodesics (worldlines of free test particles) correspond to the classical non-relativistic trajectories of particles accelerating in a gravitational field. But GR also tells us, through Einstein's equations, how matter and energy curve spacetime. The classical analogue to this is how mass density produces the gravitational field. This is most easily expressed through Poisson's equation, viz

$$\nabla^2\phi = 4\pi G\rho \tag{10.38}$$

where G is the Newtonian gravitational constant (not to be confused with the Einstein tensor!). General relativity must agree with equation (10.38) in the non-relativistic weak field limit.

10.6.1 Geodesic deviation in Newtonian theory

In a gravitational field, ϕ, the Newtonian acceleration of a test particle will be given by equation (10.36), i.e.

$$\frac{d^2x^i}{dt^2} = -\frac{\partial\phi}{\partial x^i}$$

Suppose a neighbouring particle, with coordinates $x^i+\xi^i$, accelerates in the same gravitational field. Evidently

$$\frac{d^2x^i+\xi^i}{dt^2} = -\frac{\partial\phi(x+\xi)}{\partial x^i} = -(\frac{\partial\phi}{\partial x^i} + \frac{\partial^2\phi}{\partial x^i\partial x^j}\xi^j) \tag{10.39}$$

(we have used a first order Taylor's expansion in the last step).

The relative acceleration of the neighbouring particle to the original test particle is found by subtracting equation (10.39) from equation (10.38) to give

$$\frac{d^2\xi^i}{dt^2} = -\frac{\partial^2\phi}{\partial x^i \partial x^j}\xi^j \tag{10.40}$$

10.6.2 Geodesic deviation in GR

Consider the same two neighbouring test particles, but now from the point of view of GR. In GR the geodesic deviation of the two particles is given by (c.f. equation 9.19)

$$\frac{D^2\xi^\mu}{D\tau^2} = -R^\mu_{\alpha\beta\gamma}v^\alpha\xi^\beta v^\gamma \tag{10.41}$$

Take a geodesic coordinate system in which the first test particle is instantaneously at rest. Thus $v^1 = v^2 = v^3 = 0$ and $v^0 = 1$. In this coordinate system the covariant derivative operator $D^2/D\tau^2$ can be approximated simply by d^2/dt^2. Thus, taking the spatial components of equation (10.41) yields

$$\frac{d^2\xi^i}{dt^2} = -R^i_{0\beta 0}v^0\xi^\beta v^0 = -R^i_{0\beta 0}\xi^\beta \tag{10.42}$$

If we take the initial separation such that $\xi^0 = 0$, (i.e. the test particles are released simultaneously in the chosen coordinate frame), equation (10.42) becomes

$$\frac{d^2\xi^i}{dt^2} = -R^i{}_{0j0}\xi^j \tag{10.43}$$

Comparing this with the Newtonian expression (10.40) we see immediately that

$$R^i{}_{0j0}\xi^j = \frac{\partial^2\phi}{\partial x^i \partial x^j}\xi^j \tag{10.44}$$

but since the components ξ^j were arbitrary we must have

$$R^i{}_{0j0} = \frac{\partial^2\phi}{\partial x^i \partial x^j} \tag{10.45}$$

10.6.3 Relating the Einstein and energy momentum tensors

It follows from equation (10.45) that

$$R_{00} = R^{\alpha}{}_{0\alpha 0} = R^{i}{}_{0i0} + R^{0}{}_{000} = R^{i}{}_{0i0} = \frac{\partial^2 \phi}{\partial x^i \partial x^i} \tag{10.46}$$

(Note that $R^{0}{}_{000} = 0$ from its skew-symmetry; see Tutorial 6, Q.1)

R_{00} can also be related to the mass-energy density through Einstein's equations, this time written in *covariant* form, i.e.

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = k T_{\mu\nu} \tag{10.47}$$

Taking the '00' component we have

$$R_{00} = k T_{00} + \frac{1}{2} g_{00} R \tag{10.48}$$

However, in our geodesic coordinates $g_{00} = -1$ and $g_{ii} = 1$. Thus

$$R_{00} = k T_{00} - \frac{1}{2} R \tag{10.49}$$

Suppose we contract the energy momentum tensor, defining the scalar, T, as

$$T = g^{\mu\nu} T_{\mu\nu} \tag{10.50}$$

In geodesic coordinates this is simply

$$T = -T_{00} + T_{11} + T_{22} + T_{33} \tag{10.51}$$

but for non-relativistic matter, the energy density is much greater than the momentum density (i.e. for a non-relativistic fluid, the pressure $P << \rho$). Hence, $T \simeq -T_{00} = -\rho$. Contracting Einstein's equations (10.47) with $g^{\mu\nu}$ we, therefore, obtain

$$g^{\mu\nu} R_{\mu\nu} - \frac{1}{2} g^{\mu\nu} g_{\mu\nu} R = k g^{\mu\nu} T_{\mu\nu} = kT \tag{10.52}$$

i.e.

$$R - \frac{1}{2}\delta^{\mu}{}_{\mu}R = R - 2R = -R = -kT_{00} = -k\rho \tag{10.53}$$

Substituting into equation (10.49),

$$R_{00} = kT_{00} - \frac{1}{2}R = k\rho - \frac{k}{2}\rho = \frac{k}{2}\rho \tag{10.54}$$

Finally, combining equations (10.53) and (10.54) we see that in the non-relativistic limit

$$R_{00} = \nabla^2\phi = \frac{k}{2}\rho \tag{10.55}$$

This is precisely the Newtonian equation of gravity, expressed by Poisson's equation (10.38), provided we choose the as yet arbitrary constant k to be given in terms of the Newtonian gravitation constant, G, by

$$k = 8\pi G \tag{10.56}$$

Thus the final form of Einstein's field equations is

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R = 8\pi G T^{\mu\nu} \tag{10.57}$$

or in units where $G = 1$, (i.e. mass in metres)

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R = 8\pi T^{\mu\nu} \tag{10.58}$$