Course on General Relativity

G F R Ellis
Applied Mathematics, University of Cape Town

November 16, 2009

1 Introduction

The main thrust of Einstein's theory of gravitation ("the General Theory of Relativity") is that gravity can best be understood as a theory of space-time geometry. This is a daring proposal; and it is astonishing that it works. In order to explain this, the major part of this text is concerned with space-time geometry: how to use coordinates to describe a space-time, how to represent physical quantities as space-time tensors, how to write the equations of physics as tensor equations, and which particular equations are the equations describing gravity.

This will all be developed systematically in what follows. Thus this is a text on curved space-times, describing the geometry and physics of Einstein's theory of gravity. It builds on an understanding of flat space-time and special relativity such as that given in *Flat and Curved Space-Times*, G F R Ellis and R M Williams, Oxford University Press (which will be often referred to in the sequel). The aim is to (a) show clearly the foundations on which General Relativity is built, and (b) give detailed exposition of the calculational methods that are used in examining the implications of that theory.

Technically the text is conservative: it concentrates on the index notation that is almost always the best way to carry out calculations of any complexity. However an unusual feature is that that right from the start, an emphasis is placed on the use of general bases rather than just coordinate bases. This means that the use of orthonormal and null tetrads is easily included as a special case, rather than being tacked on as an awkward after-thought. We believe this is the best approach to take, given the usefulness of tetrad methods in many calculations (and in particular, as a basis for spinors and as a way of relating physics in a curved space-time to physics in flat space-time).

This approach makes clear that this text does not aim to present the material in a formal or mathematically rigorous way, rather aiming at building up a solid working knowledge of the subject and good geometrical understanding, demonstrating clearly the interplay between geometry and physics thatis the centre of general relativity. It will serve as a solid foundation for advanced texts on relativity that present a more rigorous view (for example, *The Large Scale Structure of Space-Time*, S W Hawking and G F R Ellis, Cambridge University Press, or Clarke and de Felice). It is at about the same level as the excellent books by Schutz and d'Inverno.

Given this aim, much of the text considers foundational and calculational issues. In order to provide a useful basis for the two main astrophysical applications of the subject (stellar structure and gravitational collapse on the one hand, and cosmology on the other), the following policy has been adopted. The text develops in depth the theory of spherically symmetric static space-times, presenting many details of the calculations required to understand them, so that in particular the Schwarzschild solution and the equations for static stars are treated in some depth. This provides a good introduction to the theory of black holes, and a sound basis for proceeding to texts such as Shapiro and Teukolsky on astrophysical applications of general relativity.

To keep the text a reasonable length, we have not developed the equations of cosmology to the same extent in the main text. Rather, we have consistently used the Robertson-Walker metric (the standard basis of modern cosmology) as the basis of the examples set for the student, so that anyone who works consistently through these examples will at the end have mastered the geometric basis and dynamic equations of modern cosmology. This then provides a sound basis for proceeding to a study of the physics of cosmology.

2 Spacetime

General Relativity is the classical theory of *gravity*. This is a generalisation of special relativity, based on the idea of a space-time as a unifying feature. In general relativity we treat gravity *geometrically*, as due to spacetime curvature; this is effectively necessitated by the equivalence principle (sect. 5.3). Thus general relativity addresses itself primarily to the understanding of the *geometry of spacetime*, which is a Riemannian Geometry (discussed in detail below). The technical tool needed for this purpose is *tensor calculus*, which generalises vector calculus to quantities with multiple indices. It is important to understand that neither the space-time geometry nor its global topology are given *a priori*; we have to discover them by careful analysis.

The aim of the following is to give geometrical and calculational insight rather than to give a formal exposition. I recommend *Introducing Einstein's General Relativity*, R. D'Inverno (Oxford University Press) and *General Relativity*, H. Stephani (Cambridge University Press) for further reading at something like the same level. A more formal exposition is in *Relativity on Curved Manifolds*, F de Felice and C J S Clarke (Cambridge University Press).

2.1 Events and coordinates

Spacetime, which J.A. Wheeler calls the "Arena of physics", is the set of 'events' (i.e. places and times) where objects can move and interact. Technically, it is a 4-dimensional *manifold* M, that is, a space that is locally like a Cartesian 4-dimensional space. It gives the history of a set of objects and observers (see *Flat and Curved Space-times*, G Ellis and R Williams (Oxford University Press), for a detailed introductory discussion). For example, using ordinary coordinates, the movement of a planet around the sun can be represented as a helical curve (with radius equal to the earth's orbit) around a tube (with radius equal to the radius of the sun); at each specific time, one sees from this where the earth was at that time.

In describing spacetime (which is four dimensional) we use *coordinate systems* with 4 coordinates, such as (t, x, y, z) and (t, r, θ, ϕ), to denote the location of events (because spacetime is 4-dimensional, fewer coordinates would be inadequate to distinguish all events; more are superfluous). The general notation is

$$\{x^i\} = (x^0, x^1, x^2, x^3).$$

We can choose the coordinates arbitrarily, according to what is convenient:

Principle: we can use any coordinates we like.

It is important to note that the meaning of coordinates may not be obvious; for example just because a coordinate is labelled 't' does not necessarily mean that it is in fact a time coordinate. Furthermore in general we cannot cover the whole of space-time by one coordinate system; we will need several overlapping coordinate systems to describe the whole.

In what follows we will often come across quantities with indices, such as the components of a vector X^a. Here the index a ranges from 0 to 3, so X^a represents the set of objects (X^0, X^1, X^2, X^3). Similarly the components T_{ab} of a tensor with 2 indices represent the set of quantities where each of a, b range over 0 to 3: T_{00}, T_{01}, and so on. These can be conveniently represented as a 4×4 matrix. In principle we can have tensors with any numbers of indices; in practice we will deal with up to 5: e.g. T_{abcde}.

Three points to note about what follows.

1] In general, as it costs little, we will deal with an *n- dimensional space* with coordinates $\{x^i\}$, $i = 1, .., n$ (or $i = 0, .., n-1$). It is always helpful to think of simple examples (2- and 3-dimensional spaces are easier to visualise than higher dimensional ones!)

2] We will always use the Einstein Summation Convention, namely that *whenever there is a repeated index, it is summed over the range of the index.* Thus for example,

$$X^a Y_a = \sum_{a=1}^{a=n} X^a Y_a = X^1 Y_1 + X^2 Y_2 + X^3 Y_3$$

in a three dimensional space, and also

$$\begin{aligned} g_{ab} X^a Y^b &= g_{11} X^1 Y^1 + g_{12} X^1 Y^2 + g_{13} X^1 Y^3 + \\ &+ g_{21} X^2 Y^1 + g_{22} X^2 Y^2 + g_{23} X^2 Y^3 + \\ &+ g_{31} X^3 Y^1 + g_{32} X^3 Y^2 + g_{33} X^3 Y^3 \end{aligned}$$

It is important to note that a summed ('dummy') index can be relabelled at will (*provided* the index chosen does not already occur elsewhere in the equation); thus for example

$$X^a Y_a = X^b Y_b = X^s Y_s$$

(these being exactly the same expressions, when written out in full).

3] The Kronecker delta δ^i_j is defined by

$$\delta^i_j = \begin{array}{ll} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{array}$$

When summed with any quantity, it leaves it essentially unaffected, simply substituting a new index for an old one: for example,

$$X^j \delta^i_j = X^0 \delta^i_0 + X^1 \delta^i_1 + X^2 \delta^i_2 + X^3 \delta^i_3 = X^i, \quad w_i \delta^i_j = w_j$$

(consider each value in turn of i to see the first result is true; the second follows similarly). As a particular case,

$$\delta^i_k \delta^k_j = \delta^i_j.$$

The basic elements of spacetime are events, curves, and surfaces.

2.1.1 Events

Events (i.e. space-time points) are labelled by the chosen coordinates, which can be considered as a (local) 1-1 map Φ from the manifold M to the coordinate space R^4. This map assigns to each event P the corresponding coordinates $x^i(P)$ (the Cartesian coordinates of the image $\Phi(P)$ of P). Because the map is 1:1, this enables us to locate the event P and refer to it uniquely. It also determines what are open sets and closed sets in M (they are the images of open and closed sets in R^4, as obviously described by the coordinates) thus fixing the local space-time topological and continuity properties (i.e. determining which events are 'near' to other events in the sense of continuity), and similarly determines the differentiability properties of the space-time (expressed by differentiability in terms of local coordinates).

Spacetime is just the set of all physically connected events. We may need several coordinate charts to cover all of space-time (only in the simplest cases can it all be covered by one coordinate system in a regular way).

2.1.2 Curves

A *curve* in space-time is a one-dimensional set of events, denoted by $x^i(v)$ where v is the curve parameter. Thus a curve is a (local) map of R^1 into M. Examples are the history of a point particle in space-time, called its *world-line*, or of a light-ray. Each curve is associated with a tangent vector which gives us its direction in spacetime at each point on the curve. The components of this vector are

$$X^i = \frac{dx^i}{dv} \tag{1}$$

In the case of a particle motion, it gives its 4-velocity.

In general we can have a *congruence of curves*, that is a family of curves that fills up spacetime, with a tangent vector defined at each point (e.g. the flow-lines of a fluid filling space-time); then the tangent vector gives the fluid 4-velocity at each point. These in turn determine the family of curves, through equation (1).

2.1.3 Surfaces

A single *surface* is a 3-dimensional set of events. It is described by a non-constant function $f(x^i)$ through an equation $f(x^i) = c$, c constant, (a 'level surface' of the function). A *family of surfaces* is obtained as we let c range over a set of values in an interval I of R. Thus any family of (non-intersecting) surfaces is represented by a function $f(x^i)$, which is a map from M to R^1. For example $T(x)$, $p(x)$ represent surfaces of constant temperature and pressure in a universe model. The coordinates x^i are just a set of 4 functions, and so represent 4 independent families of surfaces in space-time.

2.1.4 Curve-Surface relations

Now we want to find the relation between curves and surfaces.

Consider a curve $x^i = x^i(v)$ and a family of surfaces $f = f(x^i)$. Combining these relations, the variation of the function along the curve is given by $f(v) = f(x^i(v))$. Its rate of change with respect to the curve parameter is

$$\frac{df}{dv} = \frac{\partial f}{\partial x^i}\frac{\partial x^i}{\partial v} = X^i \frac{\partial f}{\partial x^i} =: X(f) \tag{2}$$

where (as implied by the definition at the end of this equation)

$$\mathbf{X} = X^i(x^j)\frac{\partial}{\partial x^i} \tag{3}$$

is the directional derivative operator in the direction $\mathbf{X}$ (remember the summation convention is being used!). In geometric notation the vector $\mathbf{X}$ is an arrow, but analytically it is a differential operator (a directional derivative, uniquely characterising a direction at each point) and this equation expresses it in terms of the basis of directional derivatives $\partial/\partial x^i$ (also operators) with components $X^i(x^j)$. It approximates the value of a differentiable function f at the point $x^i(v_0 + \delta v)$ because of the Taylor expansion

$$f(x^i(v_0 + \delta v)) = f(x^i(v_0)) + \delta v X(f)|_{v_0} + O(\delta v^2). \tag{4}$$

If the curve lies in the surface $f = const$, then $f(x(v)) = constant \Leftrightarrow df/dv = 0 \Leftrightarrow X(f) = 0$. If $X(f) \neq 0$, then $\mathbf{X}$ does not lie in the surfaces of constant f, and the larger this value is, the more rapidly the function is changing (for example we can think of the vector as transvecting the set of surfaces $f = c + n\,\delta c$ where c, δc are constants and n an integer; the larger $X(f)$, the more of these surfaces are crossed by the curve with tangent $\mathbf{X}$ for a given δv, by equation (4)). Equation (2) shows that a vector field $\mathbf{X}$ maps the function f to the function $X(f)$.

The variation of the function, geometrically characterised by this set of 'level surfaces', is analytically characterised by its *differential* $\mathbf{d}f$, with components

$$df_i = \frac{\partial f}{\partial x^i} \Leftrightarrow X(f) = df_i X^i =: df(\mathbf{X}) \tag{5}$$

for any vector field $\mathbf{X}$, where the equivalence follows from (2). The last equality defines the differential as an operator acting on vector fields to give a function ($\mathbf{d}f : \mathbf{X} \to df(\mathbf{X}) \equiv X(f)$, the value of $\mathbf{df}$ when acting on $\mathbf{X}$). Geometrically it can be thought of as a pair of nearby level surfaces (cf. the previous paragraph).

2.1.5 Coordinates, Curves and Surfaces

Any set of coordinates has associated with it a family of coordinate surfaces, curves, and vectors. To understand the coordinates we should understand these curves and surfaces. They result from using the 1-1 coordinate map $\Phi : M \to R^n$ to map the coordinate surfaces and curves in R^n into the space M.

The coordinate surfaces are defined by $\{x^i = constant\}$ for $i = 1$ to n. The coordinate curves lie in the intersection of these surfaces, for example the curve $x^i(v)$ given by $x^1 = v$, $x^2 = const$, $x^3 = const$ in R^3 lies in the intersection of the surfaces $x^2 = const$, $x^3 = const$. By Equation (1), its tangent vector has components $X^i = (1, 0, 0) = \delta^i_1$, the '1' stating that x^1 is the curve parameter and the zeros that x^2 and x^3 do not vary along the curve. From (3) its tangent vector is $\mathbf{X} = \delta^i_1 \frac{\partial}{\partial x^i} = \frac{\partial}{\partial x^1}$. Similarly the curve $\{x^i \text{ only varies}\}$ implies $x^j (j \neq i)$ is constant, and this is the curve with tangent vector $\mathbf{X} = \frac{\partial}{\partial x^i} \Leftrightarrow X^j = \delta^j_i$ ('i' fixed, 'j' runs over 1 to n). Thus associated with the coordinate system there is a set of n vector fields e_k given by $e_k = \frac{\partial}{\partial x^k}$, and with components $e^i_k = \delta^i_k$ ('k' labels the vector field and 'i' labels the component). This is the coordinate basis associated with the coordinates $\{x^i\}$. Equation (3) expresses the vector field $\mathbf{X}$ in terms of this basis.

The component relations (1) are special cases of a general result: apply (2) to the coordinate x^j (that is, set $f = x^j$ in (2)). Then we find

$$\frac{dx^j}{dv} = X(x^j) = X^i \frac{\partial x^j}{\partial x^i} = X^i \delta^j_i = X^j \tag{6}$$

That is, the component X^j of a vector field $\mathbf{X}$ is just the rate of change of the coordinate x^j along the curves with tangent vector $\mathbf{X}$. If in particular we consider the coordinate curves, we obtain the results of the previous paragraph.

The differentials naturally associated with the coordinates are defined similarly. Apply (5) to $f = x^j$:

$$d(x^j)_i = \frac{\partial x^j}{\partial x^i} = \delta^j_i \Leftrightarrow X(x^j) = dx^j(\mathbf{X}) = X^j$$

that is, the first equation states that the 'i'-component of the differential $d(x^j)$ is zero unless $i = j$, when it is one; the right hand equation says that the differential dx^j is just that mapping of vector fields to functions which gives the component X^j of an arbitrary vector field $\mathbf{X}$. These differentials $d(x^j)$ form a coordinate basis of differentials. They are the basis *dual* to the vector basis $\frac{\partial}{\partial x^i}$ because if we apply (5) to this coordinate basis of vectors we find $\frac{\partial}{\partial x^j}(x^i) =: dx^j(\frac{\partial}{\partial x^i}) = \delta^j_i$. Geometrically this corresponds to the fact that the coordinate curves and surfaces exactly fit together.

To fully grasp these relations, one should look in detail at the coordinate curves and surfaces in the following cases: (i) The Euclidean plane with Cartesian coordinates (x, y); (ii) the Euclidean plane with polar coordinates (r, θ); (iii) Euclidean 3-space with cylidrical coordinates (ρ, θ, z); (iv) Euclidean 3-space with spherical polar coordinates (r, θ, ϕ).

Given an arbitrary function $f(x^j)$, on using equation (5) we can write

$$\mathbf{d}f = df_i \mathbf{d}x^i \Leftrightarrow df(\mathbf{X}) = df_i dx^i(\mathbf{X}) = \frac{\partial f}{\partial x^i} X^i \tag{7}$$

where the first equality expresses the differential $\mathbf{d}f$ (an operator on vectors) in terms of the basis differentials $\mathbf{d}x^i$ (also operators on vectors), with components $df_i(x^k)$ (a set of functions). The right hand equality explicitly shows what this map is when applied to the arbitrary vector field $\mathbf{X}$. If we choose $\mathbf{X} = \frac{\partial}{\partial x^k}$ for some coordinate x^k, we find the kth. component of the differential: $df_k = df(\partial/\partial x^k)$.

2.2 Distances, Angles and Times

In 3 dimensional Euclidean space, using Cartesian coordinates, the distance corresponding to a small coordinate displacement $dx^i = (dx, dy, dz)$ is

$$ds^2 = dx^2 + dy^2 + dz^2$$

(Pythagoras' theorem applied to this displacement), so distance along a curve $x^i(v)$ is defined as

$$S = \int_P^Q \sqrt{ds^2} = \int_P^Q \sqrt{dx^2 + dy^2 + dz^2} = \int_P^Q \sqrt{\left(\frac{dx}{dv}\right)^2 + \left(\frac{dy}{dv}\right)^2 + \left(\frac{dz}{dv}\right)^2} dv$$

In spherical coordinates, ds^2 becomes

$$ds^2 = dr^2 + r^2\{d\theta^2 + \sin^2\theta d\phi^2\}$$

The distances along the coordinate lines can immediately be read off from these metric forms; for example the latter expression shows that distances along the coordinate lines $\{r$ only varies$\}$, $\{\theta$ only varies$\}$, and $\{\phi$ only varies$\}$ are given respectively by $ds^2 = dr^2$, $ds^2 = r^2 d\theta^2$ and $ds^2 = r^2\sin^2\theta d\phi^2$ (for example, obtain the first one by setting $d\theta = 0 = d\phi$). This agrees with a simple geometric analysis of distances along these coordinate lines. In particular, it shows that the coordinates do not in general give distances along the coordinate curves; for example the coordinate θ only gives distance along the curve $\{\theta$ *only varies*$\}$ when scaled by r.

In general coordinates, this has the form

$$ds^2 = g_{ij}(x^k)dx^i dx^j, \quad g_{ij}(x^k) = g_{ji}(x^k) \tag{8}$$

where $g_{ij}(x^k)$ are the components of the *metric tensor*. Of course the summation convention has been assumed here; when written out in full in 3 dimensions this relation is

$$\begin{aligned} g_{ab}dx^a dx^b &= g_{11}dx^1dx^1 + 2g_{12}dx^1dx^2 + 2g_{13}dx^1dx^3 + \\ &+ g_{22}dx^2dx^2 + 2g_{23}dx^2dx^3 + g_{33}dx^3dx^3 + \end{aligned}$$

where the symmetry of the metric form has been used.

They take the forms $g_{ij} = diag(1,1,1)$ and $diag(1, r^2, r^2\sin^2\theta)$ for Cartesian and spherical coordinates respectively (equation (8) then reproduces the distance relations for those coordinate systems). In terms of these components, the *scalar product* of two vectors is

$$\mathbf{X}.\mathbf{Y} = g_{ij}X^iY^j \Rightarrow X^2 = g_{ij}X^iX^j = |\mathbf{X}|^2 \tag{9}$$

and so the angle between the vectors is determined by $\cos\theta = \frac{\mathbf{X}.\mathbf{Y}}{|\mathbf{X}||\mathbf{Y}|}$. In particular, they are orthogonal iff $\mathbf{X}.\mathbf{Y} = 0$.

It follows that we can easily calculate also the angles between the coordinate lines from the metric tensor, for *the metric tensor components are just the scalar products of the basis vectors.* As an example, the (x, y, z) lines in Euclidean space are orthogonal, as there are no cross terms like $dxdy$, etc. To see this, note that if $\mathbf{X}$ is the 1-basis vector (lying along the 1-direction), $\mathbf{Y}$ is the 2- basis vector (lying along the 2-direction), and $\mathbf{Z}$ is the 3- basis vector (lying along the 3-direction), then the scalar product $\mathbf{X}.\mathbf{Y} = \delta^i_1 g_{ij}\delta^j_2 = g_{12}$, so $g_{12} = \{$scalar product of $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}\}$; this vanishes because they are orthogonal. Similarly $\mathbf{Z}.\mathbf{X} = g_{ij}Z^iX^j = g_{ij}\delta^i_3\delta^j_1 = g_{31} = 0$ because those lines are orthogonal.

We can also see immediately from the metric form that the coordinate lines in spherical polars are orthogonal to each other.

In general we can regard the metric as giving a map from pairs of vectors to numbers at each point, through taking the scalar product: thus $\mathbf{g} : \mathbf{X}, \mathbf{Y} \to \mathbf{X}.\mathbf{Y} = g(\mathbf{X}, \mathbf{Y})$ where the functional notation is used in the latter expression. In this way it maps pairs of vector fields to functions on space-time.

2.2.1 Curved Spaces

These results are for Euclidean space. We obtain a *general curved (Riemannian) space* by adopting equation (8), taking it as determining distances in the space through

$$S = \int_P^Q \sqrt{ds^2} = \int_P^Q \sqrt{g_{ij}(x^k)dx^i dx^j} = \int_P^Q \sqrt{g_{ij}(x^k)\frac{dx^i}{dv}\frac{dx^j}{dv}}\, dv \qquad (10)$$

and scalar products through (9). Clearly this includes (flat) Euclidean space as a special case. It also includes the (curved) surface of a 2-sphere of radius R, obtained in standard coordinates (θ, ϕ) if $g_{ij} = R^2\, diag(1, \sin^2\theta)$. This enables us to find the physical distances along coordinate curves (the coordinates by themselves do not give distances; they only do so when we kow the metric). However we can use the metric to define a coordinate - say x - to be distance along its coordinate curves: in that case $g_{xx} = 1$. This is done for example in Cartesian coordinates, where we define such coordinates along all the orthogonal coordinate curves (this is not possible in a general space).

In general, following the same argument as above, we find that the metric tensor components are the scalar products of the coordinate basis vectors:

$$g_{ij} = (\frac{\partial}{\partial x^i}).(\frac{\partial}{\partial x^j}) = g\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right)$$

This gives us an immediate geometric interpretation of these metric components in any space-time.

When using any particular coordinate system, we will in general find there are coordinate singularities. For example, there is one at $r = 0$ for spherical coordinates; this can be seen by the fact that the metric components are singular there. When the metric is singular, this need not be a real (physical) singularity; it may just represent a poor choice of coordinates at that point. Distinguishing the one from the other is important, and is one of the aims of our analysis. However the occurrence of metric singularities associated with coordinate singularities is why coordinates are generally local, and cannot cover the whole space. Therefore, we usually need several coordinate patches.

2.2.2 Flat Spacetime

First consider the flat space-time of special relativity, in Minkowski coordinates $\{x^i\} = (t, x/c, y/c, z/c)$ where c is the speed of light (often put equal to 1 by appropriate choice of units). The metric tensor corresponding to (8) can be written

$$\eta_{ij} = diag(-1, 1, 1, 1) \tag{11}$$

so that

$$ds^2 = -dt^2 + (dx^2 + dy^2 + dz^2)/c^2 \tag{12}$$

while if we use polar coordinates with $c = 1$ we get

$$ds^2 = -dt^2 + dr^2 + r^2\{d\theta^2 + \sin^2\theta d\phi^2\}$$

This *space-time interval* conveys information not only about distances but also about time and the motion of light, in the following way. Consider a particle moving from x^i to $x^i + dx^i$ in terms of these coordinates. From (12), we find

$$ds^2 = dt^2(-1 + \frac{dr^2}{c^2dt^2}) = -dt^2(1 - \frac{v^2}{c^2}) \tag{13}$$

where $dr^2 = dx^2 + dy^2 + dz^2$ and so v is local speed of motion. Consequently,

(1) $ds^2 = 0 \Leftrightarrow v^2 = c^2$: this defines motion at the speed of light (in space-time, it determines the local light-cone at each point).

(2) $ds^2 < 0 \Leftrightarrow v^2 < c^2$: this represents motion at less than the speed of light, and so a possible space-time path of a material particle. In this case, we define the proper time τ along the particle world line by

$$\tau = \int_P^Q \sqrt{|ds^2|} = \int_P^Q \sqrt{-g_{ij}(x^i)dx^idx^j} \Leftrightarrow d\tau = \sqrt{-ds^2} \tag{14}$$

From (12,3) we then find

$$\frac{d\tau}{dt} = \sqrt{1 - \frac{v^2}{c^2}} \Leftrightarrow \frac{dt}{d\tau} = \gamma(v), \quad \gamma(v) := \frac{1}{\sqrt{1 - v^2/c^2}} \tag{15}$$

in accord with the usual Special Relativity results.

(3) $ds^2 > 0 \Leftrightarrow v^2 > c^2$: this represents motion at greater than the speed of light, and so is impossible for a material particle. In this case, we define the proper distance D (measured locally by radar) along the path by

$$D = \int_P^Q \sqrt{|ds^2|} \Leftrightarrow dD = \sqrt{ds^2}. \tag{16}$$

(4) The metric also determines simultaneity for an observer: this corresponds to space-time orthogonality. Specifically, the tangent vector T^i to the timelike curves $\{t$ only varies$\}$ in metric (11) has components $T^i = \delta^i_0$; while any curve tangent to the surfaces of simultaneity $\{t = const\}$ for an observer moving on these world-lines, has tangent vector $Y^i = (0, Y^\nu) \Leftrightarrow Y(t) = 0$. Thus the metric (11) shows $\mathbf{T}.\mathbf{Y} = T^i g_{ij} Y^j = 0$. This is an invariant relation (cf. the following section); consequently it always determines if the displacement Y^i is instantaneous for an observer with 4-velocity T^i.

(5) As before, the metric defines magnitudes for any vectors. However there is an important difference from Euclidean spaces (and indeed from positive definite spaces, that is, spaces where the metric components are all positive when diagonalised). This is that in Euclidean geometry $\mathbf{X}.\mathbf{X} = 0 \Leftrightarrow X^a = 0$, which is not true in spacetime; rather $\mathbf{X}.\mathbf{X} = 0$ implies *either* $X^i = 0$ *or* it is a null vector (e..g. the tangent to a light ray).

2.2.3 Curved Spacetimes

The idea of curved space-times is that we take over these relations for general metrics $g_{ij}(x^k)$ that can be reduced to form (11) at any point by appropriate choice of coordinates. Thus we adopt a metric as in equation (8) and the interpretations (1)-(5) of the previous section; in particular (14) determines how ideal clocks will measure time in space-time.

Example: A specific useful example is the Robertson-Walker metric

$$ds^2 = -dt^2 + S^2(t)\{dr^2 + f^2(r)(d\theta^2 + sin^2\theta d\phi^2)\} \tag{17}$$

where $f(r) = \sin r$, r, or $\sinh r$. This is the standard metric of cosmology, with positively curved, flat, or negatively curved space-sections respectively; where the matter is assumed to move along the world lines with tangent vector $u^i = \delta^i_0$ (called 'fundamental world lines'; an observer moving along them is called a 'fundamental observer').

Implications from the above are: the matter moves with constant coordinates r, θ, ϕ. Coordinate t measures proper time along these world-lines, and the surfaces $\{t = const\}$ are surfaces of instantaneity for them, and are surfaces of homogeneity of all physical quantities. Spatial distance from the origin $r = 0$ to the matter at $r = u$, measured in a surface $\{t = const\}$, is $D = S(t)u$, so the scale function $S(t)$ determines how the spatial distances between fundamental world-lines varies with time. Radial light rays ($\theta = const$, $\phi = const$) move on paths $x^i(v)$ such that $dr/dt = \pm 1/S(t)$.

Examples 1: COORDINATE TRANSFORMATIONS

1.(i) Using an appropiate coordinate transformation show that the metric

$$ds^2 = -2dudv + dy^2 + dz^2$$

describes a Minkowski spacetime.

(ii)The four velocity of a particle is defined through $u^\alpha = \frac{dx^\alpha}{d\tau}$ where $d\tau = \sqrt{-ds^2}$ is the proper time. Compute the four velocity for a particle at rest in usual Minkowski coordinates and transform it into the (u, v, y, z) coordinates.

2. The Robertson Walker metrics defined by

$$ds^2 = -dt^2 + S^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + sin^2\theta d\varphi^2)\right] \qquad ,$$

where $k = 0, 1, -1$, are the standard models of cosmology.

Show that these metrics can be rewritten as

$$ds^2 = -dt^2 + S^2(t)[d\chi^2 + f^2(\chi)(d\theta^2 + sin^2\theta\varphi^2)]$$

Find the function $f(\chi)$ for each of the three cases $k = 0, 1, -1$.

3. A metric is *conformally flat* if it differs from that of Minkowski spacetime only by a factor, that is, if $ds^2 = \Omega^2(x^\alpha)ds^2_{Mink}$. The function $\Omega(x^\alpha)$ is called the *conformal factor*.

(i) Show that the "flat" ($k = 0$) Robertson Walker metric is conformally flat.

(ii)What do you think of the other Robertson Walker metrics? Try with the coordinate transformation

$$\rho = \frac{2sin\chi}{cos\chi + cos\eta} \qquad \tau = \frac{2sin\eta}{cos\chi + cos\eta} \qquad \eta = \int dt S^{-1}(t)$$

in the closed ($k = 1$) Robertson Walker metric.

---

3 Tensors

In the description of spacetime we should be able to use an arbitrary coordinate system, that is, our equations should be valid for any coordinates.

3.1 Change of coordinates

If we change from coordinates $\{x^i\}$ to new coordinates $\{x^{i'}\}$ (note that we put the prime on the index 'i'), this corresponds to (1) choosing a new map Φ from M to R^4 to define the coordinates, or equivalently (2) choosing new coordinate surfaces and lines in M. Then

$$x^{i'} = x^{i'}(x^i) \tag{18}$$

is the coordinate expression of the forward transformation, giving the new coordinates of each point in terms of the old. This uniquely implies the inverse (backward) transformation

$$x^i = x^i(x^{i'}) \tag{19}$$

giving the old coordinates of each point in terms of the new. These must give the identity when substituted into each other; e.g. putting (18) into (19) must give $x^i = x^i(x^{i'}) = x^i(x^{i'}(x^i))$ as an identity.

Example: we may wish to describe the plane R^2 in terms of Cartesian coordinate $\{x^i\} = (x, y)$ or spherical coordinate $\{x^{i'} = (r, \theta)\}$.

Then the forward transformations (18) are $x^{1'} = r = (x^2 + y^2)^{1/2}$ and $x^{2'} = \theta = \arctan(\frac{y}{x})$, while the inverse transformations (19) are $x^1 = x = r\cos\theta$ and $x^2 = y = r\sin\theta$. On substituting the first into the second, we find e.g. that $x = (x^2 + y^2)^{1/2}\cos(\arctan(y/x))$ must be an identity.

3.1.1 Functions

By transforming coordinates, we are expressing the same space in terms of different coordinate system; this will affect the representation of all objects in M. For example, take a surface denoted by the function $f(x^i)$. Under the transformation (18),(19), $f(x^i) \Rightarrow f(x^{i'}) = f(x^{i'}(x^i))$, a new functional representation of f in terms of coordinates. Mathematically, we strictly should represent the functions by different symbols, because they have different functional forms; however we will follow the physical convention of using the same letter for the function (e.g. 'ρ' for density), even though the functional form changes. Hence the convention is that the relation

$$\rho(x^i) = \rho(x^{i'}) = \rho(x^{i'}(x^i))$$

is understood when we change coordinates (this asserts equality of the density at the same physical points, even though their coordinate representation has changed). Mathematically they are different functions but physically the same.

To show this,take the Euclidean plane example above: if the density is given by $\rho(x,y) = (x^2+y^2)^{1/2} + A$ in Cartesian coordinates, then $\rho(r,\theta) = r + A$ in polars. It is clear that $\rho = f(x,y)$ and $\rho = g(r,\theta)$ are different functions mathematically.

3.1.2 Curves and vectors

Let us denote a curve by $x^i(\tau)$; then under transformation (18) it becomes

$$x^{i'}(\tau) = x^{i'}(x^i(\tau))$$

and so the new components of the tangent vector are given by

$$X^{i'} = \frac{d}{d\tau}(x^{i'}) = \frac{d}{d\tau}(x^{i'}(x^i(\tau))) = \frac{\partial x^{i'}}{\partial x^i}\frac{dx^i}{d\tau} = \frac{\partial x^{i'}}{\partial x^i}X^i$$

on using the chain rule. Thus we can write the forward transformation,

$$X^{i'} = A^{i'}{}_i X^i, \tag{20}$$

where

$$A^{i'}{}_i = \frac{\partial x^{i'}}{\partial x^i}. \tag{21}$$

We can get the inverse transformation in one of two ways.

(1) Note that the inverse of the Jacobian matrix (21) is given by

$$A_{i'}{}^i = \frac{\partial x^i}{\partial x^{i'}} \Leftrightarrow A_{i'}{}^i A^{i'}{}_j = \delta^i_j, \quad A^{i'}{}_i A_{j'}{}^i = \delta^{i'}_{j'} \tag{22}$$

so multiplying (20) by this inverse gives

$$X^i = A_{i'}{}^i X^{i'} \tag{23}$$

It is important to realise we will have to deal with two transformation matrices (21), (22), each the inverse of the other. They are distinguished from each other by the positions of the indices (see the note after equation (35) for further explanation of this positioning).

(2) Starting with the inverse transformation $x^i(\tau) = x^i(x^{i'}(\tau))$, carry out the derivation as above to arrive at (23);

(3) Note that in the above derivation, there is nothing special about x^a as opposed to $x^{a'}$; so we can obtain the correct result by simply relabelling all indices in equations (20), (21), simultaneously setting $i \to i'$ and $i' \to i$. Then (20), (21) essentially become (22), (23) (the difference is the left-right placing of the indices on the transformation matrix; we always put the new (primed) index on the left).

Example: Consider the 2-dimensional example given above, where $x^i = x^i(x^{i'})$ takes the form $x = r\cos\theta$ and $y = r\sin\theta$.

The transformation matrix ${A_{i'}}^i$ has the form

$${A_{i'}}^i = \frac{\partial x^i}{\partial x^{i'}} = \begin{pmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{pmatrix}$$

Thus if we consider the tangential vector

$$X = \frac{\partial}{\partial\theta} \Leftrightarrow X^{i'} = \delta_2^{i'} = (0, 1)$$

then its transformation is

$$X^i = X^{i'} {A_{i'}}^i = \delta_2^{i'} {A_{i'}}^i = {A_{2'}}^i = \frac{\partial x^i}{\partial x^{2'}}$$

giving the result

$$X^i = (-r\sin\theta, r\cos\theta)$$

To find the explicit inverse transformation matrix ${A^{i'}}_i$, we can use the easiest of the following three ways:

(1) First explicitly find the inverse set of transformations $x^{i'} = x^{i'}(x^i)$ and then take its Jacobian. In this case, find the Jacobian of the transformation $r = (x^2 + y^2)^{1/2}$ and $\theta = \arctan(y/x)$.

(2) Invert the Jacobian matrix already obtained, to get

$${A^{i'}}_i = \begin{pmatrix} \cos\theta & -\sin\theta/r \\ \sin\theta & \cos\theta/r \end{pmatrix}$$

Note that one can express either Jacobian matrix in terms of either set of coordinates. The first method will give the matrix ${A^{i'}}_i$ as a function of the old coordinates; the second as a function of the new coordinates. Thus to explicitly give the matrix in terms of the old coordinates (usually required) in the second case one will still have to transform the result from the new coordinates to the old ones. This gives the result

$$A^{i'}{}_i = \begin{pmatrix} x/(x^2+y^2)^{1/2} & -y/(x^2+y^2) \\ y/(x^2+y^2)^{1/2} & x/(x^2+y^2) \end{pmatrix}$$

which has to agree with the result of (1).

It is important to notice from this analysis that we have a good transformation if and only if

$$det(A_{i'}{}^i) \neq 0 \Leftrightarrow det(A^{i'}{}_i) \neq 0.$$

This is the condition that the transformations are 1-1 (each has a regular inverse). Geometrically it is the condition that linearly independent basis vectors remain linearly independent. Thus when this is not true at some set of points U, there is some kind of problem with the coordinates there (one of the coordinate systems will have a singularity). For example, this problem will arise at the origin of polar coordinates when we transform between them and Cartesians in R^3.

An alternative way of obtaining the vector transformation laws is to note that we can write the vector $\mathbf{X}$ equally well in terms of the old and the new basis:

$$\mathbf{X} = X^i \frac{\partial}{\partial x^i} = X^{i'} \frac{\partial}{\partial x^{i'}} \tag{24}$$

However by the usual partial differential chain rule,

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

so (24) can be written

$$\mathbf{X} = X^i \frac{\partial}{\partial x^i} = X^{i'} \frac{\partial x^i}{\partial x^{i'}} \frac{\partial}{\partial x^i}.$$

Equating the components in these two ways of expressing the same vector in terms of the same basis, we obtain the transformation law (23).

3.1.3 Gradient of Function

On changing coordinates, the functional form changes: $f(x^i) \to f(x^i(x^{i'}))$, so the components of its differential will change according to

$$df_{i'} = \frac{\partial f}{\partial x^{i'}} = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial f}{\partial x^i} = \frac{\partial x^i}{\partial x^{i'}} df_i$$

that is

$$df_{i'} = A_{i'}{}^i df_i, \quad df_i = A^{i'}{}_i df_{i'} \tag{25}$$

the inverse being obtained by multiplying by the inverse matrix, calculating from scratch, or just relabelling indices (cf. the transition from (20) to (23)).

Note that the opposite matrices appear in the forward and backward transformations here, as in the vector case (equations (20) and (23)), but no confusion will arise if we note that the effect of the transformation matrix is to act as a substitution operator: in every case it cancels the old index (by summation) and replaces it with a new index in the same position. We simply place the indices on the transformation matrix to ensure that the summation is correct: it must always take place between one index that is up and one that is down, with both these indices being primed or both being unprimed.

3.1.4 Invariants

It is important that many physical quantities be expressed as invariants, that is as functions $f(x^i)$ that take the same value in all coordinate systems: $f(x^i) = f(x^{i'})$ (in the sense explained above). Thus we need to find a systematic way of creating invariants out of tensor quantities. To see how this works, consider the interval $ds^2 = g_{ij}dx^i dx^j$ where $g_{ij}(x^k)$ are the metric components. We demand that this be invariant under an arbitrary coordinate transformation. Thus we require that

$$ds^2 = ds'^2 \Leftrightarrow g_{ij}dx^i dx^j = g_{i'j'}dx^{i'}dx^{j'} \tag{26}$$

for all coordinate changes and for all displacements dx^i. Now using the transformation law for the displacements[1] dx^i, this relation becomes

$$g_{ij}dx^i dx^j = g_{i'j'}A^{i'}{}_i dx^i A^{j'}{}_j dx^j$$

$$\Leftrightarrow Z_{ij}dx^i dx^j = 0, \ \ where \ \ Z_{ij} := g_{ij} - g_{i'j'}A^{i'}{}_i A^{j'}{}_j$$

for all displacements dx^j. Choosing this displacement successively in the various coordinate directions and then in combinations of these directions, we find that the symmetric part of Z_{ij} must vanish (for example, choose $dx^i = \delta v \delta^i_1$; then $Z_{11} = 0$). But because g_{ij} is symmetric, Z_{ij} is itself symmetric, so necessarily $Z_{ij} = 0$; that is

$$g_{ij} = g_{i'j'}A^{i'}{}_i A^{j'}{}_j \quad \Leftrightarrow \quad g_{i'j'} = g_{ij}A_{i'}{}^i A_{j'}{}^j \tag{27}$$

is the necessary and sufficient condition for the invariance of the interval expressed in (26). As usual the inverse transformation follows from the forward transformation either by use of inverse matrices, by deriving it from scratch, or by relabelling indices.

We have now derived the transformation law for the metric tensor components. This also follows in a simple way by using the obvious way the displacement components transform. For example, the 2-plane Euclidean metric

[1] These are components of a vector representing a displacement along a curve, not differentials of a function. One could avoid this notational confusion by writing $\delta v\, \eta^i$ instead of dx^i to represent this vector, but the above notation is so standard that we will use it here.

$ds^2 = dx^2 + dy^2$ is easily transformed to the polar form $ds^2 = dr^2 + r^2 d\theta^2$ on assuming invariance of ds^2 (which is why we can write ds^2 in the second expression, rather than ds'^2) and using the relations $dx = \cos\theta\, dr - r \sin\theta\, d\theta$, $dy = \sin\theta\, dr + r \cos\theta\, d\theta$, which follow immediately from the transformation $x = r \cos\theta$, $y = r \sin\theta$.

3.2 Tensors and Tensor transformations

We have discussed the transformation properties of scalar functions, vectors, the metric (see (20), (23), (25), (26), (27)). The general pattern that emerges is now clear. A *tensor* is a quantity which generalizes these patterns to any number of indices upstairs and any number of indices downstairs: it transform on each index in the same way as vector (up) and differential (down) indices respectively (the position of the index correctly shows its transformation properties). As we will see, tensors represent the physical quantities of interest in space-time.

As an example, a tensor $R^{ijk}{}_{gh}$ transforms as follows.

$$R^{i'j'k'}{}_{g'h'} = R^{ijk}{}_{gh} A^{i'}{}_i A^{j'}{}_j A^{k'}{}_k A_{g'}{}^g A_{h'}{}^h \tag{28}$$

As noted above, the effect of the transformation matrix is to act as a substitution operator: in every case it cancels the old index (by summation) and replaces it with a new index in the same position. This is a tensor of type (3,2), as it has three indices upstairs and two down; a general tensor is of type (m, n), with m indices up and n indices down. It is convenient to regard functions as tensors of type (0,0). Clearly (contravariant) vectors X^i and (covariant) vectors W_i are tensors of type (1,0) and type (0,1) respectively.

Example: The quantities $I^i{}_j$ that are equal to the Kronecker delta δ^i_j in an initial coordinate system, are the components of the unit tensor **I** which is of invariant form: it has the same components in all coordinate systems (try any coordinate transformation; you will find $I^{i'}{}_{j'} = \delta^{i'}_{j'}$). This is the *unit tensor*, with the properties: (a) for any contravariant vector X^a, $I^a{}_b X^b = X^a$; (b) for any covariant vector W_a, $W_a I^a{}_b = W_b$; (c) consequently, acting on any tensor, it gives the identity transformation; and this applies in particular if it acts on itself; $I^a{}_b I^b{}_c = I^a{}_c$.

The transformation matrices satisfy the properties of a group, because the coordinate transformations also do so. Specifically,

(a) The identify transformation $x^i \to x^{i'} = x^i$ leads to the unit matrix $A^{i'}{}_i = \delta^{i'}{}_i$, giving the identity tensor transformation in (28),

(b) the inverse (19) to the arbitrary transformation (18) replaces the matrix $A^{i'}{}_i$ by its inverse $A_{i'}{}^i$, leading to a corresponding tensor transformation with each matrix replaced by its inverse (as we have already seen in particular cases). In the case of the tensor $R^{ijk}{}_{gh}$, (28) has the inverse

$$R^{ijk}{}_{gh} = R^{i'j'k'}{}_{g'h'} A_{i'}{}^{i} A_{j'}{}^{j} A_{k'}{}^{k} A^{g'}{}_{g} A^{h'}{}_{h} \tag{29}$$

(c) composition of two transformations: T_1: $x^i \to x^{i'}$ where $x^{i'} = x^{i'}(x^i)$, and T_2: $x^{i'} \to x^{i"}$ where $x^{i''} = x^{i''}(x^{i'})$, leads to the overall transformation T_3: $x^i \to x^{i"}$ where $x^{i''}(x^i) = x^{i''}(x^{i'}(x^i))$, with the corresponding composite tensor transformation matrix given by the composition formula

$$A^{i''}{}_{i} = A^{i''}{}_{i'} A^{i'}{}_{i} \tag{30}$$

(by (21), this is just the chain rule for partial derivatives). For

$$R^{i''j''k''}{}_{g''h''} = R^{i'j'k'}{}_{g'h'} A^{i''}{}_{i'} A^{j''}{}_{j'} A^{k''}{}_{k'} A^{g'}{}_{g''} A^{h'}{}_{h''} =$$

$$= (R^{ijk}{}_{gh} A^{i'}{}_{i} A^{j'}{}_{j} A^{k'}{}_{k}) A^{i''}{}_{i'} A^{j''}{}_{j'} A^{k''}{}_{k'} A^{g'}{}_{g''} A^{h'}{}_{h''} A^{g}{}_{g'} A^{h}{}_{h'}$$

and so

$$R^{i''j''k''}{}_{g''h''} = R^{ijk}{}_{gh} A^{i''}{}_{i} A^{j''}{}_{j} A^{k''}{}_{k} A^{g}{}_{g''} A^{h}{}_{h''} \tag{31}$$

with the $A^{i''}{}_{i}$ given by (30) . Thus the same result is obtained by performing first T_1 and then T_2, or just performing T_3. Incidentally, the inverse (b) above is obtained as the special case when T_2 is the inverse of T_1.

Thus the set of tensor transformations form a (local) group.

3.2.1 Quotient Law: Tensor detection

There arise occasions where we know that all quantities in an equation except one are tensors; it then follows that the remaining quantity is a tensor. This can be very useful. As a specific example:

If $W_j = X^i T_{ij}$ is a vector for all vectors X^i, then T_{ij} is a tensor.

To prove the above statement, we consider a change of coordinates:

$$W_{j'} = A_{j'}{}^{j} W_j = A_{j'}{}^{j} X^i T_{ij} = A_{j'}{}^{j} X^{i'} A^{i}{}_{i'} T_{ij}$$

(The first equality follows from fact that W_j is a vector, the second from the definition of W_j, the third because X^i is a vector). Put $W_{j'} = X^{i'} T_{i'j'}$ (the definition of W_j in the new coordinates), then

$$X^{i'}(A_{j'}{}^{j} A^{i}{}_{i'} T_{ij} - T_{i'j'}) = 0$$

for all $X^{i'}$, so that

$$T_{i'j'} = A_{j'}{}^{j} A^{i}{}_{i'} T_{ij}$$

as required (to see the last step, choose $X^{i'}$ successively in each coordinate direction, i.e. set $X^{i'} = \delta^{i'}_{k'}$ for $k' = 1, 2, .., n$). □

Similarly one can prove that if $T_{ij}X^{ij}$ is invariant for all tensors X^{ij}, then T_{ik} is a tensor. (The proof is almost identical to that above where the transformation properties of g_{ij} followed from the invariance of ds^2).

3.3 Tensor Equations

The fundamental point of tensor transformations is that

> *If a tensor equation is true in one coordinate system, it is true in all coordinate systems.*

For example, if

$$T_{ij} = R_{ij}$$

in an initial coordinate system, where T_{ij} and R_{ij} are components of tensors, then

$$T_{i'j'} = A_{i'}{}^{i} A_{j'}{}^{j} T_{ij} = A_{i'}{}^{i} A_{j'}{}^{j} R_{ij} = R_{i'j'}$$

(the first equality is because T_{ij} are components of a tensor, the second because the tensor equation is true in the first coordinate system, and the third because R_{ij} are components of a tensor). Thus $T_{i'j'} = R_{i'j'}$ in every coordinate system.

This is the property we would like to be true for physically meaningful equations: they should hold independent of the coordinate system used (otherwise we can make an effect change, or even 'go away', by simply changing the coordinate system). It implies in particular

> *If a tensor vanishes in one coordinate system, it vanishes in all coordinate systems.*

For example, consider the case when $T_{ij} = 0$ in a particular coordinate system; it will follow by the above argument that $T_{i'j'} = 0$ in every coordinate system.

A general tensor equation can have any number of free indices on the right provided the identical indices occur on the left. It implies equality of the left and right for every choice of the values of the free indices. Thus for example $R^{ijk}{}_{gh} = V^{ijk}{}_{gh}$ is a tensor equation implying that $R^{111}{}_{11} = V^{111}{}_{11}$ (on choosing $i = j = k = g = h = 1$, $R^{123}{}_{01} = V^{123}{}_{01}$ (on choosing $i = 1, j = 2, k = 3, g = 0, h = 1$), and so on. Thus if there are n indices in a 4-dimensional space, then in general such an equation represents 4^n separate relations. There will however be fewer independent relations if the tensors have specific symmetries, as discussed below.

3.3.1 Tensor operations

Tensor equations of this simple kind are relatively uninteresting. We construct interesting tensor equations by using four basic algebraic tensor operations, considered here, and tensor differentiation, considered in the following Chapter.

(1) *Linear combination*:

Given two tensors of the same type, we may form new tensors by taking linear combinations of them. For example, if ${T^{ij}}_c$ and ${R^{ij}}_c$ are tensors and α, β numbers, then

$$
{S^{ij}}_c = \alpha {T^{ij}}_c + \beta {R^{ij}}_c
$$

is also a tensor (of type indicated by the indices). In order to do this, it is essential that the indices on the tensors are the same.

In particular, the addition of two tensors ($\alpha = \beta = 1$) give us another tensor, and so does subtraction ($\alpha = 1,\ \beta = -1$).

(2) *Multiplication*:

Given any two tensors, they can be multiplied together to get a new tensor. For example, if ${R^{ij}}_c$ and ${S^{ef}}_{gh}$ are tensors, they define a new tensor ${{{K^{ij}}_c}^{ef}}_{gh}$ by the relation

$$
{R^{ij}}_c {S^{ef}}_{gh} = {{{K^{ij}}_c}^{ef}}_{gh}
$$

In particular, as functions can be regarded as tensors (with no free indices), we can multiply tensors by functions to get new tensors. Thus $T^{ij} = fR^{ij}$ gives us new tensor from a tensor R^{ij} and a function f.

(3) *Contraction*:

Given a tensor with at least one index up and one index down, one can form a new tensor by contracting such indices. For example, if ${R^{ij}}_{cd}$ is a tensor then we can form its contraction on indices 'i' and 'c':

$$
{R^{ij}}_{cd} \to {T^j}_d = {R^{ij}}_{id}.
$$

With each such summation, m indices get reduced to $m-2$ indices.

I will give the proof that ${T^j}_d$ is indeed a tensor (similar proofs can be given for the other operations defined above). Under a general change of coordinates,

$$
{T^{j'}}_{d'} = {R^{i'j'}}_{i'd'} = {R^{ij}}_{cd} {A^{i'}}_i {A^{j'}}_j {A_{i'}}^c {A_{d'}}^d = {R^{ij}}_{id} {A^{j'}}_j {A_{d'}}^d = {T^j}_d {A^{j'}}_j {A_{d'}}^d
$$

as required, where the first step is by the definition of ${T^j}_d$ (in the new coordinates), the second is because ${R^{ij}}_{cd}$ is a tensor, the third is because the first and third transformation matrices are summed on each other and so give δ^c_i, and the final step is by the definition of ${T^j}_d$ (in the old coordinates). □

The point to note specially is use of the dummy suffix 'c' instead of the 'i' one might be tempted to use at first; the point is that 'i' has at that stage *already* been used as a dummy suffix, and one must *never* use the same dummy suffix twice (doing so will lead to ambiguous equations and errors). So before choosing the name of a dummy suffix, always check that that label has not been used already in the equation!

If all the tensor indices are contracted, we get a scalar. This is a very useful way of forming invariants. For example, $T^i{}_j \to T := T^i{}_i$, a scalar, whose value is independent of the coordinates chosen.

(4) *Raising and Lowering indices*:
Finally, tensor indices can be raised or lowered by using the metric and the inverse metric. To lower indices, we contract with the metric tensor. Thus for example from X^i we can form

$$X_i = g_{ij} X^j \tag{32}$$

which is a tensor (by (20) and (27) above). The convention is to use the same kernel symbol for the quantity with an index lowered as with it raised; this is because these are regarded as different aspects of the same object. We can similarly lower any index. In space-time, despite the negative sign in the metric (11), g_{ij} is still is *non- degenerate*: that is, if $X^i \neq 0$ then $X^i g_{ij} \neq 0$, because its determinant $g = |g_{ij}|$ is non-zero.

To raise an index we need a 'metric tensor' g^{ij} with indices upstairs. Such a quantity is defined as the inverse of the metric g_{ij}, which exists because we always assume the metric g_{ij} is non-degenerate. Thus

$$g^{ij} g_{jk} = \delta^i_k \quad \Rightarrow g^{ij} = g^{ji} \tag{33}$$

This is a tensor because of the tensor detection theorem (see above) plus the fact that g_{ij} and δ^i_k are tensors. We can use it to raise indices; for example,

$$X^i = g^{ij} X_j \tag{34}$$

This procedure can be used to raise any index. This notation is compatible with (32) because if we first lower and then raise the index of X^i, we end up back with the quantities we started with: $X^i = g^{ij} X_j = g^{ij}(g_{jk} X^k) = \delta^i_k X^k = X^i$, by (33).

This enables us to write tensor equations with the indices in any position (up or down) we like. It is for this reason that we leave blanks below each upper index (and above each lower index): if we do not do so, we may run into trouble when we try to raise or lower indices (and are likely to obtain wrong answers if the tensor concerned is not totally symmetric). It also enables us to easily form scalar products and invariants; for example $\mathbf{A}.\mathbf{B} = A^iB^jg_{ij} = A^iB_i = A_jB^j$, $|\mathbf{A}|^2 = \mathbf{A}.\mathbf{A} = A^iA_i$, and $T_{ab} \to T = T^a{}_a = g^{ij}T_{ij} = g_{ij}T^{ij} = T_a{}^a$. Note that when forming such scalars, the position (up or down) of the contracted indices is immaterial.

In space time, as mentioned in section 1.2, $\mathbf{X}.\mathbf{X} = 0$ for a vector X^a implies *either* $X^i = 0$ *or* it is a null vector. The same applies in general to any tensor: vanishing of its magnitude does not in general imply that it vanishes. As a particular example, the curvature tensor R_{ijkl} gives important information about the space time (as discussed in a later chapter). Now it is possible that $R^{ibcd}R_{ibcd} = 0$ but $R_{ibcd} \neq 0$; indeed this occurs in a particular class of plane-wave spacetimes.

Two special cases of raising and lowering indices deserve attention. Firstly, consider raising the index on the metric tensor g_{ij}. If we raise one index, we obtain $g^i{}_j = g^{ik}g_{kj} = \delta^i_j$ (see equation (33)), and so if we raise the second, we obtain $g^{ik}g_{km}g^{jm} = g^{ij}$. Similarly if we lower one index of g^{ij} we obtain δ^i_j, and if we lower the second we therefore obtain g_{ij}. Thus g_{ij}, δ^i_j, g^{ik} are really all aspects of the same tensor, the unit tensor for the space-time. One should note the following fact: the scalar $g^a{}_a$ is just the dimension of the space. For example, in 4-dimensional space-time,

$$g^a{}_a = \delta^a_a = \delta^0_0 + \delta^1_1 + \delta^2_2 + \delta^3_3 = 1 + 1 + 1 + 1 = 4$$

Secondly, start with the metric transformation law $g_{i'j'} = g_{ij}A_{i'}{}^iA_{j'}{}^j$ and multiply by $A^{j'}{}_kg^{km}$. Using (22) and (33) we obtain

$$A_{i'}{}^m = g_{i'j'}A^{j'}{}_kg^{km} \qquad (35)$$

that is, the inverse matrix can be obtained by raising and lowering indices of the forward matrix. This explains why the indices on these matrices should be positioned as they are (the primed index first, and the unprimed last, whether up or down).

3.3.2 Symmetry properties

Symmetries are important in tensor equations, because a tensor transformation preserves symmetries.

Consider first a tensor T_{ij} with two indices downstairs: it can be written as a sum of symmetric and antisymmetric parts:

$$T_{ij} = \frac{1}{2}(T_{ij} + T_{ji}) + \frac{1}{2}(T_{ij} - T_{ji}) = T_{(ij)} + T_{[ij]}$$

where

$$T_{(ij)} := \frac{1}{2}(T_{ij} + T_{ji}), \quad T_{[ij]} := \frac{1}{2}(T_{ij} - T_{ji}) \tag{36}$$

are the symmetric and skew-symmetric parts respectively. Then T_{ij} is called *symmetric* if

$$T_{ij} = T_{(ij)} \Leftrightarrow T_{[ij]} = 0 \Leftrightarrow T_{ij} = T_{ji};$$

in this case it has $\frac{1}{2}n(n+1)$ independent components (for example, the metric tensor). On the other hand T_{ij} is *skew-symmetric* if

$$T_{ij} = T_{[ij]} \Leftrightarrow T_{(ij)} = 0 \Leftrightarrow T_{ij} = -T_{ij};$$

in this case it has $\frac{1}{2}n(n-1)$ independent components. In 4 dimension, there are 10 symmetric parts and 6 antisymmetric parts for a 2-index tensor.

If we change coordinates, then

$$2T_{(i'j')} = T_{i'j'} + T_{j'i'} = T_{ij}A_{i'}{}^{i}A_{j'}{}^{j} + T_{ji}A_{j'}{}^{j}A_{i'}{}^{i} =$$

$$= A_{i'}{}^{i}A_{j'}{}^{j}(T_{ij} + T_{ji}) = A_{i'}{}^{i}A_{j'}{}^{j}\,2T_{(ij)}$$

i.e. the symmetric part transforms as a tensor; so does the skew-symmetric part. Thus being symmetric or skew-symmetric is a tensor relation: it is a property that is invariant under change of basis. As an immediate application, the metric tensor is symmetric in all coordinates.

If $T_{ij} = T_{(ij)} \Rightarrow T^{ij} = T^{(ij)}$, that is, if a tensor is symmetric with its indices down, it will also be symmetric with its indices up; but this is *not* true for the mixed components (symmetry of T_{ij} does not imply symmetry of $T^{i}{}_{j}$ in general). Indeed we should not try to define symmetries with one index up and one down, for such 'symmetries' will not remain true under a general change of coordinates.

Let us now write the symmetric and antisymmetric parts of a 3-index tensor, with all indices on the same level. The symmetric part is

$$F_{(ijk)} = \frac{1}{3!}(F_{ijk} + F_{kij} + F_{jki} + F_{ikj} + F_{jik} + F_{kji})$$

and the skew-symmetric part is

$$F_{[ijk]} = \frac{1}{3!}(F_{ijk} + F_{kij} + F_{jki} - F_{ikj} - F_{jik} - F_{kji}). \tag{37}$$

Similarly, we can take the skew symmetric part over any r indices on the same level by summing over all permutations of these indices, with a plus sign for each even permutation and a minus sign for each odd one; dividing the whole by $r!$. Then interchanging any two neighbouring indices from this set will alter the value of the components by a minus sign. Hence all components in which the same index occurs twice in this set (e.g. F_{11}) will be zero. It is then clear that in four dimensions, any tensor that is skew on five or more indices must vanish; e.g. $T_{ijklm} = T_{[ijklm]} \Rightarrow T_{ijklm} = 0$ (because at least one index must be repeated).

A fundamental fact that will be used repeatedly is as follows: if a tensor T is symmetric on any two indices, e.g. $T^{ij} = T^{(ij)}$, and a tensor W is skew-symmetric on any two indices, e.g. $W_{ij} = W_{[ij]}$, then their contraction over these indices is zero: $T^{ij}W_{ij} = 0$. This follows because $T^{ij}W_{ij} = -T^{ji}W_{ji} = -T^{ij}W_{ij}$, the first step arising because $T^{ij} = T^{ji}$, $W_{ij} = -W_{ji}$, and the second by simultaneously relabelling $i \to j$, $j \to i$, which is allowed as it is exactly the same expression (to check this, write out the sums explicitly!).

Any tensor with r-indices downstairs, that is totally skew on those indices, is called an *r-form.* 4-forms play an important role in determining volumes in a 4-dimensional space (as do 3-forms in a 3- dimensional space, and 2-forms in a 2-dimensional space).

3.3.3 The volume element

Any totally skew quantity w_{ijkm} or v^{ijkm} in 4-dimensional space-time has only one independent component. Any particular component will either vanish (if two indices are the same) or will be equal to $\pm w_{0123}$ (if they are all different), the sign being determined by how many times neighbouring indices must be swapped to bring it to the standard form (e.g. $w_{2013} = -w_{0213} = +w_{0123}$).

Now the totally skew permutation quantities ϵ^{ijkm}, ϵ_{ijkm} (not the components of a tensor, see below) take the values ± 1 or 0 as follows: in any coordinate system,

$$\epsilon^{ijkm} = \epsilon^{[ijkm]}, \quad \epsilon^{0123} = 1; \quad \epsilon_{ijkm} = \epsilon_{[ijkm]}, \quad \epsilon_{0123} = 1. \tag{38}$$

They define determinants in the following way: $A^0{}_i A^1{}_j A^2{}_k A^3{}_m \epsilon^{ijkm} = |A^i{}_j|$ is the definition of the determinant $|A^i{}_j|$ of the matrix $\left(A^i{}_j\right)$, so

$$A^r{}_i A^s{}_j A^t{}_k A^u{}_m \epsilon^{ijkm} = |A^i{}_j| \epsilon^{rstu} \tag{39}$$

(the left and right are totally skew, and take the same value if $r = 0$, $s = 1$, $t = 2$, $u = 3$.) Equally we could write

$$A^r{}_i A^s{}_j A^t{}_k A^u{}_m \epsilon_{rstu} = |A^i{}_j| \epsilon_{ijkm} \tag{40}$$

Theorem: If $C^i{}_j = A^i{}_k B^k{}_j$, and $C := |C^i{}_j|$, $A := |A^i{}_j|$, $B := |B^i{}_j|$, then $C = AB$. (the determinant of a contracted product is the product of the determinants).

Proof: Using (39) for C, the definition of $C^i{}_j$, and then (40) twice for B abd then A,

$$\begin{aligned} C\epsilon^{ijkl} &= C^i{}_s C^j{}_r C^k{}_t C^l{}_u \epsilon^{srtu} \\ &= A^i{}_a B^a{}_s A^j{}_b B^b{}_r A^k{}_c B^c{}_t A^l{}_d B^d{}_u \epsilon^{srtu} \\ &= A^i{}_a A^j{}_b A^k{}_c A^l{}_d \, B\epsilon^{abcd} = AB\,\epsilon^{ijkl} \end{aligned}$$

as required. □

In particular when applied to (22) this shows that the determinants of the inverse transformation matrices are inverses of each other:

$$|A^{i'}{}_i||A_{j'}{}^j| = 1. \qquad (41)$$

Also when applied to the metric transformation law (27), defining $|g_{ij}| =: g$ (the determinant of the metric tensor), we find that on change of coordinates, provided $|A_{i'}{}^i| > 0$

$$g' = |A_{i'}{}^i|^2 g \quad \Leftrightarrow \quad \sqrt{|g'|} = |A_{i'}{}^i|\sqrt{|g|} \qquad (42)$$

where we take the positive square root on both sides, and the modulus is necessary in the square root because $g < 0$ (indeed in Minkowski coordinates in flat space-time, $g = -1$). If $|A_{i'}{}^i| < 0$ then $\sqrt{|g'|} = -|A_{i'}{}^i|\sqrt{|g|}$.

Now consider the totally skew tensor v^{ijkm} that in an initial coordinate system is equal to ϵ^{ijkm}. In a new frame we find

$$v^{i'j'k'm'} = A^{i'}{}_i A^{j'}{}_j A^{k'}{}_k A^{m'}{}_m \epsilon^{ijkm} = |A^{i'}{}_i|\epsilon^{i'j'k'm'} \qquad (43)$$

on using (39). Thus the quantities ϵ^{ijkm} are not components of a tensor. However on using (42) we can form the tensor volume element η^{ijkm} as follows: it is a tensor which in an initial frame, has the form

$$\eta^{ijkm} = \frac{1}{\sqrt{|g|}}\epsilon^{ijkm} \quad \Leftrightarrow \quad \eta^{ijkm} = \eta^{[ijkm]}, \;\; \eta^{0123} = \frac{1}{\sqrt{|g|}} \qquad (44)$$

Then in a new frame

$$\eta^{i'j'k'm'} = |A^{i'}{}_i|\frac{1}{\sqrt{|g|}}\epsilon^{i'j'k'm'} = \frac{1}{|A_{i'}{}^i|\sqrt{|g|}}\epsilon^{i'j'k'm'} = \frac{1}{\sqrt{|g'|}}\epsilon^{i'j'k'm'}$$

the first equality following from (43), the second from (41), and the third from (42) provided $|A_{i'}{}^i| > 0$. Thus it will again take the same form (44) in the new coordinates as long as the determinant of the transformation is positive (that is, as long as it preserves the orientation of the axes). Hence η^{ijkm} will have the form (44) in every coordinate system, if the orientation of the axes is preserved.

If the orientation is changed, we get the same form but with a minus sign.

If we lower the indices on η^{ijkm} (given by (44)), we find

$$\eta_{ijkl} = g_{ir}g_{js}g_{kt}g_{lu}\eta^{rstu} = g_{ir}g_{js}g_{kt}g_{lu}\frac{1}{\sqrt{|g|}}\epsilon^{rstu} = g\frac{1}{\sqrt{|g|}}\epsilon_{ijkl}$$

where the last equality arises by the same argument that lead to (39). Hence, remembering that $g < 0 \Rightarrow g = -(\sqrt{|g|})^2$, the downstairs components are

$$\eta_{ijkl} = -\sqrt{|g|}\,\epsilon_{ijkl} \quad \Leftrightarrow \quad \eta_{ijkl} = \eta_{[ijkl]}, \quad \eta_{0123} = -\sqrt{|g|} \tag{45}$$

The minus sign arises essentially because lowering the indices is equivalent to a change of basis that alters the orientation of the axes.

The pseudo-tensor η_{abcd} is essentially the *volume element* of space-time. More precisely, consider the 4-dimensional parallelepiped spanned by the arbitrary displacements dx^i_0, dx^j_1, dx^k_2, dx^l_3, its volume is

$$dV = |\eta_{ijkl}dx^i_0 dx^j_1 dx^k_2 dx^l_3| = \sqrt{|g|}\left(|\epsilon_{ijkl}dx^i_0 dx^j_1 dx^k_2 dx^l_3|\right) =: \sqrt{|g|}dv \tag{46}$$

showing that $\sqrt{|g|}$ is the factor relating coordinate volume dv to physical volume dV. If we choose dx^i_0 as a displacement a coordinate distance ϵ_0 along the 0-coordinate curve: $dx^i_0 = \epsilon_0\delta^i_0$, and similarly for the other displacements, then $dv = |\epsilon_0\epsilon_1\epsilon_2\epsilon_3|$ and $dV = |\eta_{0123}|dv$. This makes clear why we demand $g \neq 0$ for a regular space-time point.

The relation of this 4-volume to the 3-volume measured by an observer is as follows: consider a 4-velocity vector u^d of an observer, such that $u^d u_d = -1$ (this will be considered more in the next chapter). Then the effective volume element measured by such an observer is

$$\eta_{ijk} = \eta_{ijkm}u^m \quad \Rightarrow \quad \eta_{ijk} = \eta_{[ijk]}, \quad \eta_{ijk}u^k = 0 \tag{47}$$

the last relation arising because $\eta_{ijk}u^k = \eta_{ijkl}u^k u^l$, the contraction of a skew quantity with a symmetric quantity. This shows that η_{ijk} lies in the rest-space of the observer. It defines 3-space volumes by the obvious expression analogous to (46). It also defines the 3-space vector product as follows: if X^i, Y^j are in the rest-space of u^i (i.e. $u^i X_i = 0 = u^k Y_k$) then their vector product $Z \times Y$ and triple scalar product (X, Y, W) with a vector W^k (also in the rest space) are given by

$$Z^i = (X \times Y)^i = \eta^{ijk}X_j Y_k, \quad (X, Y, W) = \eta_{ijk}X^i Y^j W^k \tag{48}$$

There are a series of *identities* involving the tensor η^{ijkl} that are often useful. The fundamental one is the following:

$$\eta^{abcd}\eta_{efgh} = -4!\delta^a_{[e}\delta^b_f\delta^c_g\delta^d_{h]} \tag{49}$$

This is proved by noting that this is a tensor equation with only one independent component (both the right and the left are totally skew in all the top indices, and all the bottom indices). Setting $a = 0$, $b = 1$, $c = 2$, $d = 3$, $e = 0$, $f = 1$, $g = 2$, $h = 3$ we find on the left $\eta^{0123}\eta_{0123} = \left(1/\sqrt{|g|}\right)\left(-\sqrt{|g|}\right) = -1$ on using (26), (27); while on the right we find $-4!\delta^0_{[0}\delta^1_1\delta^2_2\delta^3_{3]} = -(4!/4!)(\delta^0_0\delta^1_1\delta^2_2\delta^3_3 + \delta^0_1\delta^1_0\delta^2_2\delta^3_3 + ...) = -(1 + 0 + 0 + ...) = -1$ where the factor 1/4! comes from the definition of the skew- symmetrisation square brackets, and only one of all the permutations implied by those brackets is non-zero when one writes them all out. Thus the left and right agree. □

A further series of useful identities arise if we now contract systematically, remembering that $\delta^a_a = 4$. We find (put $e = a$, then $f = b$, then $g = c$, and finally $h = d$):

$$\eta^{abcd}\eta_{afgh} = -3!\delta^b_{[f}\delta^c_g\delta^d_{h]} \tag{50}$$
$$= -\delta^b_f\delta^c_g\delta^d_h - \delta^b_h\delta^c_f\delta^d_g - \delta^b_g\delta^c_h\delta^d_f + \delta^b_f\delta^c_h\delta^d_g + \delta^b_g\delta^c_f\delta^d_h + \delta^b_h\delta^c_g\delta^d_f$$
$$\eta^{abcd}\eta_{abgh} = -4\delta^c_{[g}\delta^d_{h]} = 2(-\delta^c_g\delta^d_h + \delta^c_h\delta^d_g) \tag{51}$$
$$\eta^{abcd}\eta_{abch} = -3!\delta^d_h \tag{52}$$
$$\eta^{abcd}\eta_{abcd} = -4! \tag{53}$$

[nb: Considerable care is needed in deducing the corresponding 3-dimensional relations from the 4-d ones, bcause of the space-time signature. Direct deduction of the corresponding 3-dimensional relations in a positive definite 3-dimensional space-time is straightforward.]

As an example of their application, given any skew tensor $F_{ij} = F_{[ij]}$ we can define its *dual* $^*F_{ij}$ by the relation

$$^*F_{ij} = \frac{1}{2}\eta_{ij}{}^{cd}F_{cd} = \frac{1}{2}\eta_{ijcd}F^{cd} \tag{54}$$

Then (34) shows that

$$^*(^*F_{ij}) = \frac{1}{2}\eta_{ij}{}^{cd}(\frac{1}{2}\eta_{cd}{}^{ef}F_{ef}) = \frac{1}{4}(\eta_{ijcd}\eta^{cdef})F_{ef} =$$

$$= -\frac{4}{4}(\delta^e_{[i}\delta^f_{j]}F_{ef}) = -F_{ij}$$

that is, the dual * is an *involutive operator* on skew tensors: $^*(^*F) = -F$.

[n.b.: Don't confuse this type of (bivector) dual with dual basis vectors, and with dual vectors].

Examples 2: Vectors and Tensors

1. (i) Prove that if X_aY^a is invariant for all vectors Y^a, then X_a is a vector.
(ii)Consider the vector A^a and transformation $A^{a'} = \Lambda^{a'}_a A^a$, where $\Lambda^{a'}_a$ is an arbitrary transformation. Show that $A.B = A'.B'$.

2. Let $h_{ab} = g_{ab} + u_a u_b$, where $u_a u^a = -1$. This is the projection tensor. Prove that

$$h^b_a h^c_b = h^c_a \qquad h_{ab}u^b = 0 \qquad h^a_a = 3$$

What is h_{ab} in an orthonormal frame?

3. Let $F_{ab} = -2E_{[a}u_{b]} + \eta_{ab}{}^{cd}H_c u_d$. Find the vectors E_a and H_a in terms of the tensor F_{ab} (Hint: contract F_{ab} with u^b and $\eta^{abst}u_t$). Interpret your results in an orthonormal frame.

4. Let us define the product $(X \times Y)^a = \eta^{abc}X_bY_c$ where $\eta^{abc} = \eta^{abcd}u_d$. Using the identity $\eta^{abc}\eta_{def} = 3!h^a_{[d}h^b_e h^c_{f]}$ find an expression for the components of the triple product $((X \times Y) \times Z)^a$, where $X_a u^a = Y_a u^a = Z_a u^a = 0$.

5. Show that $V_{ab} = Y_{a,b}$ is not a tensor if Y_a is a vector.

---

4 General Bases

When we use a coordinate basis (as we have done up to now) the components of a tensor do not directly tell us the magnitude of any physical quantity, because the basis vectors have an arbitrary magnitude. For example if a vector $\mathbf{X}$ points in the 1-coordinate direction, then $\mathbf{X} = X^1 \frac{\partial}{\partial x^1}$ and the magnitude of $\mathbf{X}$ is given by $|\mathbf{X}|^2 = g_{ab}X^a X^b = g_{11}(X^1)^2$. Hence if say $g_{11} = 100$ then $X^1 = |\mathbf{X}|/10$: the component is a multiple of the magnitude. However if $g_{11} = 1$, then $X^1 = |\mathbf{X}|$: the component equals its magnitude. For a general vector, the components will only give the magnitude of the displacement in each component direction if the corresponding metric component is unity. Furthermore they will only give the standard 'sum of squares' form for magnitudes (with unit coefficients) if the basis vectors are orthogonal to each other. Thus it is useful to be able to find basis vectors with such metric properties (often called a 'physical basis'). We can always do this.

It is helpful to consider two simple examples. First, consider the Euclidean plane in the polar coordinates discussed earlier. Then $ds^2 = dr^2 + r^2 d\theta^2 \Leftrightarrow g_{ij} = diag(1, r^2)$ so the magnitudes of the coordinate basis vectors are $g_{rr} = (\frac{\partial}{\partial r}).(\frac{\partial}{\partial r}) = 1$, and $g_{\theta\theta} = (\frac{\partial}{\partial \theta}).(\frac{\partial}{\partial \theta}) = r^2$, with scalar product $g_{r\theta} = (\frac{\partial}{\partial r}).(\frac{\partial}{\partial \theta}) = 0$. We now choose an orthonormal basis $\{\mathbf{e}_a\}$ parallel to this coordinate basis; we do so by dividing each vector by its magnitude. Thus such an orthonormal basis e_a is given by

$$\mathbf{e}_r = \frac{\partial}{\partial r}, \quad \mathbf{e}_\theta = \frac{1}{r}\frac{\partial}{\partial \theta}.$$

Then $\mathbf{e}_r.\mathbf{e}_r = (\frac{\partial}{\partial r}).(\frac{\partial}{\partial r}) = g_{rr} = 1$, $\mathbf{e}_\theta.\mathbf{e}_\theta = (\frac{1}{r})^2(\frac{\partial}{\partial \theta}).(\frac{\partial}{\partial \theta}) = \frac{1}{r^2}g_{rr} = \frac{1}{r^2}r^2 = 1$ showing the basis vectors are unit vectors, and $\mathbf{e}_r.\mathbf{e}_\theta = \frac{1}{r}(\frac{\partial}{\partial r}).(\frac{\partial}{\partial \theta}) = \frac{1}{r}g_{r\theta} = 0$, showing they are also orthogonal. This is thus indeed an orthonormal basis, such that $g_{ab} = diag(1, 1)$.

We can write the above transformation in the form

$$\mathbf{e}_a = \Lambda_a{}^i(x^k)\frac{\partial}{\partial x^i} \quad \Leftrightarrow \quad \frac{\partial}{\partial x^i} = \Lambda^a{}_i(x^k)\mathbf{e}_a \tag{55}$$

where

$$\Lambda_a{}^i = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{r} \end{pmatrix}$$

and the inverse is

$$\Lambda^a{}_i = \begin{pmatrix} 1 & 0 \\ 0 & r \end{pmatrix}$$

This is very similar to the equation following (24), and implies that components of vectors will transform in the obvious way:

$$\mathbf{X} = X^a\mathbf{e}_a = X^i\frac{\partial}{\partial x^i} \tag{56}$$

which shows that

$$\mathbf{X} = X^a \Lambda_a{}^i(x^k)\frac{\partial}{\partial x^i} = X^i \frac{\partial}{\partial x^i}$$

and so

$$X^a = \Lambda^a{}_i X^i \quad \Leftrightarrow \quad X^i = \Lambda_a{}^i X^a \tag{57}$$

Thus vectors transform just as before, but with transformation matrix $\Lambda_a{}^i$ instead of $A_{i'}{}^i$. In fact this will hold for all tensor transformations: the same equations (28), (29) hold as before, but with the new transformation matrix.

The second example is the (curved) surface of a sphere. Then $ds^2 = dr^2 + \sin^2 r d\theta^2 \Leftrightarrow g_{ij} = diag(1, \sin^2 r)$. Proceeding just as above, we choose an orthonormal basis $\{\mathbf{e}_a\}$ parallel to this coordinate basis, given by

$$\mathbf{e}_r = \frac{\partial}{\partial r}, \quad \mathbf{e}_\theta = \frac{1}{\sin r}\frac{\partial}{\partial \theta}.$$

Then as above, $g_{ab} = diag(1, 1)$ for this basis. Again we can write the transformation in the form (1), where now

$$\Lambda_a{}^i = diag(1, \frac{1}{\sin r}), \quad \Lambda^a{}_i = diag(1, \sin r)$$

and (56), (57) hold with this new transformation matrix.

Life is simplified in each case by having this new basis with a simple metric form. For an arbitrary displacement

$$\eta^a = \Lambda^a{}_i(x^k)dx^i \quad \Leftrightarrow \quad dx^i = \Lambda_a{}^i(x^k)\eta^a \tag{58}$$

we have that

$$ds^2 = g_{ab}\eta^a\eta^b = g_{ab}\Lambda^a{}_i(x^k)dx^i\Lambda^b{}_j(x^k)dx^j \tag{59}$$

Because this is an orthonormal basis, the left hand equality is

$$ds^2 = (\eta^1)^2 + (\eta^2)^2$$

as in a Euclidean basis, and for the scalar product of two vectors, $\mathbf{X}.\mathbf{Y} = X^a g_{ab} Y^b = X^1Y^1 + X^2Y^2$, so in particular $\mathbf{X}^2 = (X^1)^2 + (X^2)^2$. What then is the penalty? The cost is that, as will be shown below, in each case the new basis is not a coordinate basis: that is, there do not exist any coordinates z^1, z^2 such that $\mathbf{e}_a = \partial/\partial z^a$. Correspondingly, we cannot express an arbitrary displacement $\eta^a = \Lambda^a{}_i(x^k)dx^i$ as an exact differential $\eta^a = dz^a$. If we could do so, we could express the metric as $ds^2 = (dz^1)^2 + (dz^2)^2$. In the first case (the Euclidean plane) this is possible but only if we change to quite different coordinates; in the second case (the surface of a sphere), it is not possible. What we do always have is (from (58) and (59))

$$ds^2 = (\Lambda^1{}_i(x^k)dx^i)^2 + (\Lambda^2{}_j(x^k)dx^j)^2$$

clearly expressing the local Pythagorean distance relation for the space.

These examples show the possibility of using more general bases than coordinate bases; they are often very useful in general relativity. In the next section, we will consider changing to completely general bases, and then will consider the particular properties of coordinate and tetrad (orthonormal) bases.

4.1 General basis

A general basis $\{\mathbf{e}_a\}$ can be determined from a coordinate basis by the relations (55). Given a general basis, we can change to another general basis by making the change

$$\mathbf{e}_{a'} = \Lambda_{a'}{}^{a}(x^k)\mathbf{e}_a \quad \Leftrightarrow \mathbf{e}_a = \Lambda^{a'}{}_{a}(x^k)\mathbf{e}_{a'} \tag{60}$$

where

$$\Lambda_{a'}{}^{a}(x^k)\Lambda^{a'}{}_{b}(x^k) = \delta^a_b \quad \Rightarrow \quad det(\Lambda^{a'}{}_{a}) \neq 0. \tag{61}$$

(and clearly, (55) is a special case, as is the change from one coordinate basis to another). As before (cf. (24)),

$$\mathbf{X} = X^a\mathbf{e}_a = X^{a'}\mathbf{e}_{a'} \tag{62}$$

shows that

$$X^{a'} = \Lambda^{a'}{}_{a}X^a \quad \Leftrightarrow \quad X^a = \Lambda_{a'}{}^{a}X^{a'} \tag{63}$$

(with (20,23,57) as special cases). Consequently for the tensor transformation law we find (just as in the previous chapter) that for example

$$T^{a'b'}{}_{c'd'} = \Lambda^{a'}{}_{a}\Lambda^{b'}{}_{b}\Lambda_{c'}{}^{c}\Lambda_{d'}{}^{d}T^{ab}{}_{cd} \quad \Leftrightarrow \quad T^{ab}{}_{cd} = \Lambda_{a'}{}^{a}\Lambda_{b'}{}^{b}\Lambda^{c'}{}_{c}\Lambda^{d'}{}_{d}T^{a'b'}{}_{c'd'} \tag{64}$$

Indeed this form is required to allow the other tensor operations defined in the previous chapter to be good operations defined in any basis, for example to preserve contractions, so that $X^aW_a = X^{a'}W_{a'}$ for this change of basis. In particular the transformation for the metric is

$$g_{a'b'} = \Lambda_{a'}{}^{a}\Lambda_{b'}{}^{b}g_{ab} \quad \Leftrightarrow \quad g_{ab} = \Lambda^{a'}{}_{a}\Lambda^{b'}{}_{b}g_{a'b'} \tag{65}$$

giving the invariant interval (59) for displacements (58), whatever basis is used. Then all the tensor operations can be performed as in the last chapter and are maintained under change of basis, for example in any base, there is a unique inverse metric defined by

$$g^{ab}g_{bc} = \delta^a_c = g^a{}_c \tag{66}$$

and g_{ab}, g^{bc} can be used to raise and lower the indices defined for that base, e.g.

$$X^a = g^{ab}X_b \Leftrightarrow X_a = g_{ab}X^b. \tag{67}$$

As particular examples, (66) raises and lowers the indices on the metric, and exactly as in (35) it follows from (65) that we can raise and lower indices of $\Lambda^{a'}{}_a$ to get

$$\Lambda_{a'}{}^a = g_{a'b'}\Lambda^{b'}{}_b g^{ab} \quad \Leftrightarrow \quad \Lambda^{a'}{}_a = g^{a'b'}\Lambda_{b'}{}^b g_{ab} \tag{68}$$

This equation is related to the following fact: the change of basis idea applies in particular to raising and lowering indices. This can be thought of as giving a change of basis $\mathbf{e}_a \to \mathbf{e}^a$ and vice versa, where $\mathbf{e}^a$ is a new set of basis vectors (the 'dual basis' to $\mathbf{e}_a$) defined by

$$\mathbf{e}^a = g^{ab}\mathbf{e}_b, \quad \mathbf{e}_a = g_{ab}\mathbf{e}^b \quad \Leftrightarrow \quad (\mathbf{e}^a).(\mathbf{e}_b) = \delta^a_b \tag{69}$$

The last equality states that each vector of the basis $\mathbf{e}^a$ is orthogonal to every vector of the basis $\mathbf{e}_a$ except the one with the same label, with which it has scalar product 1; this uniquely defines the geometrical relation between the two sets of basis vectors at each point of any space or space-time.

Now (69) is in fact a particular form of (60), and raising and lowering indices is (from this viewpoint) just a change of basis for the index concerned, from the basis $\mathbf{e}_a$ to the dual basis (or vice versa). For example (67) is then a particular case of (63).

Finally it will be useful to interpret the change matrix in two different ways. Firstly, as in the previous chapters, a vector $\mathbf{X}$ is precisely a basis vector $\mathbf{e}_k$ relative to the basis set $\mathbf{e}_a$ iff $X^a = \delta^a_k$. In this case, equation (63) says that

$$X^{a'} = \Lambda^{a'}{}_a \delta^a_k = \Lambda^{a'}{}_k$$

Thus, $\Lambda^{a'}{}_k$ is the a'-component of the (old) basis vector $\mathbf{e}_k$ relative to the new basis $\mathbf{e}_{b'}$; conversely, $\Lambda_{k'}{}^a$ is the a-component of the (new) basis vector $\mathbf{e}_{k'}$ relative to the old basis $\mathbf{e}_b$. Thus we can write

$$\Lambda^{a'}{}_k = (e_k)^{a'}, \quad \Lambda_{k'}{}^a = (e_{k'})^a \tag{70}$$

expressing the change matrix in terms of basis vector components. This always gives us a geometric interpretation of the change matrix (at each point it relates the directions of the new basis to those of the old). If the basis $\mathbf{e}_a$ is a coordinate basis, then this relates the directions of the new basis to the coordinates (cf the interpretation of vector components based on equation (6)). Also, taking the components of the first of equations (69) with respect to some basis $\mathbf{e}_i$, we find

$$(e^a)_i = g^{ab}(e_b)_i = g^{ab}(e_b)^j g_{ij},$$

which is just the same as equation (68) written in different notation; similarly the second of (69) gives $(e_a)^j = g_{ab}(e^b)^j = g_{ab}(e^b)_i g^{ij}$.

Secondly, if we express the unit tensor I (with components $I^a{}_b = \delta^a_b$) in terms of mixed components $I^a{}_{b'}$, we find $I^a{}_{b'} = I^a{}_b \Lambda_{b'}{}^b = \delta^a_b \Lambda_{b'}{}^b = \Lambda_{b'}{}^a$, so

$$\Lambda_{b'}{}^a = I^a{}_{b'}, \quad \Lambda^{b'}{}_a = I^{b'}{}_a \tag{71}$$

From this viewpoint, the change laws (63)-(65) and (67) just express the identity resulting when a tensor is contracted with the unit tensor, with the latter chosen with mixed components as desired. This shows we need not use the same bases for the different components of a tensor, and (65,66,68) are all just the identity relation $I^a{}_b I^b{}_c = I^a{}_c$ expressed in different bases.

Whether or not there are coordinates for a basis e_a, we will use a subscript comma for the directional derivatives associated with a basis. When this is a coordinate basis, this is the same as $f_{,i} = \frac{\partial}{\partial x^i} f$; more generally, if $\mathbf{e}_a = e_a{}^i \frac{\partial}{\partial x^i}$ then

$$f_{,a} = e_a(f) = e_a{}^i \frac{\partial}{\partial x^i} f \quad \Leftrightarrow \quad Z(f) = Z^a f_{,a}$$

4.1.1 Coordinate bases

Now we want to establish a criterion: given an arbitrary basis $\mathbf{e}_i$, what condition ensures the existence of coordinates x^i such that $\mathbf{e}_i = \frac{\partial}{\partial x^i}$?

In order to do this, we define the *commutator* $[\mathbf{X}, \mathbf{Y}]$ of two vector fields $\mathbf{X} = X^i \frac{\partial}{\partial x^i}$ and $\mathbf{Y} = Y^j \frac{\partial}{\partial x^j}$ as follows:

$$[\mathbf{X}, \mathbf{Y}] = \mathbf{XY} - \mathbf{YX} \quad \Leftrightarrow \quad [\mathbf{X}, \mathbf{Y}]f = X(Y(f)) - Y(X(f)) \tag{72}$$

the right hand equality being true for all functions $f(x^i)$. In particular if we choose $f = x^k$ and let $\mathbf{Z} = [\mathbf{X}, \mathbf{Y}]$, then (cf. (6)) we find the components

$$Z^k = Z(x^k) = [X, Y](x^k) = X(Y(x^k)) - Y(X(x^k)) = X(Y^k) - Y(X^k) =$$

$$= X^j \frac{\partial}{\partial x^j} Y^k - Y^j \frac{\partial}{\partial x^j} X^k$$

where we have used (2) twice in the last equality (with $f = Y^k$ and $f = X^k$ respectively). Thus when a coordinate basis is used,

$$\mathbf{Z} = [\mathbf{X}, \mathbf{Y}] \quad \Leftrightarrow \quad Z^i = Y^i{}_{,j} X^j - X^i{}_{,j} Y^j \tag{73}$$

gives the components of the commutator.

We can now formulate the criterion as follows: *for an arbitrary basis* $\mathbf{e}_i$*, the vanishing of the commutators of the basis vectors determines if it is a coordinate basis or not*:

$$[\mathbf{e}_i, \mathbf{e}_j] = 0 \quad \Leftrightarrow \quad \exists y^i : \quad \mathbf{e}_i = \frac{\partial}{\partial y^i}$$

Thus in particular, if $[\mathbf{e}_i, \mathbf{e}_j] \neq 0$ then there exist no coordinates for which $\mathbf{e}_i$ is a coordinate basis. This is because if it is a coordinate basis, then coordinates exist such that $[e_i, e_j]f = \frac{\partial}{\partial y^i}(\frac{\partial}{\partial y^j}f) - \frac{\partial}{\partial y^j}(\frac{\partial}{\partial y^i}f) = 0$. This can also be obtained from (19) on choosing $X^k = \delta^k_i$, $Y^k = \delta^k_j$.

As a simple example, consider the orthonormal basis $\mathbf{e}_a$ for R^2 discussed at the beginning of this chapter. In that case, $[e_r, e_\theta]f = \frac{\partial}{\partial r}(\frac{1}{r}\frac{\partial f}{\partial \theta}) - \frac{1}{r}\frac{\partial}{\partial \theta}(\frac{\partial f}{\partial r}) = -\frac{1}{r^2}\frac{\partial f}{\partial \theta}$ for any function f; thus

$$[\mathbf{e}_r, \mathbf{e}_\theta] = -\frac{1}{r}\mathbf{e}_\theta$$

This is not zero. Therefore $\mathbf{e}_r$ and $\mathbf{e}_\theta$ do not form a coordinate basis.

If one basis is a coordinate basis and we make a transformation (55) to a new basis, then the criterion that the new basis also be a coordinate basis is that the transformation matrix be a matrix of partial derivatives: $\Lambda^{a'}{}_a = \frac{\partial x^{a'}}{\partial x^a}$ (see (21), (22)). This will be the case iff

$$\frac{\partial}{\partial x^b}\Lambda^{a'}{}_a = \frac{\partial}{\partial x^a}\Lambda^{a'}{}_b \quad \Leftrightarrow \quad \frac{\partial^2 x^{a'}}{\partial x^a \partial x^b} = \frac{\partial^2 x^{a'}}{\partial x^b \partial x^a}.$$

A general transformation does not satisfy this restriction.

4.2 Tetrad bases

We will use the name 'tetrad basis' (or simply 'tetrad') for any basis where the metric components $g_{ab} = \mathbf{e}_a.\mathbf{e}_b$ are constants.

4.2.1 Orthonormal bases

A special role is played in space-time by an orthonormal basis, that is a tetrad basis for which

$$g_{ab} = \mathbf{e}_a.\mathbf{e}_b = diag(-1,1,1,1) =: \eta_{ab} \quad \Leftrightarrow \quad g^{ab} = \eta^{ab} = diag(-1,1,1,1) \quad (74)$$

for then each component of a vector relative to this basis gives the actual magnitude of the component in that direction, and the magnitudes of vectors are given by

$$\mathbf{X}.\mathbf{X} = X^a g_{ab} X^b = -(X^0)^2 + (X^1)^2 + (X^2)^2 + (X^3)^2 \quad (75)$$

Furthermore the invariant interval (59) becomes

$$ds^2 = g_{ab}\eta^a\eta^b = \quad (76)$$

$$= -(\Lambda^0{}_i(x^k)dx^i)^2 + (\Lambda^1{}_i(x^k)dx^j)^2 + (\Lambda^2{}_i(x^k)dx^j)^2 + (\Lambda^3{}_i(x^k)dx^j)^2$$

(which is a way of writing (65)), showing clearly the local special-relativistic structure of an arbitrary curved space-time. We can always locally choose an orthonormal basis, but it cannot be a coordinate basis unless space-time is flat (this will be proved in a following chapter). Clearly $g := |g_{ab}| = -1$, $\eta^{abcd} = \epsilon^{abcd}$, $\eta_{abcd} = -\epsilon_{abcd}$ in such a basis.

To see the effect of raising and lowering indices of the basis vectors (equation (69) above) choose an orthogonal basis of vectors $(\mathbf{t}, \mathbf{x}, \mathbf{y}, \mathbf{z})$ such that $\mathbf{t}.\mathbf{t} = -1$, $\mathbf{x}.\mathbf{x} = \mathbf{y}.\mathbf{y} = \mathbf{z}.\mathbf{z} = 1$, $\mathbf{t}.\mathbf{x} = \mathbf{t}.\mathbf{y} = \mathbf{t}.\mathbf{z} = \mathbf{x}.\mathbf{y} = \mathbf{y}.\mathbf{z} = \mathbf{z}.\mathbf{x} = 0$. Then we can define $\mathbf{e}_a$ by $\{\mathbf{e}_a\} = \{\mathbf{t}, \mathbf{x}, \mathbf{y}, \mathbf{z}\} \Leftrightarrow {e_0}^i = t^i, {e_i}^i = x^i, {e_2}^i = y^i, {e_3}^i = z^i$. Then $\mathbf{e}^a = g^{ab}e_b$ shows that $\{\mathbf{e}^a\} = \{-\mathbf{t}, \mathbf{x}, \mathbf{y}, \mathbf{z}\}$. Thus the dual basis is parallel to the original basis, but with the time vector pointing in the opposite time direction. Another way of writing (76) (obtained by using (70)) is then

$$g_{ij} = -t_i t_j + x_i x_j + y_i y_j + z_i z_j \quad \Leftrightarrow \quad {g^i}_j = -t^i t_j + x^i x_j + y^i y_j + z^i z_j = \delta^i_j \quad (77)$$

We can confirm that is correct by noting that it is the unit tensor, correctly giving each vector when contracted with any of them: for example $g_{ij}t^j = (-t_i t_j + x_i x_j + y_i y_j + z_i z_j)t^j = -t_i(-1) = t_i$, all the other scalar product vanishing.

4.2.2 Lorentz transformations

The transformation (65) is a Lorentz transformation if it preserves the orthonormality of the basis: $g_{ab} = \eta_{ab}$, $g_{a'b'} = \eta_{a'b'}$. Thus the condition for a Lorentz transformation is

$$\eta_{a'b'} = {\Lambda_{a'}}^a {\Lambda_{b'}}^b \eta_{ab} \quad (78)$$

ensuring that (74)-(76) hold in both the old and the new frame. The set of Lorentz transformations form a group (because general tensor transformations form a group, see section 2.2 above).

Arbitrary Lorentz transformations can be composed of spatial rotations and 'boosts' or *velocity transformations*, such as

$$\Lambda^i_{i'}(v) = \begin{pmatrix} \gamma & -\gamma\, v & 0 & 0 \\ -\gamma\, v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (79)$$

with $\gamma(v) = (1-v^2)^{-1/2}$ as usual (see (15)), as we have set $c = 1$. The condition (78) can be verified term by term (set $a' = 0$, $b' = 0$, and carry out the summations, then set $a' = 1$, $b' = 0$, and so on), or by matrix multiplication: write (78) in the matrix form $(\eta_{a'b'}) = ({\Lambda_{a'}}^a)(\eta_{ab})({\Lambda_{b'}}^b)^T$ where in each case left-hand index denotes a row and the right hand one a column, and the transpose of the right hand matrix is taken so that matrix multiplication correctly corresponds to the summation convention in (78) (matrix multiplication corresponds to the summed indices being next to each other). Explicitly,

$$\eta_{a'b'} = \begin{pmatrix} \gamma & -\gamma\, v & 0 & 0 \\ -\gamma\, v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & -\gamma\, v & 0 & 0 \\ -\gamma\, v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

which can easily be verified to give the correct form for $\eta_{a'b'}$. At the end of this chapter, we will give a systematic way of generating continuous Lorentz transformations.

By the group property (see (30,31)), a Lorentz transformation generates an associated family of transformations that preserve *any* metric form. The point is that there will be some (non-unique) transformation $\lambda_a{}^i(x^k)$ that locally transforms an arbitrary metric to the orthonormal form: so given an initial metric g_{ij}, we have

$$\eta_{ab} = \lambda_a{}^i(x^k)\ \lambda_b{}^j(x^k)\ g_{ij}(x^k) \quad \Leftrightarrow \quad g_{ij}(x^k) = \lambda^a{}_i(x^k)\ \lambda^b{}_j(x^k)\ \eta_{ab}. \tag{80}$$

Then the transformation

$$A_j{}^i(x^k) = \lambda^a{}_j(x^k)\ \Lambda_a{}^b(x^k)\ \lambda_b{}^i(x^k) \tag{81}$$

will preserve the form of g_{ij} if $\Lambda^a{}_b(x^k)$ is a (possibly position dependent) Lorentz transformation (the right hand factor $\lambda_b{}^i(x^k)$ brings the metric to the orthonormal form, the Lorentz transformation preserves that form, and the left hand factor $\lambda^a{}_j(x^k)$ brings the metric back to the original form).

4.3 Other tetrads

Given the tetrad $\{\mathbf{e_a}\} = \{\mathbf{t}, \mathbf{x}, \mathbf{y}, \mathbf{z}\}$ as defined above, we can form the null vectors $\mathbf{k}, \mathbf{n}$ by $k^i = (t^i + x^i)/\sqrt{2}$, $n^i = (t_i - x^i)/\sqrt{2}$. It follows immediately that $\mathbf{k}.\mathbf{k} = (\mathbf{t}+\mathbf{x}).(\mathbf{t}+\mathbf{x})/2 = 0$, $\mathbf{k}.\mathbf{n} = (\mathbf{t}+\mathbf{x}).(\mathbf{t}-\mathbf{x})/2 = -1$, $\mathbf{n}.\mathbf{n} = (\mathbf{t}-\mathbf{x}).(\mathbf{t}-\mathbf{x})/2 = 0$ (or in component form, $k^i k_i = (t^i + x^i)(t_i + x_i)/2 = 0$, etc.) This defines a new tetrad $\mathbf{e_i} = \{\mathbf{k}, \mathbf{n}, \mathbf{y}, \mathbf{z}\}$ obtained from the old one by the transformation matrix $\lambda_i{}^a$ where

$$\lambda_i{}^a = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Hence for this tetrad, the metric form is

$$g_{ij} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \mathbf{e_i}.\mathbf{e_j}$$

(obtained either by (65) or directly from the scalar products above) The corresponding expression to (77) in this semi-null basis is

$$g_{ij} = -k_i n_j - n_i k_j + y_i y_j + z_i z_j \quad \Leftrightarrow \quad g^i{}_j = -k^i n_j - n^i k_j + y^i y_j + z^i z_j = \delta^i_j$$

We can go one stage further and define the complex conjugate vectors $m^i = (y^i + iz^i)/\sqrt{2}$, $\overline{m}i = (y_i - iz^i)/\sqrt{2}$, such that $m.\overline{m} = 1$, $m.m = 0$. Then the null tetrad $\{k, n, m, \overline{m}\}$ has the metric components

$$g_{ij} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \mathbf{e_i}.\mathbf{e_j}$$

and the metric can be written as

$$g_{ij} = -k_i n_j - n_i k_j + m_i \overline{m}_j + \overline{m}_i m_j$$

These tetrads have proved very useful in examining gravitational radiation and geometric optics; and null tetrads provide the basis for defining spinors in a curved space-time.

4.4 Physics in an orthonormal basis

An orthonormal basis is a locally physical basis, and so is ideal for examining local physical behaviour. We look at some examples. In each case, their components in an orthonormal frame are physically significant, uniting in one 4-dimensional quantity various disparate 3-dimensional entities. Furthermore important physical information comes from the fact that they transform as tensors under Lorentz transformations, and have specific invariants under such transformations.

4.4.1 4 velocity of particle

Let us denote the 4-velocity of a particle moving on the world line $x^a(\tau)$ by

$$U^i = \frac{dx^i}{d\tau} = (\frac{dx^0}{d\tau}, \frac{dx^1}{d\tau}, \frac{dx^2}{d\tau}, \frac{dx^3}{d\tau}) \qquad (82)$$

where τ is proper time. To understand its meaning consider flat space-time with Minkowski coordinates (t, x, y, z) (so the metric is orthonormal). Then

$$U^i = \frac{dt}{d\tau}(\frac{dx^0}{dt}, \frac{dx^1}{dt}, \frac{dx^2}{dt}, \frac{dx^3}{dt}) = \gamma\ (1, \mathbf{v}) \qquad (83)$$

by (15), where v is the speed of motion of the particle relatve to the Lorentz reference frame. This shows that (82) is a unit vector: by (75)

$$U^a g_{ab} U^b = -(U^0)^2 + (U^1)^2 + (U^2)^2 + (U^3)^2 = \gamma^2(-1 + v^2) = -1$$

The meaning of the components will locally be the same in any orthonormal frame in any space-time.

If we choose an orthonormal frame in a flat or curved spacetime, then the components of the 4-velocity (82) are $U^a = \Lambda^a{}_i dx^i/d\tau$ where $g_{ab} = \eta_{ab}$; and these take the same form as in (83). We can see this in two steps. First, choosing the timelike tetrad vector parallel to the particle's 4-velocity: $\mathbf{U} = \mathbf{e}_0$ then

$$U^a = (1, 0, 0, 0) \quad \Rightarrow U_a = g_{ab}U^b = (-1, 0, 0, 0), \quad U^a U_a = -1 \tag{84}$$

Comparing with (83), this is just the case where the relative velocity $\mathbf{v}$ is zero, so $\gamma = 1$ (and (83) reduces to (84)). Now remembering that U^a is a 4-vector, we can make a Lorentz transformation to any other frame to find the effect of relative motion. Then $U^{a'} = \Lambda^{a'}{}_a U^a = \Lambda^{a'}{}_0$ so if the transformation takes the particular form (79) for $v = -V$, we find $U^{a'} = \gamma(V)\,(1, V, 0, 0)$ in agreement with (83); and a spatial rotation will now bring this special form to the general form (83). Thus we see that the form (83) is really a result of the nature of Lorentz transformations. Because it is an invariant, $U^{i'}U_{i'} = -1$ whatever frame we may transform to. However there is always a Lorentz transformation that will bring the 4-velocity to the form (84).

4.4.2 The 4-momentum

The 4-momentum vector P^i of a particle with non-zero rest-mass is defined from the 4-velocity (82) by

$$P^i = m_0 U^i = m_0 \frac{dx^i}{d\tau} \tag{85}$$

where m_0 is the particle's rest-mass. Hence in an orthonormal frame, by (83)

$$P^a = m_0\gamma\,(1, \mathbf{v}) = m\,(1, \mathbf{v}) =: (E, \pi) \tag{86}$$

where $m := m_0\gamma$, $E = m$ (remember we have set $c = 1$) and $\pi = m\mathbf{v}$ are respectively the relativistic mass, energy, and momentum.

In the particle's rest frame (84), $P^a = (m_0, 0, 0, 0)$. If we make a transformation from this rest frame to another one, then $P^{a'} = \Lambda^{a'}{}_a P^a = m_0\Lambda^{a'}{}_0$ so if the transformation takes the particular form (79) for $v = -V$, we find $P^{a'} = m\,(1, V, 0, 0)$ in agreement with (86). The invariant associated with the 4-momentum is

$$P^a P_a = m_0^2 U^a U_a = -m_0^2 \tag{87}$$

which will take this form whatever basis is used. In an orthonormal basis, by (86)

$$P^a P_a = -E^2 + \pi^2 = -m_0^2 \tag{88}$$

The fundamental importance of the 4-momentum is that (a) the total 4-momentum is preserved in collisions between particles; then (87) is very useful in determining the outcome; and (b) it is the basis of the force law: in flat space time and Minkowski coordinates,

$$\frac{dP^a}{d\tau} = F^a \tag{89}$$

where F^a is the 4-force. This has to be modified in curved space- times (see the section 5.1).

Finally we can consider particles with zero rest-mass, that is, particles that have a 4-momentum $P^a = (E, \pi)$ but with $m_0 = 0$. Then (85) is modified to $P^i = \mu k^i$ where $k^i = dx^i/dv$ is the tangent to the particle path $x^a(v)$, and μ is a constant. Equation (86) is modified to just give the identification $P^a = (E, \pi)$. Now (87), (88) show

$$P^a P_a = 0 \quad \Leftrightarrow \quad E^2 = \pi^2 \tag{90}$$

showing such particles (for example, photons) must always move at the speed of light; so $k^a k_a = 0$ and μ has no invariant meaning. As they move at the speed of light, we cannot perform a Lorentz transformation to bring them to rest (bringing their 4-velocity to the form (84) in an orthonormal frame): they always move at the speed of light in any frame.

4.4.3 Electromagnetism

The electromagnetic field tensor (combining both electricity and magnetism) is described by the skew tensor $F_{ij} = F_{[ij]}$, where in an orthonormal frame it is related to the electric field E_ν and magnetic field B_ν by

$$F^{ij} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{pmatrix}$$

which gives (a) the correct Lorentz force law exerted on a charge e with 4-velocity u^a, given by the expression $F^a = (e/c)F^{ab}u_b$, (b) the correct transformation properties of E_ν and B_ν when we change to a moving frame, following from the fact that F_{ab} is a tensor, and leading to identification of the two invariants $F^{ab}F_{ab} = 2(-E^2 + B^2)$ and $F_{ab}\eta^{abcd}F_{cd} = -8\mathbf{E}.\mathbf{B}$, and (c) a simple form for Maxwell's equations (discussed in the next chapter).

These properties (see *Flat and Curved Space-Times*, G. F. R. Ellis and R. M. Williams, pp. 322-336) follow directly from the tensor properties of F_{ab}. We look here at the second only.

$F^{a'b'} = \Lambda^{a'}{}_a \Lambda^{b'}{}_b F^{ab}$ where if we choose a boost in the x-direction, then $\Lambda^{0'}{}_0 = \gamma = \Lambda^{1'}_1$, $\Lambda^{1'}{}_0 = -\gamma v = \Lambda^{0'}{}_1$. Thus we find

$$\begin{aligned}
F^{0'1'} &= \Lambda^{0'}{}_a \Lambda^{1'}{}_b F^{ab} \\
&= \Lambda^{0'}{}_0 \Lambda^{1'}{}_b F^{0b} + \Lambda^{0'}{}_1 \Lambda^{1'}{}_b F^{1b} \\
&= \gamma(\Lambda^{1'}{}_1 F^{01}) - \gamma v(\Lambda^{1'}{}_0 F^{10})
\end{aligned}$$

so

$$E'_x = \gamma^2 E_x - \gamma^2 (v/c) E_x = E_x.$$

Similarly

$$\begin{aligned}
E'_y &= \Lambda^{0'}{}_a \Lambda^{2'}{}_b F^{ab} \\
&= \Lambda^{0'}{}_0 F^{02} + \Lambda^{0'}{}_1 F^{12} \\
&= \gamma(E_y - vB_z),
\end{aligned}$$

and so on. Putting these together we find

$$\begin{aligned}
E'_x = E_x, \quad & E'_y = \gamma(E_y - vB_z), \quad E'_z = \gamma(E_z + vB_y), \\
B'_x = B_x, \quad & B'_y = \gamma(B_y + \frac{v}{c^2}E_z), \quad B'_z = \gamma(B_z - \frac{v}{c^2}E_y).
\end{aligned}$$

This demonstrates the fundamental result: *magnetism is generated by motion relative to an electric field* (consider a situation where $E_a \neq 0$ but $B_a = 0$; we find $B'_a \neq 0$ in general).

An interesting question then is what invariants there are under such changes. We can find them by looking for scalars formed from F_{ab}, leading to identification of two invariants: $F_{ab}F^{ab} = 2(-E^2 + B^2)$ and $F_{ab}\eta^{abcd}F_{cd} = -8\mathbf{E}.\mathbf{B}$ (Note that we get no useful invariant from $F^a{}_a$ because it vanishes). By construction these are invariant and so are the same for all observers. This leads to the characterisation of a *plane wave* as the case when both invariants vanish: $E^2 = B^2$, $\mathbf{E}.\mathbf{B} = 0$. If this is true in one frame then it is true in all. Thus if one observer measures a field to have this plane wave form, then so do all observers.

4.4.4 Energy-momentum

The energy, momentum, and pressures of matter are described by the symmetric energy-momentum-stress tensor $T^{ab} = T^{(ab)}$, given in an orthonormal frame by

$$T^{ij} = \begin{pmatrix} \mu & q_x & q_y & q_z \\ q_x & \Pi_{xx} & \Pi_{xy} & \Pi_{xz} \\ q_y & \Pi_{yx} & \Pi_{yy} & \Pi_{yz} \\ q_z & \Pi_{zx} & \Pi_{zy} & \Pi_{zz} \end{pmatrix}$$

where μ is the relativistic energy density, q_ν is the flux of energy across a surface perpendicular to the ν-direction, or equivalently the ν-component of momentum density, and $\Pi_{\mu\nu}$ the symmetric stress tensor representing the ν component of stress, or of flux of momentum, across a surface perpendicular to the μ-direction. It is striking how all these quantities – saparate in Newtonian physics - ar eunified in one tensor when one takes the space-time viewpoint.

This tensor has vanishing divergences in Minkowski coordinates in flat space-time: that is $\frac{\partial}{\partial x^j} T^{ij} = 0$. This expresses conservation of energy ($i = 0$) and also conservation of momentum ($i = 1, 2, 3$). Generalisation of this relation to a curved space-time is considered in section 5.1.

The simplest example of a stress-tensor is that of a pressure-free matter with density ρ and 4-velocity U^i. The stress tensor is

$$T^{ab} = \rho U^a U^b.$$

In the rest-frame of the fluid, $T^{00} = \rho$ and all the other components are zero. If the fluid is moving with 3-velocity v^i relative to an orthonormal reference frame, we see immediately from the above expressions for 4-velocities that

$$\begin{aligned} T^{00} &= \rho\gamma^2 = \mu \\ T^{0i} &= \rho\gamma^2 \frac{v^i}{c} = \mu \frac{v^i}{c} \\ T^{ij} &= \rho\gamma^2 \frac{v^i}{c}\frac{v^j}{c} = \mu \frac{v^i}{c}\frac{v^j}{c} \end{aligned}$$

This form also follows from the Lorentz transformation propertis of T^{ab}, which show the relations between its components in different frames in general (cf. the cases of P^a and F_{ab} above). In the case of a cloud of particles moving with different 4-velocities, we get the same expressions as just given, but summed over all the particles, for example $T^{00} = \Sigma \rho_* \gamma_*^2 = \mu$; this is the beginning of relativistic kintic theory.

The trace of the stress tensor is the invariant

$$T \equiv T^a{}_a = -\mu + \frac{3p}{c^2}$$

where the isotropic pressure p is defined by $p \equiv \Pi^i{}_i/3$. This trace vanishes if and only if $p = (\mu/3)c^2$, the condition for isotropic radiation. Indeed this will be true for example if $T^{ab} = \Sigma \mu_* k_*^a k_*^b$ where $k_*^a k_{*a} = 0$ - a sum of contributions from zero-rest mass particles (each moving at the speed of light). There are three other independent invariants: $T^a{}_b T^b{}_a$, $T^a{}_b T^b{}_c T^c{}_a$, and $T^a{}_b T^b{}_c T^c{}_d T^d{}_a$ (equivalently, there are four independent eigenvalues for T_{ab}).

More details and specific examples of stress tensors are given in *Flat and Curved Space-Times*, G. F. R. Ellis and R. M. Williams, pp. 336-340. We discuss electromagnetism, a scalar field, and a 'perfect fluid' in Chapter 5.

4.5 Lorentz Transformations

We close this chapter with a brief discussion of the nature of families of Lorentz transformations depending on a parameter β. These are generated by skew matrices, as follows. Let the transformation be ${\Lambda_a}^b(\beta)$, where we now regard the transformation as *active*, that is as causing a 'rotation' of the space-time around fixed axes (as opposed to the passive transformations we have so far assumed, where the space-time events have remained fixed while the coordinates moved around relative to them). Its action on vectors will be: $X^a \to X'^a = {\Lambda^a}_b(\beta)X^b$, and its action on the metric will be

$$g_{ab} \to g'_{ab} = {\Lambda_a}^c(\beta)\ {\Lambda_b}^d(\beta)\ g_{cd} \tag{91}$$

The generator of the transformation is defined as

$${F_a}^b = \frac{d}{d\beta}{\Lambda_a}^b(\beta)|_{\beta=0} \tag{92}$$

which can be integrated to give

$${\Lambda_a}^b(\beta) = \exp(\beta {F_a}^b) = \delta_a^b + \beta {F_a}^b + \frac{1}{2!}\beta^2\, {}^{(2)}{F_a}^b + \frac{1}{3!}\,\beta^{3\,(3)}{F_a}^b + ... \tag{93}$$

where ${}^{(2)}{F_a}^b := {F_a}^c{F_c}^b$, ${}^{(3)}{F_a}^b = {F_a}^c{F_c}^d{F_d}^b$, etc.

Now substituting (93) into (91) and considering transformations near the identity $\beta = 0$, we keep only zero order and linear terms in β to obtain

$$g'_{ab} = ({\delta_a}^c + \beta {F_a}^c)({\delta_b}^d + \beta {F_b}^d)g_{cd} = g_{ab} + \beta({F_a}^c g_{cb} + {F_b}^d g_{ad}) + O(\beta^2)$$

Thus demanding invariance of the metric for all small β, we find

$$g'_{ab} = g_{ab} \quad \Leftrightarrow \quad F_{ab} + F_{ba} = 0 \quad \Leftrightarrow \quad F_{ab} = -F_{ba} \tag{94}$$

- the generator must be skew. This is the condition to preserve an arbitrary metric form (not only an orthonormal one)

Now we can easily find out the nature of the Lorentz transformations by choosing particular skew matrices to generate it; the specific orthonormal form of the metric enters in the raising and lowering of indices on F_{ab}. Considering the orthonormal basis of vectors (t, x, y, z) referred to above, we obtain spatial rotations in the (x, y) plane if

$$F_{ab} = x_a y_b - y_a x_b$$

so that

$$F_{ab}t^a = 0,\ F_{ab}x^b = -y_a,\ F_{ab}y^b = x_a,\ F_{ab}z^b = 0.$$

One can then easily calculate the power series (93), as follows:

$$\begin{aligned}
F_a{}^b &= x_a y^b - y_a x^b \\
{}^{(2)}F_a{}^b &= F_a{}^c(x_c y^b - y_c x^b) = -x_a x^b - y_a y^b \\
{}^{(3)}F_a{}^b &= F_a{}^c\,{}^{(2)}F_c{}^b = F_a{}^c(-x_c x^b - y_c y^b) = -x_a y^b + y_a x^b = -F_a{}^b \\
{}^{(4)}F_a{}^b &= F_a{}^c\,{}^{(3)}F_c{}^b = F_a{}^c(-F_c{}^b) = -{}^{(2)}F_a{}^b = +x_a x^b + y_a y^b \\
{}^{(5)}F_a{}^b &= F_a{}^c\,{}^{(4)}F_c{}^b = F_a{}^c(-{}^{(2)}F_c{}^b) = -{}^{(3)}F_a{}^b = F_a{}^b
\end{aligned} \tag{95}$$

and then proceeding cyclically. Hence (93) becomes

$$\Lambda_a{}^b(\beta) = \delta_a^b + F_a{}^b(\beta - \frac{1}{3!}\beta^3 + ..) + (x_a x^b + y_a y^b)(-\frac{1}{2!}\beta^2 + \frac{1}{4!}\beta^4 ...)$$

giving (on using (77))

$$\Lambda_a{}^b(\beta) = -t_a t^b + z_a z^b + F_a{}^b \sin\beta + (x_a x^b + y_a y^b)\cos\beta \tag{96}$$

This gives the results we expect for a rotation:

$$\begin{aligned}
\Lambda_a{}^b(\beta)\,t_b = t_a, &\qquad \Lambda_a{}^b(\beta)\,x_b = -y_a\sin\beta + x_a\cos\beta \\
\Lambda_a{}^b(\beta)\,z_b = z_a, &\qquad \Lambda_a{}^b(\beta)\,y_b = x_a\sin\beta + y_a\cos\beta
\end{aligned}$$

In a similar way we can calculate the velocity transformations ('boosts') generated by the timelike 2-plane $F_{ab} = t_a z_b - z_a t_b$, the calculation being identical except that some of the signs are different; and the null rotations generated by the null 2-plane $F_{ab} = k_a y_b - y_a k_b$ where k^a and n^a are the null vectors defined above (in section 3.2.3). In the latter case

$$F_a{}^b = k_a y^b - y_a k^b \Rightarrow F_a{}^b k_b = 0, F_a{}^b n_b = y_a, F_a{}^b y_b = k_a, F_a{}^b z_b = 0$$

Hence

$${}^{(2)}F_a{}^b = F_a{}^c(k_c y^b - y_c k^b) = -k_a k^b,\quad {}^{(3)}F_a{}^b = F_a{}^c\,{}^{(2)}F_c{}^b = -F_a{}^c k_c k^b = 0$$

so the Lorentz transformation (93) is the finite (quadratic) series

$$\Lambda_a{}^b(\beta) = \delta_a^b + \beta F_a{}^b - \frac{1}{2}\beta^2 k_a k^b \tag{97}$$

This has the following effect on the basis vectors:

$$\begin{aligned}
\Lambda_a{}^b(\beta)\,k_b = k_a, &\quad \Lambda_a{}^b(\beta)\,n_b = n_a + \beta y_a + \tfrac{1}{2!}\beta^2 k_a, \\
\Lambda_a{}^b(\beta)\,z_b = z_a, &\quad \Lambda_a{}^b(\beta)\,y_b = y_a + \beta k_a
\end{aligned}$$

One can explicitly check from this form that this transformation preserves the scalar products of the basis vectors. A null rotation is equivalent to a combination of spatial rotations and boosts, but is exceptionally simple when written in the above form.

These are the simplest 1-parameter groups of Lorentz transformations, generated by simple bivectors. More complex examples occur if the bivector is not simple, for example if $F_{ab} = a(x_a y_b - y_a x_b) + b(t_a z_b - z_a t_b)$ where a and b are constants.

Examples 3: Orthonormal Frames

1. (i) Consider the vector A^a and the transformation $A^{a'} = \Lambda^{a'}_a A^a$, where $\Lambda^{a'}{}_a$ is a Lorentz transformation. Show that $A.A = A'.A'$ in two different ways. Explain how to interpret this in both an active and a passive way.

(ii) Let $h_{ab} = g_{ab} + u_a u_b$ where u^a is the 4-velocity of an observer ($u^a u_a = -1$). What is h_{ab} in an appropriate orthonormal frame? Find such a frame in the case of a Robertson-Walker metric.

2. (i) Confirm that the field F_{ab} gives the correct Lorentz force law for a particle with charge e and 4-velocity u^a through the relation $F^a = (e/c)F^{ab}u_b$.

(ii) Let $F_{ab} = -2E_{[a}u_{b]} + \eta_{ab}{}^{cd}B_c u_d$. Write out these relations in an orthonormal frame. Calculate the invariants $F_{ab}F^{ab}$ and $F_{ab}\eta^{abcd}F_{cd}$ *either* from the above covariant expression *or* by using an orthonormal basis.

3. Determine the change in the quantities μ, q_i and Π_{ij} in the stress tensor T_{ab} when a change of velocity takes place in the x-direction. [Use the 'boost' form of the Lorentz transformation]. Hence (i) verify the form that pressure-free matter takes if seen from a non-comoving frame, (ii) determine the low-speed version of these transformation laws.

4. Find the Lorentz transformation generated by

$$F_{ab} = a(x_a y_b - y_a x_b) + b(t_a z_b - z_a t_b)$$

where a and b are constants. Find two real eigenvectors of this transformation [Hint: look for null vectors].

---

5 Covariant Differentiation

In this chapter we develop the apparatus needed to carry out covariant differentiation in a curved-space time. This enables us to set up differential equations for tensors, that are valid in arbitrary coordinate systems or for any tensor basis.

Why is this problematic? Analytically considered the issue is simple: the obvious derivatives of tensors are not tensors. Consider a vector field $\mathbf{X}$ with components $X^i(x^j)$ relative to a coordinate basis. Then the obvious derivatives of X^i are given by the matrix of quantities $Y^i{}_j = \frac{\partial X^i}{\partial x^j} = X^i{}_{,j}$. However these are not the components of a tensor, under a general transformation of coordinates. For changing to new coordinates, with $A_{j'}{}^j = \frac{\partial x^j}{\partial x^{j'}}$ (see (20-23)), we find

$$\begin{aligned} Y^{i'}{}_{j'} &= \frac{\partial}{\partial x^{j'}} X^{i'} = \frac{\partial x^j}{\partial x^{j'}} \frac{\partial}{\partial x^j} (A^{i'}{}_i X^i) \\ &= A_{j'}{}^j (\frac{\partial}{\partial x^j} A^{i'}{}_i X^i + A^{i'}{}_i \frac{\partial}{\partial x^j} X^i) = A_{j'}{}^j \frac{\partial^2 x^{i'}}{\partial x^j \partial x^i} X^i + A_{j'}{}^j A^{i'}{}_i Y^i{}_j \end{aligned}$$

The second term is what we would have if $Y^i{}_j$ were a tensor; the first term however is non-zero in general. Thus we do not have a tensor transformation unless $\frac{\partial^2 x^{j'}}{\partial x^i \partial x^j} = 0$, a very restricted class of coordinate changes.

We will consider the geometric reason for this later. For the present the problem is to find the tensorial quantity that corresponds to $Y^i{}_j$.

5.1 Parallel transport

It is convenient to approach the issue by considering the derivative $\mathbf{Z}$ of a vector field $\mathbf{Y}$ along a curve $\gamma(v)$ with tangent vector $X^i = dx^i/dv$. We shall do so first in flat space, and then generalise our results to a curved space and to space-time. We will introduce two different notations (both in common use) for this vector: we write $\mathbf{Z} = \nabla_{\mathbf{X}}(\mathbf{Y})$ if we wish to emphasize that it is the covariant (i.e. tensorial) derivative in the direction $\mathbf{X}$; or we write $Z^i = \delta Y^i/\delta v$ if we wish to emphasize it is the covariant derivative relative to the curve parameter v.

5.1.1 Flat space

Using standard Cartesian coordinates, the rate of change of Y^i along $\gamma(v)$ is dY^i/dv; this vanishes if and only if $\mathbf{Y}$ is constant along the curve. However, by essentially the same calculation as just given, this is not a tensorial quantity; it does not transform as a tensor when we change to a general coordinate system.

We now define $\delta Y^i/\delta v$ to be that tensorial quantity which is equal to dY^i/dv in a Cartesian frame. Changing from Cartesian coordinates x^i to general coordinates $x^{i'}$, we find

$$\begin{aligned}
\frac{\delta Y^{i'}}{\delta v} &= A^{i'}{}_i \frac{\delta Y^i}{\delta v} = A^{i'}{}_i \frac{dY^i}{dv} = A^{i'}{}_i \frac{\partial Y^i}{\partial x^{j'}} X^{j'} \\
&= A^{i'}{}_i \frac{\partial}{\partial x^{j'}} (A^i{}_{k'} Y^{k'})\, X^{j'} = A^{i'}{}_i (\frac{\partial A^i{}_{k'}}{\partial x^{j'}} Y^{k'} + A^i{}_{k'} \frac{\partial}{\partial x^{j'}} Y^{k'})\, X^{j'} \\
&= A^{i'}{}_i \frac{\partial^2 x^i}{\partial x^{j'} \partial x^{k'}} Y^{k'} X^{j'} + A^{i'}{}_i A^i{}_{k'} dY^{k'}/dv \\
&= \Gamma^{i'}{}_{j'k'} Y^{k'} X^{j'} + \frac{dY^{i'}}{dv}
\end{aligned}$$

where $\Gamma^{i'}{}_{j'k'} := A^{i'}{}_i \frac{\partial^2 x^i}{\partial x^{j'} \partial x^{k'}} = \Gamma^{i'}{}_{k'j'}$ are the Christoffel symbols (in a general frame they are called connection coefficients or connectiuon components). This gives the expression for $\frac{\delta Y^{i'}}{\delta v}$ in any coordinates whatever. Thus let us now drop the primes, to obtain the general expression for this quantity in flat space-time in any coordinates:

$$\frac{\delta Y^i}{\delta v} = \frac{dY^i}{dv} + \Gamma^i{}_{jk} Y^k X^j, \quad \frac{dY^i}{dv} := Y^i{}_{,j} X^j \tag{98}$$

where dY^i/dv is the rate of change of the components of Y^i along the curve and $\Gamma^i{}_{jk} = \Gamma^i{}_{kj}$. Because (from its derivation) this vector is equal to dY^i/dv in a Cartesian system, we can interpret (98) in the following way: this gives the components of the covariant rate of change of Y^i along the curve with tangent vector X^j, i.e.

$$(\nabla_{\mathbf{X}}(\mathbf{Y}))^i = \frac{\delta Y^i}{\delta v} \quad \text{when} \quad X^i = \frac{dx^i}{dv} \tag{99}$$

which vanishes if and only if Y^i is parallel along the curve γ. Combining (98), (99) we obtain the expression

$$(\nabla_{\mathbf{X}}(\mathbf{Y}))^i = (Y^i{}_{,j} + \Gamma^i{}_{jk} Y^k) X^j \tag{100}$$

Why do we need the extra terms in (98)? This expression takes the form it does because we have of necessity to express $\mathbf{Y}$ in terms of a basis $\mathbf{e}_i$: $\mathbf{Y} = Y^i \mathbf{e}_i$; and the rate of change of $\mathbf{Y}$ along the curve has two parts, one given by the rate of change of components Y^i relative to the basis $\mathbf{e}_i$ (the first term), and the second given by the rate of change of the basis vectors themselves along the curve (the second term). We can make this explicit after establishing the following three basic properties of the covariant derivative operator (99): for any constants α, β, functions f, g and vectors $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$,

$$\begin{aligned}
\nabla_{f\mathbf{Y}+g\mathbf{Z}}(\mathbf{X}) &= f\nabla_{\mathbf{Y}}(\mathbf{X}) + g\nabla_{\mathbf{Z}}(\mathbf{X}) & (101)\\
\nabla_{\mathbf{Z}}(\alpha\mathbf{X} + \beta\mathbf{Y}) &= \alpha\nabla_{\mathbf{Z}}(\mathbf{X}) + \beta\nabla_{\mathbf{Z}}(\mathbf{Y}) & (102)\\
\nabla_{\mathbf{X}}(f\mathbf{Y}) &= X(f)\mathbf{Y} + f\nabla_{\mathbf{X}}(\mathbf{Y}) & (103)
\end{aligned}$$

We prove the last: from (100), because $(fY)^i = fY^i$,

$$(\nabla_{\mathbf{X}}(f\,Y))^i = ((f\,Y)^i{}_{,j} + \Gamma^i{}_{jk} f\,Y^k)X^j =$$
$$= (f_{,j}Y^i + f\,(Y)^i{}_{,j} + \Gamma^i{}_{jk} f\,Y^k)X^j = X(f)Y^i + f\,(\nabla_{\mathbf{X}}(Y))^i$$

as required. □

Now choose a basis $\mathbf{e}_i$, and write $\mathbf{X} = X^j\mathbf{e}_j$, $\mathbf{Y} = Y^i\mathbf{e}_i$. Then

$$\begin{aligned} \nabla_{\mathbf{X}}(\mathbf{Y}) &= \nabla_{\mathbf{X}}(Y^i\mathbf{e}_i) = X(Y^i)\mathbf{e}_i + Y^i\nabla_{\mathbf{X}}(\mathbf{e}_i) = \\ &= Y^i{}_{,j}X^j\mathbf{e}_i + Y^i\nabla_{X^j\mathbf{e}_j}(\mathbf{e}_i) = Y^i{}_{,j}X^j\mathbf{e}_i + Y^iX^j\nabla_{\mathbf{e}_i}(\mathbf{e}_j) = \\ &= (Y^i{}_{,j} + Y^iZ^i))X^j\mathbf{e}_i, \quad \text{where} \quad Z^i := (\nabla_{\mathbf{e}_j}(\mathbf{e}_k))^i \end{aligned}$$

Thus we recover (100), with the two important features. Firstly, this derivation works for any basis at all, not only a coordinate basis. Thus (100) is generally valid. However for a general basis there is no reason to believe that the quantities $\Gamma^i{}_{jk}$ (generically, these can be called the connection components) will be symmetric in j and k, and indeed (as we see below) in general they are not so. Secondly, we obtain the important identification

$$\Gamma^i{}_{kj} = (\nabla_{\mathbf{e}_k})(\mathbf{e}_j)^i \Leftrightarrow \nabla_{\mathbf{e}_k}(\mathbf{e}_j) = \Gamma^i{}_{kj}\mathbf{e}_i\,, \tag{104}$$

that is $\Gamma^i{}_{kj}$ is the i-component of the k-derivative of the j-basis vector (which can also be obtained from (100) on setting $\mathbf{X} = \mathbf{e}_k \Leftrightarrow X^j = \delta^j_k$, $\mathbf{Y} = \mathbf{e}_j \Leftrightarrow Y^i = \delta^i_j$). This gives us an immediate interpretation of the connection components in any specific case. This identification together with (101-103) shows how these quantities transform under a change of basis (they are non-tensorial in the third index).

As an example, consider the 2-dimensional Euclidean plane. In Cartesian coordinates, let $\mathbf{e}_x$, $\mathbf{e}_y$ by the Cartesian basis vectors (i.e. $\mathbf{e}_x = \frac{\partial}{\partial x}$, $\mathbf{e}_y = \frac{\partial}{\partial y}$). By their nature they are all parallel along each other; interpretation (104) shows that $\Gamma^i{}_{jk} = 0$ for this basis. Now consider this plane in polar coordinates, and denote the coordinate basis $\mathbf{e}_1 = \frac{\partial}{\partial r}$, $\mathbf{e}_2 = \frac{\partial}{\partial \theta}$ (see example after (23)). Geometrically it is clear that

$$\nabla_{\mathbf{e}_1}(\mathbf{e}_1) = 0, \quad \nabla_{\mathbf{e}_1}(\mathbf{e}_2) = a\mathbf{e}_2, \quad \nabla_{\mathbf{e}_2}(\mathbf{e}_1) = b\mathbf{e}_2, \quad \nabla_{\mathbf{e}_2}(\mathbf{e}_2) = c\mathbf{e}_1,$$

for some functions a, b, c. We can verify this and find these functions by exactly the same calculation as lead to (104). Using the methods following (23), we find

$$\mathbf{e}_1 = \cos\theta\mathbf{e}_x + \sin\theta\mathbf{e}_y, \quad \mathbf{e}_2 = -r\sin\theta\mathbf{e}_x + r\cos\theta\mathbf{e}_y,$$

so for example

$$\nabla_{\mathbf{e}_1}(\mathbf{e}_2) = \nabla_{\mathbf{e}_1}(-r\sin\theta\mathbf{e}_x + r\cos\theta\mathbf{e}_y) =$$

$$= -e_1(r\sin\theta)\mathbf{e}_x + e_1(r\cos\theta)\mathbf{e}_y - r\sin\theta\nabla_{\mathbf{e}_1}\mathbf{e}_x + r\cos\theta\nabla_{\mathbf{e}_1}\mathbf{e}_y$$

Now the last two terms both vanish, for example for the last we obtain

$$\nabla_{\mathbf{e}_1}\mathbf{e}_y = \nabla_{\cos\theta\mathbf{e}_x+\sin\theta\mathbf{e}_y}(\mathbf{e}_y) =$$

$$= \cos\theta\nabla_{\mathbf{e}_x}(\mathbf{e}_x) + \sin\theta\nabla_{\mathbf{e}_y}(\mathbf{e}_y) = 0$$

from the properties of the Euclidean basis. Thus we find

$$\nabla_{\mathbf{e}_1}(\mathbf{e}_2) = -\sin\theta\mathbf{e}_x + \cos\theta\mathbf{e}_y = \frac{1}{r}\mathbf{e}_2$$

as expected, showing $\Gamma^1{}_{12} = 0$, $\Gamma^2{}_{12} = 1/r = a$. Similarly we can find

$$\nabla_{\mathbf{e}_1}(\mathbf{e}_1) = 0, \quad \nabla_{\mathbf{e}_2}(\mathbf{e}_1) = \frac{1}{r}\mathbf{e}_2, \quad \nabla_{\mathbf{e}_1}(\mathbf{e}_2) = -r\mathbf{e}_1.$$

in agreement with their geometric interpretation.

We can see from this geometric interpretation that if the basis vectors $\mathbf{e}_a$ are parallel along the curve γ, then the connection components give no contribution to the expression for the covariant derivative. In detail, in this case by (101)

$$\nabla_{\mathbf{X}}(\mathbf{e}_a) = 0 \quad \Leftrightarrow \quad \nabla_{X^b\mathbf{e}_b}(\mathbf{e}_a) = 0 \quad \Leftrightarrow \quad X^b\nabla_{\mathbf{e}_b}(\mathbf{e}_a) = 0$$

$$\Leftrightarrow X^b(\nabla_{\mathbf{e}_b}(\mathbf{e}_a))^c = 0 \quad \Leftrightarrow \quad X^b\Gamma^c{}_{ba} = 0 \quad \Leftrightarrow \quad \frac{\delta Y^i}{\delta v} = \frac{dY^i}{dv} \qquad (105)$$

for all vectors Y^i (using (3)). That is, in this case (and only in this case) the covariant and apparent derivatives are the same along the curve.

5.1.2 The Parallel Transport Operator

Finally, it is sometimes useful to explicitly introduce the *Parallel Transport Operator* τ along a curve $\gamma(v)$: if a vector field $\mathbf{Y}$ is parallel transported a parameter distance v from P to Q along a curve $x^i(v)$ with tangent vector $X^i = dx^i/dv$, we can write $\tau_v(\mathbf{X})$ for the parallel transported vector at Q. Then the covariant derivative is just the derivative operator defined by parallel transport: i.e.

$$(\nabla_{\mathbf{X}}(Y))_Q = \frac{d}{dv}(\tau_v(\mathbf{Y}_P) - Y_Q)|_{v=0} = \lim_{v\to 0}\frac{1}{v}(\tau_v(\mathbf{Y}_P) - \mathbf{Y}_Q)$$

so that in particular $\nabla_{\mathbf{X}}(Y) = 0$ everywhere $\Leftrightarrow \tau_v(\mathbf{Y}) = \mathbf{Y}$ for all P, Q. The parallel transfer operator is just the solution to this equation.

5.2 Curved space

The idea now is that we take over the essential structure above unchanged to curved spaces. That is, covariant derivative of vectors along curves will be governed by (98)-(104). In particular we characterise *parallel transfer* of a vector $\mathbf{Y}$ along a curve with tangent vector $\mathbf{X}$ by

$$(\nabla_{\mathbf{X}}(Y))^i = \frac{\delta Y^i}{\delta v} = 0 \tag{106}$$

What we still have to discover is the rule determining the values of the connection coefficients; this is considered below. However consideration of the simple example of parallel transfer of vectors along the great circles on a sphere will suffice to show two features of importance.

First, in general parallel transfer of vectors along curves is not integrable: i.e. parallel transfer from P to Q along two different curves γ_1, γ_2 will in general give different results (see section 5.7 of *Flat and Curved Space-Times* by G F R Ellis and R M Williams). Thus the concept of 'parallel at a distance' (independent of the curve chosen) is only meaningful in a flat space. This is considered further in the chapter below on curvature.

Second, in any space we can ask, what is the nature of a curve whose direction is unchanging? These are clearly very special curves: they are the generalisation of a straight line to curved spaces. The rule is clear: such a curve, called a *geodesic*, is a curve whose tangent vector (the direction of the curve) is parallel propagated along the curve:

$$(\nabla_{\mathbf{X}}(X))^i = \frac{\delta X^i}{\delta v} = 0 \tag{107}$$

We will explore this more below. For the moment the point is simply to comment that in a curved space, the distance between geodesics that are initially parallel does not usually stay constant; indeed they are able to intersect each other (as shown by the example of the great circles on the surface of a sphere).

5.2.1 Curved space-time

The geometrical structure of parallel transfer in curved space-times is the same as that in curved spaces. The main point is the physical interpretation of parallel transfer and of geodesics. This will be developed below, but for the moment we may note the following:

1: an ordinary body in free fall (i.e. moving under only gravity and inertia) moves on a timelike geodesic, that is, a curve obeying (9) with $X^i X_i < 0$.

2: A zero-rest mass particle in free motion moves on a null geodesic that is, a curve obeying (9) with $X^i X_i = 0$.

3: A set of parallel propagated axes along a timelike geodesic correspond to a physically non-rotating basis of vectors for an observer moving on the geodesic. These could be realised for example by a set of ideal gyroscopes that point in a physically unchanging direction as time advances.

Chapter 5 of *Flat and Curved Space-Times* by G F R Ellis and R M Williams discusses these issues further.

5.3 Covariant Differentiation

We develop covariant differentiation first for vector fields,and then for arbitrary tensor fields.

5.3.1 Vector fields

We can use the linearity in terms of X^i (equation (101)) to write (100) in the form

$$\frac{\delta Y^i}{\delta v} = (\nabla_{\mathbf{X}}(Y))^i = Y^i{}_{;j} X^j \tag{108}$$

where

$$Y^i{}_{;j} := Y^i{}_{,j} + \Gamma^i{}_{jk} Y^k \tag{109}$$

is the covariant derivative of the vector field Y^i. This states that the covariant directional derivative of a vector field Y^i in the arbitrary direction X^j is given by contracting X^j with the tensor field $Y^i{}_{;j}$, given by (109). This is the tensor we have been looking for, generalising $Y^i{}_{,j}$ to arbitrary coordinates. We can see it is a tensor by the tensor detection theorem (the left hand side of (108) is a tensor for arbitrary vectors X^i). From (108) we can see the geometrical meaning of $Y^i{}_{;j}$: it is the covariant derivative of Y^i in the direction of the basis vector $\mathbf{e}_j$ (to see this, set $X^i = \delta^i_j$ in (108)).

5.3.2 General Tensors

The aim now is to extend the ideas of parallel transport and the covariant derivative to arbitrary tensors. The way this is done, is by demanding that the tensor operations defined in the previous chapter are preserved under parallel transfer. Then for an arbitrary tensor $\mathbf{T}$, say with components $T^{ab}{}_{cd}$, we define

$$\frac{\delta}{\delta v} T^{ab}{}_{cd} = (\nabla_{\mathbf{X}}(T))^{ab}{}_{cd} = T^{ab}{}_{cd;j} X^j \tag{110}$$

to be the covariant derivative of T^{ab}_{cd} in the direction X^a, that is, the derivative related to parallel transfer of $\mathbf{T}$ along the integral curves of X^a (so that $\frac{\delta}{\delta v} T^{ab}{}_{cd} = 0$ on a particular such curve γ iff $\mathbf{T}$ is parallel transported along γ). The problem is to determine explicit expressions for $T^{ab}{}_{cd;j}$

These follow from the properties of the covariant derivative that arise from the above characterisation of parallel transport.

1: *Linearity*: Parallel transport preserves linear combinations of tensors, consequently the covariant derivative is linear (as usual: $D(\alpha f + \beta g) = \alpha Df + \beta Dg$.) For example, if $T^{ij}{}_c$ and $R^{ij}{}_c$ are tensors and α, β numbers, then

$$(\alpha T^{ij}{}_c + \beta R^{ij}{}_c)_{;d} = \alpha T^{ij}{}_{c;d} + \beta R^{ij}{}_{c;d}$$

2: *Leibniz*: Parallel transport preserves tensor products, consequently the standard Leibniz derivation rule ($D(fg) = Df.g + f.Dg$) applies. For example,

$$(R^{ij}{}_c S^{ef}{}_{gh})_{;s} = R^{ij}{}_{c;s} S^{ef}{}_{gh} + R^{ij}{}_c S^{ef}{}_{gh;s}$$

3: *Commutes with contraction*: we get the same result by taking the derivative and then contracting, or contracting and then taking the derivative. For example, if $R^{ij}{}_{cd} \to T^j{}_d = R^{ij}{}_{id}$, then $R^{ij}{}_{cd;e} \to R^{ij}{}_{id;e} = (T^j{}_d)_{;e}$.

4: *Covariant derivative of functions and vectors are as expected*: for any function f,

$$f_{;a} = f_{,a} \quad \Leftrightarrow \quad \nabla_{\mathbf{X}}(f) = X(f) = f_{,a} X^a \tag{111}$$

while for any vector, the covariant derivative is given by (109).

From these properties, we can find the derivative of any tensor. For example, consider a 1-form (covariant vector) W_a. To obtain its covariant derivative, form the scalar $X^a W_a$ with an arbitrary (contravariant) vector field X^a. Then

$$(W_a X^a)_{,b} = (W_a X^a)_{;b} = W_{a;b} X^a + W_a X^a{}_{;b}$$

(the first equality by (14), the second by Leibniz), and so by (12)

$$W_{a;b} X^a = (W_a X^a)_{,b} - W_a (X^a{}_{,b} + \Gamma^a{}_{bc} X^c)$$

holds for all vectors X^a. Choose $X^a = \delta^a_k$ to find

$$W_{a;b} \delta^a_k = (W_a \delta^a_k)_{,b} - W_a ((\delta^a_k)_{,b} + \Gamma^a{}_{bc} \delta^c_k)$$

that is

$$W_{a;k} = W_{a,k} - W_b \Gamma^b{}_{ka} \tag{112}$$

Similarly for a tensor $T_a{}^b$, form the scalar $T_a{}^b X^a W_b$ with arbitrary vectors X^a,W_b. Then

$$(T_a{}^b X^a W_b)_{,c} = (T_a{}^b X^a W_b)_{;c} = T_a{}^b{}_{;c} X^a W_b + T_a{}^b X^a{}_{;c} W_b + T_a{}^b X^a W_{b;c}$$

and so by (12), (15)

$$T_a{}^b{}_{;c} X^a W_b = (T_a{}^b X^a W_b)_{,c} - T_a{}^b (X^a{}_{,c} + \Gamma^a{}_{cd} X^d) W_b - T_a{}^b X^a (W_{b,c} - W_d \Gamma^d{}_{cb})$$

holds for all vectors X^a, W_b. Choose $X^a = \delta^a_k$, $W_b = \delta^m_b$ to find

$$T_a{}^{b;c}\delta^a_k\delta^m_b = (T_a{}^b\delta^a_k\delta^m_b)_{,c} - T_a{}^b((\delta^a_k)_{,c} + \Gamma^a{}_{cd}\delta^d_k)\delta^m_b - T_a{}^b\delta^a_k((\delta^m_b)_{,c} - \delta^m_d\Gamma^d{}_{cb})$$

that is

$$T_k{}^m{}_{;c} = T_k{}^m{}_{,c} - T_a{}^m\Gamma^a{}_{ck} + T_k{}^b\Gamma^m{}_{cb} \tag{113}$$

As a specific example: consider the unit tensor $I^a{}_b$ with components given by the Kronecker delta ($I^a{}_b = \delta^a_b$): $I_k{}^m{}_{;c} = (\delta_k{}^m)_{;c} = (\delta_k{}^m)_{,c} - \delta_a{}^m\Gamma^a{}_{ck} + \delta_k{}^b\Gamma^m{}_{cb} = 0 - \Gamma^m{}_{ck} + \Gamma^m{}_{ck}$, so

$$\delta_k{}^m{}_{;c} = 0 \tag{114}$$

On considering the examples (109), (112), (115), the general rule is clear. As in (109), we correct each index in turn for the effects of change of the basis vectors on that index, by 'summing it out' onto a Christoffel symbol, with a '+' if the index is up and a '-' if it is down. In each case the index to be corrected appears on the Christoffel symbol, its previous position being summed onto that symbol, always summing one index up with one down. As the middle index on these symbols is always the derivative position, the placing of indices is unique. For example,

$$T^{ab}{}_{cd;e} = T^{ab}{}_{cd,e} + T^{sb}{}_{cd}\Gamma^a{}_{es} + T^{as}{}_{cd}\Gamma^b{}_{es} - T^{ab}{}_{sd}\Gamma^s{}_{ec} - T^{ab}{}_{cs}\Gamma^s{}_{ed}$$

and

$$T_{cd;e} = T_{cd,e} - T_{sd}\Gamma^s{}_{ec} - T_{cs}\Gamma^s{}_{ed} \tag{115}$$

The proof in each case is as above: contract the tensor to form a scalar, then use Leibniz' rule and the known forms (109), (111), (112). Finally it is important to notice the following: because parallel transfer preserves tensor relations, it will preserve tensor symmetries also. Thus *the covariant derivative of a tensor will have the same symmetries as the tensor.* For example,

$$T_{ab} = T_{(ab)} \quad \Rightarrow \quad T_{ab;c} = T_{(ab);c}.$$

This may be seen from (115): interchanging (c, d) and subtracting,

$$(T_{cd;e} - T_{dc;e}) = (T_{cd,e} - T_{dc,e}) - (T_{sd}\Gamma^s{}_{ec} - T_{sc}\Gamma^s{}_{ed}) - (T_{cs}\Gamma^s{}_{ed} - T_{ds}\Gamma^s{}_{ec}) =$$

$$= (T_{cd} - T_{dc})_{,e} - \Gamma^s{}_{ec}(T_{sd} - T_{ds}) - \Gamma^s{}_{ed}(T_{sc} - T_{cs})$$

which clearly vanishes if T_{ab} is symmetric.

5.4 Relation to the metric

So far, the metric of the space and the parallel transfer structure (determined by the $\Gamma^a{}_{bc}$) have been treated as independent. However they determine each other in Euclidean space, and also in the Riemannian geometry that is the basis of General Relativity. This relation is determined by two new relations we now impose.

5.4.1 The derivative of the metric

Firstly, we assume the metric tensor is invariant under parallel transport along any curve:

$$\forall X^a : \quad \frac{\delta g_{cd}}{\delta v} = g_{cd;e}X^e = (g_{cd,e} - g_{sd}\Gamma^s{}_{ec} - g_{cs}\Gamma^s{}_{ed})X^e = 0 \tag{116}$$

(using (115)). This is true if and only if

$$g_{cd;e} = 0 \quad \Leftrightarrow \quad g_{cd,e} = \Gamma_{dec} + \Gamma_{ced} \tag{117}$$

where for convenience we allow raising and lowering on the first index of the $\Gamma^s{}_{ec}$ (but not the other ones). This is true in Euclidean space and in flat space-time.

The geometric meaning of this assumption is that magnitudes and scalar products are preserved under parallel transfer. Consider for example two vectors Y^i and Z^j that are parallel transferred along a curve $\gamma(v)$ with tangent vector X^k: then $\frac{\delta Y^i}{\delta v} = Y^i{}_{;j}X^j = 0$, $\frac{\delta Z^i}{\delta v} = Z^i{}_{;j}X^j = 0$. Then

$$\frac{\delta(Y.Z)}{\delta v} = (Y^i g_{ij} Z^j)_{;k}X^k =$$

$$(Y^i{}_{;k}X^k)g_{ij}Z^j + Y^i(g_{ij;k}X^k)Z^j + Y^i g_{ij}(Z^j{}_{;k}X^k)$$
$$= Y^i(g_{ij;k}X^k)Z^j$$

which vanishes for all X^i, Y^j, and Z^k if and only if (117) holds. As a particular application, we may note that the magnitude of a geodesic vector is always constant along itself:

$$X^a{}_{;b}X^b = 0 \quad \Rightarrow \quad \frac{\delta(\mathbf{X}.\mathbf{X})}{\delta v} = (X^a g_{ab} X^b)_{;c}X^c = 0 \tag{118}$$

We can consider spaces where this is not true; they form generalisations of Riemannian geometry, in which scalar products and magnitudes are not preserved under parallel transfer, in general. Thus these (Weyl) geometries are more complex than those of General Relativity. The resulting geometrical relations are given in the book *Ricci Calculus* by J A Schouten (Springer, 1954).

When (117) is true, as we will assume from now on, it has two important consequences. Firstly, take the derivative of the definition (33) to find

$$g^{ij}{}_{;s}g_{jk} + g^{ij}g_{jk;s} = \delta^i_{k;s} = 0 \quad \Rightarrow \quad g^{ij}{}_{;s}g_{jk} = 0$$

by (114) and (115). Multiply by g^{km} to find

$$g^{im}{}_{;s} = 0 \tag{119}$$

Equations (117), (119) have the very useful consequence that covariant differentiation commutes with raising and lowering of indices, so that tensor indices can be raised or lowered at will even if covariant derivatives occur. For example,

$$X_{a;k} = (X^b g_{ab})_{;k} = X^b{}_{;k} g_{ab} + X^b g_{ab;k} = X^b{}_{;k} g_{ab}$$

by (117), and so

$$X^b{}_{;k} = 0 \Leftrightarrow X_{a;k} = 0$$

(In the case of the metric tensor, for example, we have (117), (114), (119)). Again,

$$A^i B_{i;k} = A^i (B^j g_{ij})_{;k} = A^i (B^j{}_{;k} g_{ij} + B^j g_{ij;k}) = A^i g_{ij} B^j{}_{;k}$$

and so

$$A^i B_{i;k} = A_j B^j{}_{;k} \tag{120}$$

This implies in particular that

$$(X^i X_i)_{;k} = X^i{}_{;k} X_i + X^i X_{i;k} = 2X^i{}_{;k} X_i = 2X^j X_{j;k} \tag{121}$$

Secondly, we can determine the derivative of η^{abcd} as follows: from the product formula (49)

$$(\eta^{abcd}\ \eta_{ijkl})_{;e} = -4!(\delta^a_{[i}\delta^b_j\delta^c_k\delta^d_{l]})_{;e} = 0$$

by Leibniz on the right and (114). Using Leibniz on the left

$$\eta^{abcd}{}_{;e}\ \eta_{ijkl} + \eta^{abcd}\ \eta_{ijkl;e} = 0$$

However as in (121), (53) shows $\eta_{abcd}\ \eta^{abcd} = 4! \quad \Rightarrow \quad \eta_{abcd}\ \eta^{abcd}{}_{;e} = 0$. Thus multiplying the previous equation by η_{abcd}, we find

$$\eta_{ijkl;e} = 0 \quad \Leftrightarrow \quad \eta^{ijkl}{}_{;e} = 0 \tag{122}$$

This is as we would expect: the volume element is conserved under parallel transport, because this is true of the metric. It gives us an interesting formula as follows: write out the first of (122) explicitly (remembering (44)). Then

$$\left(\frac{1}{\sqrt{|g|}}\epsilon^{abcd}\right)_{,e} + \eta^{sbcd}\Gamma^a{}_{es} + \eta^{ascd}\Gamma^b{}_{es} + \eta^{absd}\Gamma^c{}_{es} + \eta^{abcs}\Gamma^d{}_{es} = 0$$

The first term is $-\frac{1}{2}\frac{|g|_{,e}}{|g|}\ \eta^{abcd}$ and the other terms are totally skew in (a, b, c, d). Thus multiplying by η_{abcd}, we find $-4!(\log\sqrt{|g|})_{,e} = 4\eta_{abcd}\ \eta^{sbcd}\Gamma^a{}_{es}$. Using (52) this gives

$$\left(\log\left(\sqrt{|g|}\right)\right)_{,e} = \Gamma^s{}_{es} \tag{123}$$

which is very useful in calculating divergences.

5.4.2 Vanishing torsion

The second basic assumption we make is that there is no torsion. This is expressed in the covariant equation

$$\forall\ f(x^i):\quad f_{;ab}=f_{;ba},\quad f_{;ab}:=(f_{;a})_{;b} \tag{124}$$

(covariant second derivatives of a function commute). From (12) and (14), $f_{;ab}=(f_{,a})_{,b}-f_{,c}\Gamma^c{}_{ba}$ so

$$f_{;ab}-f_{;ba}=f_{,ab}-f_{,ba}+f_{,c}(\Gamma^c{}_{ba}-\Gamma^c{}_{ab})=0 \tag{125}$$

where by definition $f_{,a}=e_a(f)$, as on page 36 below (71). Again, it is possible to find spaces where this is not true (for example in the Einstein-Cartan-Sciama-Kibble theories); however it is true in flat space, and in Riemannian geometry we adopt it also (so it holds in General Relativity).

We will work out what this means first in a coordinate basis, and then generalise to a general basis. In a coordinate basis,

$$\forall f(x^i):\quad f_{,ab}=f_{,ba}\quad\Rightarrow\quad \Gamma^c{}_{ba}=\Gamma^c{}_{ab} \tag{126}$$

That is, the $\Gamma^c{}_{ab}$, called Christoffel symbols in this case, are symmetric. Hence (still using a coordinate basis), we find from (12) that

$$Y^i{}_{;j}X^j-X^i{}_{;j}Y^j=Y^i{}_{,j}X^j+Y^s\Gamma^i{}_{js}X^j-X^i{}_{,j}Y^j-X^s\Gamma^i{}_{js}Y^j=Y^i{}_{,j}X^j-X^i{}_{,j}Y^j=[X,Y]^i$$

because of (126). Thus from (73), for every vector field $\mathbf{X}$, $\mathbf{Y}$,

$$[X,Y]^i=Y^i{}_{;j}X^j-X^i{}_{;j}Y^j\quad\Leftrightarrow\quad[\mathbf{X},\mathbf{Y}]=\nabla_{\mathbf{X}}(\mathbf{Y})-\nabla_{\mathbf{Y}}(\mathbf{X}) \tag{127}$$

Applying this now to a general basis $\mathbf{e}_a$, we choose $\mathbf{X}=\mathbf{e}_a$, $\mathbf{Y}=\mathbf{e}_b$ and using (104) we obtain

$$[\mathbf{e}_a,\mathbf{e}_b]=\nabla_{\mathbf{e}_a}(\mathbf{e}_b)-\nabla_{\mathbf{e}_b}(\mathbf{e}_a)=\gamma^c{}_{ab}\mathbf{e}_c\quad\Leftrightarrow\quad\gamma^c{}_{ab}=\Gamma^c{}_{ab}-\Gamma^c{}_{ba} \tag{128}$$

whih defines the *commutation coefficients* $\gamma^c{}_{ab}$ of the basis vectors (see after (3.19) for a specific example), which vanish if and only if the basis is a coordinate basis. Indeed

$$f_{;ab}-f_{;ba}=0\forall f\quad\Leftrightarrow\quad\forall f:\ f_{,ba}-f_{,ab}=\gamma^c{}_{ab}f_{,c} \tag{129}$$

This will for example apply in particular when f is chosen to be any of the components X^a of a vector field; thus $X^f{}_{,bc}-X^f{}_{,cb}=X^f{}_{,a}\gamma^a{}_{cb}$.

Writing this basis in terms of local coordinates: $e_a=e_a{}^i(x^j)\,\frac{\partial}{\partial x^i}$, implying $e_a(f)=e_a{}^i\frac{\partial f}{\partial x^i}$, and with components of the dual basis $e^a{}_i$ given by $e_a{}^je^b{}_j=\delta^a_b$ (see (69)), we find from (73) the explicit expressions

$$\gamma^c{}_{ab}=-(e_a{}^i{}_{,j}e_b{}^j-e_b{}^i{}_{,j}e_a{}^j)e^c{}_i=(e^c{}_{i,j}-e^c_{j,i})e_a{}^ie_b{}^j \tag{130}$$

for the commutation coefficients (notice that the derivatives here are partial derivatives with respect to the coordinates chosen, because the indices i, j are coordinate indices; it is the indices a, b, c that belong to the general basis). Equations (128), (129) generalise (125,126) to an arbitrary basis.

5.4.3 The Christoffel relations

To calculate the Γ-symbols, we write down (117) three times, with the indices permuted cyclically: $(c, d, e) \to (e, c, d) \to (d, e, c)$:

$$\begin{aligned} g_{cd;e} = 0 \;&\Leftrightarrow\; g_{cd,e} = \Gamma_{dec} + \Gamma_{ced} \\ g_{ec;d} = 0 \;&\Leftrightarrow\; g_{ec,d} = \Gamma_{cde} + \Gamma_{edc} \\ g_{de;c} = 0 \;&\Leftrightarrow\; g_{de,c} = \Gamma_{ecd} + \Gamma_{dce} \end{aligned}$$

Now we add the first two and subtract the third, to get

$$g_{cd,e} + g_{ec,d} - g_{de,c} = \Gamma_{dec} + \Gamma_{ced} + \Gamma_{cde} + \Gamma_{edc} - \Gamma_{ecd} - \Gamma_{dce}$$

We now combine the 1st and 6th, 2nd and 3rd, 4th and 5th terms on the right, using (128) in the form with first indices lowered:

$$\gamma_{cab} = \Gamma_{cab} - \Gamma_{cba}$$

(we can lower the first index on $\gamma^c{}_{ab}$ also without problems). We find

$$\Gamma_{ced} = \frac{1}{2}\{g_{cd,e} + g_{ec,d} - g_{de,c}\} + \frac{1}{2}\{\gamma_{edc} + \gamma_{dec} - \gamma_{ced}\} \tag{131}$$

which determines $\Gamma^a{}_{ed}$ by

$$\Gamma^a{}_{ed} = g^{ac}\Gamma_{ced} \tag{132}$$

This set of equations, the (generalised) *Christoffel relations*, determine the connection (that is, the $\Gamma^a{}_{bc}$) from the metric components (given the choice of basis vectors, which determines the commutation coefficients). Thus this is the desired relation between parallel transfer and distances in space-time. As a specific simple example, flat space-time in Cartesian coordinates has a coordinate basis ($\gamma_{edc} = 0$) which is also an orthonormal basis (so $g_{cd,e} = 0$); then it follows immediately that $\Gamma_{ced} = 0$ also.

Two particular cases are important: if we choose a coordinate basis, then (126) shows that the $\Gamma^a{}_{ed}$ (called Christoffel symbols in this case) are symmetric on the last two indices, and the last bracket in (131) vanishes. We find then

$$\Gamma^a{}_{ed} = g^{ac}\Gamma_{ced}, \quad \Gamma_{ced} = \frac{1}{2}\{g_{cd,e} + g_{ec,d} - g_{de,c}\}\,.$$

If we choose an orthonormal basis, then (117) show that the Γ_{aed} (called Ricci rotation coefficients in this case) are skew-symmetric in the first and last indices (this is because they generate a Lorentz transformation, cf. the last section of the last Chapter), and the first bracket in (131) vanishes. We find then

$$\Gamma^a{}_{ed} = g^{ac}\Gamma_{ced}, \quad \Gamma_{ced} = -\frac{1}{2}\{\gamma_{edc} + \gamma_{dec} - \gamma_{ced}\}\,.$$

5.4.4 Euclidean space quantities

It is useful to relate the formalism above to the usual Euclidean formalism of grad, div, curl. The relations are as follows:

$grad\, f = \nabla f \Leftrightarrow \mathbf{d}f$ with covariant components $f_{,a}$, contravariant components $g^{ab}f_{,b}$.

$div\, \mathbf{X} = X^a{}_{;a} = X^a{}_{,a} + X^d\Gamma^a{}_{ad} = \frac{1}{\sqrt{|g|}}\frac{\partial(X^a\sqrt{|g|})}{\partial x^a}$ (the last following from (123)).

$(curl\, \mathbf{X})^i = \eta^{ijk}X_{j;k}$ where η^{ijk} is the 3- dimensional analogue of the volume element η^{abcd} In a coordinate basis, $(curl\, \mathbf{X})^i = \eta^{ijk}X_{j,k}$ where $X_j = g_{js}X^s$.

$\nabla^2 f = f^{;a}{}_{;a} = f_{;ab}g^{ab} = f_{,ab}g^{ab} - f_{,d}\Gamma^d{}_{ba}g^{ab}$.

The detailed form of any of these operators in curved coordinates, or relative to a 'physical basis' (i.e. an orthonormal basis) associated with such coordinates, is found by determining the $\Gamma^d{}_{ba}$ for the appropriate basis.

5.5 The Lie Derivative

There is another derivative, similar to the covariant derivative in many ways, that can be useful to us. This is the *Lie derivative*, which is the tensor derivative associated with the process of dragging along.

Consider a vector field X^a, and the transformations Φ_V of the space into itself generated by X^a: Φ_V is the mapping where each point P is transported a parameter distance V to $Q = \Phi_V(P)$ along the curve $x^i(v)$ through P with tangent vector $X^i = dx^i/dv$. Then surfaces, curves, and vectors will all be dragged along with the points too, moving along the integral curves of X^i from P to Q. The prescription for finding what they are transformed to is to insist that all relations between points, functions, curves, and vectors are preserved under dragging along; for example $f(Q) = \Phi_V f(P)$ if f is mapped to $\Phi_V f$ by the transformation.

The Lie derivative $L_{\mathbf{X}}$ is just the derivative associated with dragging along, i.e. it gives the difference between the actual object at Q and the dragged along one. In the case of a function this just gives the ordinary directional derivative: $L_X(f) = X(f)$ (which is also the same as the covariant derivative of a function). The key point is to determine the Lie derivative of vectors. Then the extension to arbitrary tensors is made analogously to the case of the covariant derivative.

Consider the vector field X^i generating curves $\gamma(v)$ along which Y^j is dragged along to give $\Phi_V(Y^i)$. Consider the case when P is mapped a small parameter distance δv to R along $\gamma(v)$, and then a small parameter distance $\delta\tau$ along the integral curve of $\Phi_{\delta v}(Y^i)$ from R to T. On the other hand, P is mapped the small parameter distance $\delta\tau$ along the integral curve of Y^i to Q, and then the small parameter distance δv to T along $\gamma(v)$ (these four small transformations exactly fitting together because by definition $\Phi_V(Y^i)$ is dragged along by X^i).

In this case the coordinates of the four points are related as follows:

$$x^i(R) = x^i(P) + \delta v\, X^i(P)$$

$$x^i(T) = x^i(R) + \delta\tau\, \Phi_{\delta v}(Y^i)(R)$$

$$x^i(Q) = x^i(P) + \delta\tau\, Y^i(P)$$

$$x^i(T) = x^i(Q) + \delta v\, X^i(Q)$$

Substituting the first in the second gives

$$x^i(T) = x^i(P) + \delta v\, X^i(P) + \delta\tau\, \Phi_{\delta v}(Y^i)(R)$$

while substituting the third in the fourth gives

$$x^i(T) = x^i(P) + \delta\tau\, Y^i(P) + \delta v\, X^i(Q)$$

$$= x^i(P) + \delta\tau\, Y^i(P) + \delta v\, \left(X^i(P) + X^i{}_{,j}(P)\, Y^j(P)\delta\tau\right)$$

where in the last step we use the first term of a Taylor expansion to replace the value at Q by the value at P. Equating these two values for $x^i(T)$ gives

$$\Phi_{\delta v}(Y^i)(R) = Y^i(P) + \delta v\, X^i{}_{,j}(P)\, Y^j(P)$$

Now by the definition of the Lie derivative,

$$-\delta v(L_X(Y))^i = \Phi_{\delta v}(Y^i)(R) - Y^i(R)$$

where $Y^i(R) = Y^i(P) + \delta v Y^i{}_{,j}(P)\, X^j(P)$ so we find

$$-\delta v(L_X(Y))^i = Y^i(P) + \delta v\, X^i{}_{,j}(P)\, Y^j(P) - Y^i(P) - \delta v Y^i{}_{,j}(P)\, X^j(P) =$$

$$= \delta v\, (X^i{}_{,j}(P) Y^j(P) - Y^i{}_{,j}(P)\, X^j(P))$$

hence

$$(L_X(Y))^i = Y^i{}_{,j} X^j - X^i{}_{,j} Y^j = Y^i{}_{;j} X^j - X^i{}_{;j} Y^j \tag{133}$$

the last following because all calculations are being carried out in a local coordinate basis (see (127) above). Now comparing this with (73), we find the Lie derivative is just the commutator of the vectors:

$$L_{\mathbf{X}}(\mathbf{Y}) = [\mathbf{X}, \mathbf{Y}] = \nabla_{\mathbf{X}}(\mathbf{Y}) - \nabla_{\mathbf{Y}}(\mathbf{X}) \tag{134}$$

the last equality coming from (127). In particular, the two vector fields are dragged into each other by each other (they exactly fit together) if and only if they commute: $[X, Y] = 0$. This is in fact why the commutators of coordinate basis vectors are all zero.

The extension to other tensors is done using the same methods as in the case of covariant derivatives; that is Requirements (1)-(4) of section 4.2.2 hold, but now with (109) replaced by (133,134). However we can choose whether to use the covariant derivative or partial derivative expressions for the Lie derivative, obtaining different forms of the final result. For example, to determine the Lie derivative of T_{ab}, we form the scalar $T_{ab}Y^aZ^b$ and observe that

$$L_X(T_{ab}Y^aZ^b) = L_X(T_{ab})Y^aZ^b + T_{ab}L_X(Y^a)Z^b + T_{ab}Y^aL_X(Z^b) \Leftrightarrow$$

$$X(T_{ab}Y^aZ^b) = L_X(T_{ab})Y^aZ^b + T_{ab}(Y^a{}_{;c}X^c - X^a{}_{;c}Y^c)Z^b + T_{ab}Y^a(Z^b{}_{;c}X^c - X^b{}_{;c}Z^c)$$

But the left hand side is the same as

$$\nabla_X(T_{ab}Y^aZ^b) = \nabla_X(T_{ab})Y^aZ^b + T_{ab}Y^a{}_{;c}X^cZ^b + T_{ab}Y^aZ^b{}_{;c}X^c$$

Equating the two right hand sides, we find

$$L_X(T_{ab})Y^aZ^b = \nabla_X(T_{ab})Y^aZ^b + T_{db}X^d{}_{;a}Y^aZ^b + T_{ad}Y^aX^d{}_{;b}Z^b$$

for all Y^a, Z^b so finally

$$L_X(T_{ab}) = \nabla_X(T_{ab}) + T_{db}X^d{}_{;a} + T_{ad}X^d{}_{;b} \tag{135}$$

The important application of this is to the metric: in that case we see

$$L_X(g_{ab}) = g_{db}X^d{}_{;a} + g_{ad}X^d{}_{;b} = X_{b;a} + X_{a;b} \tag{136}$$

Thus the Lie derivative of the metric tensor vanishes when

$$L_X(g_{ab}) = 0 \quad \Leftrightarrow \quad X_{b;a} + X_{a;b} = 0 \tag{137}$$

These are *Killings equations*, the definition of a symmetry of a space-time, in covariant form.

To obtain a coordinate expression for the Lie derivative, note if we use a coordinate basis, the second line could alternatively be written

$$X(T_{ab}Y^aZ^b) = L_X(T_{ab})Y^aZ^b + T_{ab}(Y^a{}_{,c}X^c - X^a{}_{,c}Y^c)Z^b + T_{ab}Y^a(Z^b{}_{,c}X^c - X^b{}_{,c}Z^c)$$

If we now choose the arbitrary vectors as coordinate basis vectors, setting $Y^i = \delta^i_k$ and $Z^i = \delta^i_m$, we find

$$L_X(T_{km}) = T_{km,s}X^s + T_{am}X^a{}_{,l} + T_{lb}X^b{}_{,m} \tag{138}$$

so for the metric tensor we find an alternative form of Killing's equations: using a coordinate basis,

$$L_X(g_{km}) = g_{km,c}X^c + g_{am}X^a{}_{,l} + g_{lb}X^b{}_{,m} \tag{139}$$

In particular, if the vector X^i is chosen to lie along a coordinate curve: $X^i = \delta^i_s$, we find

$$L_s g_{km} = g_{km,s} \tag{140}$$

so the result corresponds to common sense: if the metric is independent of the coordinate x^s, then the space has a symmetry along the x^s-coordinate lines.

The solutions of Killing's equations in any particular space-time form a (finite dimensional) Lie algebra (possibly empty, if there are no symmetries). To fully explain this we need the concept of curvature. Finally, note that unlike the covariant derivative, in general we cannot raise and lower indices with impunity when using the Lie derivative; for in general the Lie derivative of the metric is not zero.

Examples 4: Covariant Differentiation

.

1.Given that $\nabla_a X^b = X^b_{;a} = X^b_{,a} + \Gamma^b_{ac}X^c$ is a tensor whenever X^a is a vector, find the transformation law for Γ^a_{bc}. Confirm that Γ^a_{bc} can be zero in a coordinate system and different from zero in another.

2. (i) Write down expressions for the covariant derivative of $R^{ab}{}_{cd}$ and of $S^a{}_c$. Contract the expression for the covariant derivative of $R^{ab}{}_{cd}$ on b and d, and compare with the expression for $S^a{}_c$. Hence prove that in this case the operation of taking the covarinat derivative commutes with contraction.

(ii) Prove that $T_{ab;c}$ is skew symmetric in a and b is T_{ab} is.

(iii) Suppose that S_{ab} is trace-free in a and b (i.e. $S^a{}_a = 0$). Determine if the same is true for $S_{ab;c}$ or not.

3. In the euclidean plane, check that the vectors $\mathbf{e}_r = \frac{\partial}{\partial r}$ and $\mathbf{e}_\theta = \frac{1}{r}\frac{\partial}{\partial\theta}$ form an orthonormal basis. Find their commutation coefficients γ^c_{ab} from (129) or (130). Find also the Γ^c_{ab} from (131). Confirm that they agree with the geometrical interpretation: Γ^c_{ab} is the c component of the $\mathbf{e}_a$ covariant derivative of the basis vector $\mathbf{e}_b$ (see (104)).

---

6 Physics in a curved space-time

Given our understanding of covariant differentiation, we can now reconsider the way to describe local physics in a curved space-time. The way this is usually done is by trying the assumption of *minimal coupling*, i.e. making the simplest possible transition from known physics in a flat space-time to covariant equations in a curved space-time.

In more detail, we aim for a space-time (4-dimensional) tensor form of all physical equations, so that they are manifestly true in all coordinate systems or reference frames if they are true in one (cf. Section 2.3)[2].

Now by minimal coupling, given the form of algebraic physical laws (i.e. those that do not involve differentiation) in flat space-time, we can assume the same form holds in an orthonormal local system in curved space-time (cf. Section 3.3); but what about equations that involve derivatives? The idea is that given a known physical law in Cartesian/Minkowski coordinates in flat space time, we

(a) change all partial derivatives to covariant derivatives, obtaining the general tensor form of the equations in flat space-time (valid in any coordinate system); and then

(b) assume the same form holds unchanged in curved space-time. In this way we obtain the simplest covariant form of the equations that reduces to the known flat-space form; in particular, we thereby assume there is no explicit coupling to space-time curvature in the equations.

Thus for example we might have an equation involving the Wave Operator, which is

$$\Box\phi = \phi_{,ab}g^{ab} = -\phi_{,00} + \phi_{,11} + \phi_{,22} + \phi_{,33}$$

that is

$$\Box\phi = -\frac{\partial^2\phi}{\partial t^2} + c^2\left(\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2}\right)$$

in flat space-time and Cartesian coordinates (using the metric (12)); then we assume that in curved space-time this becomes

$$\Box\phi = \phi_{;ab}g^{ab} = \phi_{,ab}g^{ab} - \phi_{,d}\Gamma^d{}_{ab}g^{ab}$$

which follows on simply changing the partial to covariant derivatives, and is the general tensor form valid in all coordinate systems in flat space-time.

This procedure will usually, but not always, give us a unique way of extending known physics from flat to curved space-time. However

[2] One can also write them in spinor form: this is essentially equivalent to using combinations of tensor equations written in terms of an orthonormal frame.

(a) in some cases minimal coupling is not unique, and we have to make a choice between alternatives;

(b) sometimes theoretical reasons may suggest non-minimal coupling (e.g. this is being proposed at present in relation to the equations for scalar fields in the early universe); the challenge then is to give experimental evidence that this is physically correct;

(c) in some cases, the minimal coupling idea may be quite incorrect. The most striking example of this kind is gravity, considered in the last section of this chapter.

We consider in turn, the force law; conservation equations; electromagnetism; and scalar fields, in a curved space-time.

6.1 Force law

Consider a particle with rest-mass $m_0 \neq 0$, whose world line is $x^a(\tau)$ where τ is proper time along that world line (given by (14)). Its 4-velocity (see section 3.3.1 and equation (82)) is

$$u^a = dx^a/d\tau, \quad u^a u_a = -1 \tag{141}$$

Its 4-momentum (see section 3.3.2 and equation (85)) is

$$P^a = m_o u^a \quad \Rightarrow \quad P^a P_a = -m_0^2 \tag{142}$$

(these equations being interpreted in a local orthonormal frame as in (86), (88)). Now the force law in flat space time is (89). To implement minimal coupling, we note first that the form taken by the force law in flat space-time in arbitrary coordinates (obtained by changing the ordinary derivative to a covariant derivative) is

$$F^a \;=\; \frac{\delta P^a}{\delta\tau} \;=\; \frac{dP^a}{d\tau} + P^c \Gamma^a{}_{dc} U^d \,. \tag{143}$$

We then take this form over unchanged to curved space-time. Thus equation (143) is the (tensorial) form we assume for the force law in general. If we consider in particular the case of constant rest mass ($m_0 = const$), then

$$F^a = m_0 A^a, \quad A^a := \frac{\delta u^a}{\delta\tau} = u^a{}_{;b} u^b \tag{144}$$

This equation motivates calling A^a the *acceleration vector*. However it has a somewhat different meaning than the Newtonian idea of an acceleration vector (see the discussion of gravity below); we must remember that its properties are only what follows from the definition (144). One consequence of these definitions is that this vector is orthogonal to u^a:

$$u^a u_a = -1 \;\Rightarrow\; u^a u_{a;b} = 0 \;\Rightarrow\; u^a u_{a;b} u^b = u^a A_a = 0$$

The force law governs the motion of objects in a curved space-time. It has one immediate important consequence:

$$F^a = 0 \quad \Leftrightarrow \quad \frac{\delta P^a}{\delta\tau} = 0 \quad \Leftrightarrow \quad \frac{\delta(m_0 u^a)}{\delta\tau} = 0$$

so if no force acts on a particle of constant mass then its motion is geodesic:

$$\{F^a = 0, m_0 = const\} \quad \Rightarrow \quad \frac{\delta u^a}{\delta\tau} = \frac{du^a}{d\tau} + u^c\Gamma^a{}_{dc}u^d = 0. \qquad (145)$$

We discuss this further in the last section of this chapter.

Finally we can consider again the idea of a particle with rest-mass $m_0 = 0$, see (90). If its world-line is $x^a(v)$ where v is some convenient parameter, then (cf. (141,142)) its 4-velocity and 4-momentum are

$$u^a = dx^a/dv, \quad P^a = \mu u^a, \quad \mu \neq 0 \qquad (146)$$

where from (142) the zero rest-mass condition is

$$P^a P_a = 0 \quad \Rightarrow \quad u^a u_a = 0 \qquad (147)$$

Thus such a particle of necessity has to move at the speed of light. Indeed an example is a photon ('a particle of light').

6.2 Conservation Equations

Conservation is concerned with quantities integrated over a surface: when do they remain constant as we vary the surface in space-time? The fundamental issue in developing conservation laws over finite surfaces is that one has to integrate over a volume or surface to obtain such laws; and the only integrals that are well-defined in a general curved space-time are integrals of scalars (any attempt to define an integral of a vector over a volume, for example, will not be invariant under general changes of coordinates that are position dependent through the volume; and so will not be well-defined in a tensorial sense). Thus covariantly valid conservation laws have to relate to integrals of scalars (see Chapter 7 of *General Relativity*, H. Stephani, for further discussion).

6.2.1 Scalar conservation

If we have a 4-vector with vanishing divergence, then this expresses conservation of some quantity (mass, charge, etc.). Consider a timelike vector J^a, and a 3-surface volume element dS_a (a vector, as shown by its index); contracting J^a with dS_a and integrating over the surface S gives the flux I of J^a through S: $I = \int_S J^a dS_a$. In the case of a spacelike surface element, with unit normal n^a and volume d^3V, the (vector) surface element dS_a is $dS_a = -n_a d^3V$

(so dS_a is normal to the 3-volume, with $dS_a n^a = d^3V$); thus in this case, $I = -\int_S J^a n_a d^3V$.

Now the Gauss Law in a curved manifold is

$$\int_{\partial V} J^a dS_a = \int_V J^a{}_{;a} dV$$

where the 3-surface ∂V (spacelike in parts, timelike in parts) bounds the volume V (see e.g. Chapters 35 and 36 of *Mathematical Methods of Classical mechanics*, V. I. Arnold, for a discussion). If we choose this volume as a tube which has timelike sides and is bounded top and bottom by spacelike surfaces labelled S_1 and S_2 respectively, and for which either (a) the sides are parallel to J^a, or (b) the sides are in a region where $J^a = 0$ (for example it lies very far away in an asymptotically flat space- time), then there is no flux across the sides and the surface integral becomes $I_1 - I_2 := \int_{S_1} J^a dS_a - \int_{S_2} J^a dS_a$ (in case (a), there is no flux across a surface that is parallel to J^a; in case (b) there is no flux across timelike surfaces where $J^a = 0$) and the Gauss law shows

$$I_1 = I_2 + \int_V J^a{}_{;a} dV\,.$$

Consequently

$$J^a{}_{;a} = 0 \quad \Rightarrow \quad I_1 = I_2\,, \tag{148}$$

that is, vanishing divergence of J^a implies I is a conserved quantity for a 4-volume with sides parallel to J^a or with sides located where $J^a = 0$ (it is independent of the particular choices of S_1 and S_2).

To see what this means physically, split J^a into its space-time direction u^a and its magnitude ρ by the equation

$$J^a = \rho u^a, \quad u^a u_a = -1 \tag{149}$$

This defines the average 4-velocity of the conserved quantity represented by J^a, and its density ρ measured by an observer moving at that average velocity (rest mass density, electric charge density, etc.). Now for a spacelike 3-surface S (with unit normal n_a) bounding a 4-volume V, by (149)

$$I = \int_S J^a dS_a = \int \rho(-u^a n_a) d^3V$$

where the 3-volume is $d^3V = \eta_{ijk} dx_1^i dx_2^j dx_3^k = \sqrt{|^3g|} d^3v$ with coordinate volume $d^3v = \epsilon_{ijk} dx_1^i dx_2^j dx_3^k$ (cf. (46) and 3-dimensional volume element $\eta_{ijk} = \eta_{ijkl} n^l$ (cf. (47)). Thus

$$I = \int_S \rho \cosh\beta \sqrt{|^3g|}\, d^3v, \quad \cosh\beta \equiv -u^a n_a$$

is the total conserved quantity crossing the surface (rest mass, electric charge, number of particles, etc., depending on the nature of the conserved current J^a; in these different cases, ρ is the rest mass density, electric charge density, number density respectively, measured by an observer moving with velocity u^a). In terms of these quantities (see (149)), the divergence equation (148) is

$$J^a{}_{;a} = (\rho u^a)_{;a} = 0 \quad \Leftrightarrow \quad \dot{\rho} + \rho\Theta = 0 \tag{150}$$

where we have defined the expansion of u^a by $\Theta = u^a{}_{;a}$.

It is convenient to define the representative length $R(x^i)$ by $R^3 = \sqrt{|{}^3g|}$, that is, comoving volumes (with constant d^3v) scale like R^3. If we consider a very narrow tube around a particular world line C that is an integral curve of u^a, with n^a chosen parallel to u^a in this tube in the limit as it shrinks to zero (so that $u^a n_a = -1$ there) , we find

$$I = \int_S \rho R^3 d^3v = \rho R^3 \epsilon$$

is constant, where $\epsilon \equiv \int_S d^3v$ is the (constant) comoving coordinate volume of this thin tube. Thus (as in Newtonian theory) conservation is expressed by the differential relation

$$\rho R^3 = M, \quad \dot{M} = 0 \quad \Leftrightarrow \quad \dot{\rho} + 3\rho\frac{\dot{R}}{R} = 0 \tag{151}$$

where the latter is obtained by taking the covariant derivative of the former along this tube (in the direction u^a), and we denote the (comoving) time derivative measured by u^a by a dot: $\dot{\rho} = \rho_{;a}u^a$, $\dot{M} = M_{;a}u^a$, etc. Equations (150) and (151) together show that the expansion Θ gives the rate of change of volume:

$$\Theta = 3\dot{R}/R = (d^3V)^{\dot{}}/(d^3V) \tag{152}$$

which also follows either from detailed analysis of the fluid flow characterised by u^a (see 'Relativistic cosmology", G F R Ellis in Enrico Fermi School XLVII, ed. R K Sachs, Academic Press 1971, from now on referred to as 'RC') or from the useful identity

$$u^a{}_{;a} = \frac{1}{\sqrt{|g|}}\frac{\partial}{\partial x^a}(\sqrt{|g|}\,)u^a)$$

(which follows from (109) together with (123), expressing the divergence in terms of partial derivatives only (to obtain (12) use comoving coordinates such that $u^a = \delta^a_0 = n^a$, noting that then $g = -{}^3g = -R^6$). This equation gives a quick way of calculating Θ.

Conservation of mass, charge, particle number, etc., can thus be expressed in a variety of forms. We will usually come across them either in the form (150) or (151), which for example describe the conservation of rest-mass in cosmology (see RC).

6.3 Energy-momentum mass conservation

The energy-momentum-stress tensor (section 4.4.4) gives the energy and 4-momentum crossing a surface element dS_a by the relation $T^a = T^{ab}dS_b$. Conservation of energy and momentum is given by the minimal coupling form

$$T^{ab}{}_{;b} = 0 \tag{153}$$

generalising the flat-space conservation laws to curved space in the standard way. However we can't integrate the quantities T^a over a finite surface to get a vector conserved quantity in general, because (as mentioned above) we can't integrate a vector covariantly over a volume. Nevertheless (153) represents the local conservation of energy and momentum. We illustrate this with a specific example.

Consider a perfect fluid, with stress tensor

$$T^{ab} = (\mu + p)u^a u^b + pg^{ab}, \quad u^a u_a = -1 \tag{154}$$

and suitable equations of state relating the pressure p and relativistic energy density μ (see RC); the fluid 4-velocity u^a is uniquely determined by the stress tensor if and only if $\mu + p \neq 0$. Then

$$\begin{aligned} T^{ab}{}_{;b} &= \left((\mu + p)u^a u^b + pg^{ab}\right)_{;b} \\ &= (\mu + p)_{;b}u^b u^a + (\mu + p)u^a{}_{;b}u^b + (\mu + p)u^a u^b{}_{;b} + p_{;b}g^{ab} \\ &= (\dot{\mu} + \dot{p})u^a + (\mu + p)A^a + (\mu + p)\Theta u^a + p_{;b}g^{ab} \end{aligned}$$

where $\dot{\mu} = \mu_{,a}u^a$, A^a is the acceleration (see (144)), and $\theta = u^a{}_{;a}$ is the expansion. Contracting with u^a

$$u_a T^{ab}{}_{;b} = -(\dot{\mu} + \dot{p}) - (\mu + p)\Theta + \dot{p} = 0$$

so conservation of energy along the world lines is given by:

$$\dot{\mu} + (\mu + p)\Theta = 0 \Leftrightarrow \frac{\dot{\mu}}{(\mu + p)} + 3\frac{\dot{S}}{S} = 0 \tag{155}$$

(the implications of which depend on the equation of state, see RC). Substituting back and using (144), (155), we find

$$T^{ab}{}_{;b} = p_{;b}u^a u^b + (\mu + p)A^a + p_{;b}g^{ab} = 0$$

so conservation of momentum takes the form

$$(\mu + p)A^a + p_{;b}h^{ab} = 0, \quad h^{ab} := g^{ab} + u^a u^b \tag{156}$$

where h^{ab} is the projection tensor orthogonal to u^a. This can be regarded as the form of equation (144) that is applicable to the fluid: a spatial pressure gradient generates a force acting on the fluid that causes a non-zero acceleration, with the inertial mass density being given by $(\mu + p)$. In the case of pressure-free matter (often called 'dust'), no force acts and so the motion is geodesic (cf. (145) above): from (156),

$$p = 0 \quad \Rightarrow \quad \mu A^a = 0 \quad \Rightarrow \quad A^a = 0 \tag{157}$$

These equations are important in studying cosmology and stellar structure.

Three points about the stress tensor.

Firstly, because of the principle of equivalence (see the last section of this chapter), T^{ab} does not include a contribution from gravitational energy in any simple way (indeed gravitational energy is a complex and difficult subject that will not be dealt with here).

Secondly, if there is a Killing vector ξ^a in a space-time, then we get a conserved vectorial quantity by contracting the Killing vector with the stress tensor: by (137) and (153),

$$\xi_{a;b} = \xi_{[a;b]}, \quad J^a := T^{ab}\xi_b \quad \Rightarrow \quad J^a{}_{;a} = 0$$

Thus there is an associated conserved quantity for each Killing vector (i.e. for each space-time symmetry). This is helpful in understanding specific exact solutions of the Einstein equations; however a realistic space-time has no Killing vectors.

Thirdly, so far we have considered only one type of matter. However in general there will be various matter components present in space-time (baryons, photons, neutrinos, a scalar field, etc.) The total stress- tensor of such multi-component systems is obtained by adding the stress-tensors of the components: $T_{ab} = \sum_A {}_AT^{ab}$ where A labels the different components. Now while total energy and momentum are necessarily conserved, interchange of energy and momentum between the components is of course possible, so the energy and momentum of the individual components is not necessarily conserved. This is represented by interchange vectors ${}_AI^a$ showing how much energy and momentum has been gained or lost by each component, their total summing to zero to guarantee conservation of total energy- momentum:

$${}_AT^{ab}{}_{;b} = {}_AI^a \quad \Rightarrow \quad T^{ab}{}_{;b} = \sum_A {}_AI^a = 0$$

The quantities ${}_AI^a$ will be determined by the physics of interactions between the components.

6.4 Maxwell's equations

As mentioned in section (3.3.3), the electromagnetic field is represented by a skew tensor $F_{ab} = F_{[ab]}$, related to the electric and magnetic fields measured by an observer with 4-velocity u^a by

$$F^{ab} = -2E^{[a}u^{b]} + \eta^{abcd}H_c u_d, \quad E_a u^a = 0 = H_a u^a \tag{158}$$

(which gives the correct form of the force-law on a charged particle, the correct transformation properties for E_a and H_a on change of 4-velocity, and the correct values for the two invariants noted in section (3.3.3)). It is convenient to form the dual ${}^*F^{ab}$ of F^{ab} (see (54)), finding from (158)

$${}^*F^{ab} = \frac{1}{2}\eta^{abcd}F_{cd} = -\eta^{abcd}H_c u_d - 2H^{[a}u^{b]} \tag{159}$$

(on using (51)). Comparing this with (158), we see that the electric and magnetic fields change roles when we take the dual: more precisely if we swap E to H and H to $-E$ we change F^{ab} to its dual. These two forms easily enable us to obtain the covariant inverse relations: contracting (158), (159) with u_b we see

$$E^a = F^{ab}u_b, \quad H^a = {}^*F^{ab}u_b = \frac{1}{2}\eta^{abcd}u_b F_{cd} \tag{160}$$

Now the minimal coupling form for Maxwell's equations in a curved space-time is given by the pair of equations

$$F^{ab}{}_{;b} = J^a \tag{161}$$

$$F_{[ab;c]} = 0 \tag{162}$$

To see this, note that in flat space for the canonical observers (moving at constant x, y, z coordinate values in Minkowski coordinates) $u_{a;b} = 0$; so from (158), and using $\dot{E}^a = u^b E^a{}_{;b}$,

$$F^{ab}{}_{;b} = -\dot{E}^a + u^a E^b{}_{;b} + \eta^{abcd}H_{c;b}u_d.$$

Contracting with u_a and noting $E_a u^a = 0 \Rightarrow u^a \dot{E}_a = 0$ we find

$$j = E^b{}_{;b} \;\text{ where }\; j \equiv -u_a J^a.$$

Substituting back,

$$j^a = -\dot{E}^a + \eta^{abcd}H_{c;b}u_d \;\text{ where }\; j^a := h^a b J_b.$$

Thus we have obtained two of the four standard Maxwell's equations in flat space-time. In particular we see that

$$J^a = 0 \quad \Rightarrow \quad E^b{}_{;b} = 0,\ \dot{E}^a = \eta^{abcd} H_{c;b} u_d \tag{163}$$

as usual. To show that (162) leads to the other two Maxwell's equations, it is convenient to note the following:

$$\eta^{abcd} F_{[ab;c]} = \eta^{abcd} F_{ab;c} = (\eta^{abcd} F_{ab})_{;c} = \ {}^*F^{cd}{}_{;c}$$

so we can rewrite (162) as

$${}^*F^{cd}{}_{;c} = 0 \tag{164}$$

which is exactly the same form as (161) but without a source term. Thus using the symmetry leading from (158) to (159), we can immediately deduce that (164) leads to the usual other pair of Maxwell equations: swapping E to H and H to $-E$ in (163) (equivalent to the source- free version of (161)), we find

$$H^b{}_{;b} = 0,\ \dot{H}^a = -\eta^{abcd} E_{c;b} u_d$$

which are equivalent to (164).

It is useful to note three other features of the covariant Maxwell equations.

Firstly, we can rewrite (162) in cyclic form: using the skew symmetry of F_{ab},

$$F_{[ab;c]} = F_{ab;c} + F_{ca;b} + F_{bc;a} - F_{ac;b} - F_{ba;c} - F_{cb;a} = 2(F_{ab;c} + F_{ca;b} + F_{bc;a})$$

(in the last step, we combine the 1st and 5th, 2nd and 4th, 3rd and 6th terms). Thus (162) is the same as

$$F_{ab;c} + F_{ca;b} + F_{bc;a} = 0 \tag{165}$$

an alternative form of the equations that is often encountered.

Secondly, on using a coordinate basis, (162) is the same as $F_{[ab,c]} = 0$ (the Christoffel terms cancel because of their symmetry), which is just the integrability condition for existence of a vector potential for F_{ab}. Thus (162) are locally equivalent to

$$\exists \Phi_a(x^i): \quad F_{ab} = \Phi_{[a,b]} = \Phi_{[a;b]} \tag{166}$$

The other Maxwell equations (161) can be re-written in terms of this potential (but this involves the curvature, and so is deferred to the following chapter).

Thirdly, the stress tensor for the electromagnetic field is

$$T_{ab} = \alpha(F_a{}^c F_{bc} - \frac{1}{4} F_{cd} F^{cd} g_{ab}) \quad \Rightarrow \quad T^a{}_a = 0 \tag{167}$$

(α is a constant depending on the units chosen), incorporating in one tensor the electromagnetic energy density, Poynting vector, and Maxwell stress (see Appendix C of *Flat and Curved Space-times* for more details). Using (161), (165) one can show from (167) that

$$T^{ab}{}_{;b} = \alpha F^{ab} J_b \tag{168}$$

the term on the right representing the interchange of energy and momentum between the current and the electromagnetic field. It follows that if there is no current, the energy and momentum of the electromagnetic field is locally conserved:

$$J^a = 0 \quad \Rightarrow \quad T^{ab}{}_{;b} = 0.$$

Some further aspects of the electromagnetic equations can only be discussed after the concept of curvature has been introduced (in the next chapter). At present we note that *inter alia* these equations underlie the role of magnetic fields in astrophysics and cosmology, and also the geometric optics approximation that is the basis of the properties of light (classically considered), see RC. For the moment we note only one point: the geometric optics approximation confirms that light rays are null geodesics of space-time.

6.5 Scalar Field

A classical scalar field $\phi(x^j)$ has its motion determined by its potential $V(\phi)$. If minimally coupled, its equation of motion is

$$\Box\phi + \frac{dV}{d\phi} = 0 \tag{169}$$

where $\Box$ is the wave operator described at the beginning of the chapter. The corresponding stress tensor is

$$T_{ab} = \phi_{,a}\phi_{,b} - \left(V(\phi) + \frac{1}{2}\phi_{,c}\phi^{,c} \right) g_{ab} \tag{170}$$

From this we may note the following:

1] If $\phi = const$ in an open set U, then in that set $\phi_{,c} = 0 \Rightarrow T_{ab} = -V(\phi)g_{ab}$ i.e. we have a perfect fluid with $\mu + p = 0$.

2] If $\phi \neq const$ in an open set U, and $\phi_{,c}$ is timelike there, we can define a unique 4-velocity $u_a = \phi_{,a}/A$ where $A^2 = -\phi_{,a}g^{ab}\phi_{,b}$. Then the stress-tensor (170) takes the perfect fluid form (154) with this vector as the 4-velocity and with

$$\mu = \frac{1}{2}\dot{\phi}^2 + V(\phi), \quad p = \frac{1}{2}\dot{\phi}^2 - V(\phi) \tag{171}$$

3] Provided $\phi_{,c} \neq 0$ in an open set U, then in U the equation of motion (169) is satisfied if and only if the conservation equations $T^{ab}{}_{;b} = 0$ hold for the stress tensor (170); then (155) will hold in U with μ, p given by (171).

These equations have been the subject of intensive investigation recently because of their role in the inflationary universe models of cosmology. At present investigations are under way of non-minimally coupled scalar fields (where (169) is modified by terms explicitly containing the space-time curvature). These are motivated by quantum gravity or cosmological considerations.

6.6 Geodesics

Geodesics are the very special curves in space-time whose direction is unchanging, with the tangent vector parallel propagated along themselves (see section 4.1.2). Thus they are the nearest thing one can get to a straight line in a curved space-time. Consequently they play a special role in geometry and physics; indeed we have seen they represent the motion of freely-moving particles (145) and of pressure- free matter (157). Their physical meaning will be discussed further in the following section; this section considers their geometrical properties.

6.6.1 Parametrisation

A geodesic vector $X^a = dx^a/dv$ as defined above (equation (4.10)) obeys the equation

$$X^a{}_{;b}X^b = 0 \quad \Leftrightarrow \quad \ddot{x}^k + \Gamma^k{}_{ij}\dot{x}^i\dot{x}^j = 0 \tag{172}$$

where $\dot{x}^i = dx^i/dv$, $\ddot{x}^i = d^2x^i/dv^2$ and the second form follows on using (109,110). As usual, the $\Gamma's$ can be written in terms of the derivatives of the metric plus the rotation coefficients of the basis vectors.

However this form is not invariant under all reparametrisations of the curve. A curve parameter v for which the equation has the form (172) is called an *affine parameter*; we now investigate its invariance under parameter change.

Consider a change of parameter $v \to v' = v'(v)$ for a curve $x^a(v)$ with tangent vector $X^a = dx^a/dv$. Then

$$X^a = \frac{dx^a}{dv} \quad \to \quad X'^a = \frac{dx^a}{dv'} = \frac{dx^a}{dv}\frac{dv}{dv'} = X^a\frac{dv}{dv'}$$

a parallel vector but with different magnitude. Now $X'^a{}_{;b}X'^b = (X^a\frac{dv}{dv'})_{;b}X^b\frac{dv}{dv'}$ so

$$X'^a{}_{;b}X'^b = X^a{}_{;b}X^b(\frac{dv}{dv'})^2 + X^a\frac{d^2v}{dv'^2} \tag{173}$$

Consequently given a geodesic vector X^a, then

$$X^a{}_{;b}X^b = 0 \quad \Rightarrow \quad X'^a{}_{;b}X'^b = X'^a f, \quad f := \frac{d^2v}{dv'^2}/(\frac{dv}{dv'}) \tag{174}$$

As the change of X'^a along the curve is parallel to itself, this indicates this vector has a direction that is unchanging; but its magnitude changes along the curve for a general choice of parameter, for this will occur if $f \neq 0$. In that case it is a non-affinely parametrized geodesic.

However we can always remedy such a situation. If $X^a{}_{;b}X^b = g(v)X^b$ then we have an unchanging direction but changing magnitude (as just discussed). Changing the parameter as above, we can make the resulting vector X'^a satisfy (172) for appropriate choice of the new parameter. From (173), we want

$$X'^a{}_{;b}X'^b = 0 \quad \Leftrightarrow \quad \left(g(v)(\frac{dv}{dv'})^2 + \frac{d^2v}{dv'^2} \right) X^a = 0$$

and we can always satisfy this by choosing $v'(v)$ to solve the equation $\frac{d^2v}{dv'^2} + g(v)(\frac{dv}{dv'})^2 = 0$. Thus in this case we can always reparametrise to get a form in which both the direction and the magnitude are unchanged along the curve.

Finally suppose we have an affinely parametrized geodesic; what changes of parameter will preserve this form? From (173), $X^a{}_{;b}X^b = 0 \Rightarrow X'^a{}_{;b}X'^b = 0$ iff

$$\frac{d^2v}{dv'^2} = 0 \Leftrightarrow v' = av + b, \quad a,\ b\ const \tag{175}$$

This is the group of *affine transformations*. Thus in upshot: given a geodesic curve, we can always choose an affine parameter; the set of affine parameters are arbitrary by the affine transformations (175).

6.6.2 Existence

Equation (172) is a 2nd order equation for $x^i(v)$. Given a point p and a direction $X^i(p) = \frac{dx^i}{d\tau}(p)$ there, there will be a unique solution of the equation which is the geodesic through that point and with that starting direction. Formal mathematical theorems prove this, but geometrically it is clear: we just continue the curve without changing direction. Furthermore there will be some neighbourhood V of p such that none of the geodesics emanating from p intersect each other within V (eventually they may intersect, as happens for example in gravitational lensing, but this cannot happen arbitrarily close to p). In fact each point p is contained in a *normal neighbourhood* U: that is, an open set containing p such that one and only one geodesic contained in U connects each pair of points in U. Then none of the geodesics starting from any point q in U can intersect any other geodesic starting from q, while remaining in U.

6.6.3 Geodesic coordinates

The existence theorem for geodesics enables us to prove the existence of various geodesic coordinates in a general space-time. Three are particularly useful.

Firstly, we can locally choose geodesic coordinates based on some surface S by drawing the normal geodesics to that surface, using proper time or proper distance as the affine parameter, and using these as coordinate curves. As an example if the surface is spacelike then we can choose it to be $t = t_0$ and the metric will take the form

$$ds^2 = -dt^2 + h_{\alpha\beta}dx^\alpha dx^\beta \tag{176}$$

where α, β run from 1 to 3, and the curves with tangent vector $X^0 = \delta^a_0$ are orthogonal to all the surfaces $\{t = const\}$. The geodesic equation (172) for these normal curves reduces to $\Gamma^a{}_{00} = 0$, which is identically satisfied whenever the metric has the form (176). These coordinates are often used in cosmology.

Secondly we can use geodesic coordinates based on a point p. We choose coordinates so that the geodesics through p ('radial geodesics') are the curves $x^i = const$ (we can simply label these geodesics in this way). Then $dx^i/dv = 0 \Rightarrow d^2x^i/dv^2 = 0$ along these curves, so from the geodesic equation (172), $\Gamma^a{}_{bc}(dx^b/dv)(dx^c/dv) = 0$ along them. Away from p, this sets to zero just those combinations of Γ's corresponding to the direction of the radial geodesics there; however at p, this includes all directions, so at that point all the components of $\Gamma^a{}_{bc}$ are zero. By the Christoffel relations, this sets to zero the derivatives of the metric there. As we can also choose the coordinate basis vectors to be orthogonal to each other at p, we can always choose coordinates so that at a particular chosen point p,

$$g_{ab}|_p = \eta_{ab}, \quad g_{ab,c}|_p = 0 \;\Leftrightarrow\; \Gamma^a{}_{bc}|_p = 0 \tag{177}$$

that is the coordinates have the Minkowski form there (they will not have this form away from p, however, unless the space-time is flat). Such coordinates are called *Normal Coordinates* based on p.

Thirdly we can similarly choose normal coordinates based on a geodesic world-line γ. We proceed as above at each point p on the world line; we obtain coordinates where (177) holds at each point on γ. Such coordinates provide a local 'Newtonian-like' frame in which to examine local physics in a curved space-time. They are very convenient to use because the vanishing of the Christoffel symbols means that *partial and covariant derivatives of tensors are the same at each point on γ when we use such coordinates* (e.g. $T^a{}_{;bc}|\gamma = T^a{}_{,bc}|\gamma$.). Therefore the minimal coupling assumption discussed in the first section of this chapter implies *when we use normal coordinates based on a world-line γ, the equations of physics in the curved spacetime are unchanged from their flat space-time form at each point on γ.*

6.6.4 Extremal properties

In Euclidean space, a straight line is the shortest distance between two points. In a curved space-time, *geodesics are the local maxima or minima of space-time distance*: maximum in the case of timelike curves, minimum in the case of space-like curves (the difference occurring because of the negative sign in the proper time integral (14), cf. the proper distance integral (16)).

To see this extremal property, we remember that the Euler-Lagrange functional $\Phi = \int_P^Q L(x^i, \dot{x}^j)dt$ is a local extremal on the space of curves $x^i(v)$ joining P to Q if and only if

$$\left(\frac{\partial L}{\partial x^s}\right) - \frac{d}{dv}\left(\frac{\partial L}{\partial \dot{x}^s}\right) = 0 \tag{178}$$

where $\dot{x}^i = dx^i/dv$. Consider a Lagrangian

$$L(x^i, \dot{x}^j) = F(S), \quad S(x^i, \dot{x}^j) := g_{ij}(x^k)\dot{x}^i\dot{x}^j\,, \quad dF/dS \neq 0, \tag{179}$$

This gives proper time if $F(S) = (-S)^{1/2}$, and proper distance if $F(S) = S^{1/2}$. From (178), using commas to represent coordinate derivatives,

$$\frac{\partial S}{\partial x^s} = g_{ij,s}\dot{x}^i\dot{x}^j, \quad \frac{\partial S}{\partial \dot{x}^s} = 2g_{ij}\dot{x}^i. \tag{180}$$

Substituting (179) and (180) into (178), we find

$$\frac{dF}{dS}\frac{\partial S}{\partial x^s} - \frac{d}{dv}\left(\frac{dF}{dS}\frac{\partial S}{\partial \dot{x}^s}\right) = \frac{dF}{dS}g_{ij,s}\dot{x}^i\dot{x}^j - \frac{d}{dv}\left(\frac{dF}{dS}2g_{is}\dot{x}^i\right) = 0$$

Thus

$$\frac{dF}{dS}g_{ij,s}\dot{x}^i\dot{x}^j - \frac{d^2F}{dS^2}\frac{dS}{dv}2g_{is}\dot{x}^i - \frac{dF}{ds}(2g_{is,j}\dot{x}^i\dot{x}^j + 2g_{is}\ddot{x}^i) = 0$$

Dividing by $(-2dF/dS)$, we get

$$g_{is}\ddot{x}^i + \frac{1}{2}(g_{is,j} + g_{js,i} - g_{ij,s})\dot{x}^i\dot{x}^j = G(S)g_{is}\dot{x}^i$$

where the second term has been taken to the right hand side and the third split into two equal parts; and

$$G(S) := -\left(\frac{d^2F}{dS^2}/\frac{dF}{dS}\right)\frac{dS}{dv}.$$

Finally multiplying by g^{sk}, we obtain from (131) (as we are using a coordinate basis)

$$\ddot{x}^k + \Gamma^k{}_{ij}\dot{x}^i\dot{x}^j = G(S)\dot{x}^k \tag{181}$$

which has the form (174) showing this is a geodesic (although in general not affinely parametrized).

Now on a timelike curve, proper time τ is given by $\tau = \int(-S)^{1/2}dv$, so $d\tau/dv = (-S)^{1/2}$ and $d^2\tau/dv^2 = -\frac{1}{2}(-S)^{-1/2}dS/dv$. Hence

$$\tau = av + b \quad \Leftrightarrow \quad d^2\tau/dv^2 = 0 \quad \Leftrightarrow \quad dS/dv = 0 \quad \Rightarrow G(S) = 0$$

where a, b are constants (recovering the affine transformations (175)). Thus in particular choosing $v = \tau$ gives an affinely parametrized geodesic, for all $F(S)$. If we choose $F(S) = (-S)^{1/2}$, we have shown that *a curve locally maximising proper time is a geodesic, with proper time an affine parameter* (we know already from the time dilation formula (15) that this must be a maximum rather than a minimum; indeed one can always make a timelike curve shorter by going closer to the light cone). Similarly in the spacelike case, if we choose $F(S) = (S)^{1/2}$, we have shown that *a curve locally minimising proper distance is a geodesic, with proper distance an affine parameter.*

The fact that a geodesic (locally) gives the longest timelike curve in space-time between two events is the mathematical formulation of the 'twin paradox' (*Flat and Curved Space-Time*, p.32-34, 88-91, 168- 170), resolving once and for all the unique feature of the 'longest- lived' twin: she is the one who moves on a geodesic curve in space- time (i.e. has never undergone an acceleration). The fact that the shortest spacelike curve is a geodesic shows (as a special case) that a straight line is the shortest distance between two points in Euclidean space, and that a great circle is the shortest distance between two points on sphere.

Three comments:

1] Having shown that geodesics locally maximise or minimise distance, to actually find them, it is simplest to use the variational principle above but choosing $F = S$. The above analysis shows the curve is the same, indeed by (181) and the definition of G, we again obtain an affinely parametrized geodesic. However the calculations will be simpler than with any other choice of $F(S)$.

2] It has been emphasized in the discussion that the variational results given are local. It can be shown that the extremal property no longer holds after focusing of geodesics occurs (when there are caustics or cusps on the geodesics emanating from a point, and several geodesics join the same pair of points). While the variational result may still be true in a small neighbourhood of the original curve (small variations of a timelike curve always decrease the proper time between two events, for example) there may now be quite different geodesic curves joining the two events. Clearly if one of these geodesics is a maximum, then in general any others joining the same events cannot also be maximal (for only in very exceptional circumstances will the distances along different geodesics joining the same events be the same).

3] The extremal principle for geodesics given above is the basis for a relativistic version of *Fermat's theorem*, characterising null geodesics as the null curves from a given event that give extrema of arrival time on a timelike curve (Section 3.3 of *Gravitational Lenses*, P. Schneider, J. Ehlers and E. Falco, Springer). This is very helpful in the study of gravitational lenses (see sections 4.3, 4.6, 5.3, and 5.5 of that book).

6.6.5 Calculation of Gammas

The variational principle for geodesics can be used to calculate Christoffel terms with ease. As an example consider the metric form

$$ds^2 = -A(r)dt^2 + B(r)dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)$$

where $x^0 = t$, $x^1 = r$, $x^2 = \theta$, $x^3 = \phi$. Choosing $F = S$, (179) shows

$$L = -A(r)\dot{t}^2 + B(r)\dot{r}^2 + r^2(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2)$$

where for any function f, $\dot{f} = df/dv$. Then

$$\begin{array}{lcl}
\dfrac{\partial L}{\partial x^0} & = & 0 \\
\dfrac{\partial L}{\partial \dot{x}^0} & = & -2A(r)\dot{t} \\
\dfrac{\partial L}{\partial x^1} & = & -A'(r)\dot{t}^2 + B'(r)\dot{r}^2 + 2r(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2) \\
\dfrac{\partial L}{\partial \dot{x}^1} & = & 2B(r)\dot{r} \\
\dfrac{\partial L}{\partial x^2} & = & r^2 2\sin\theta\cos\theta\dot{\phi}^2 \\
\dfrac{\partial L}{\partial \dot{x}^2} & =\frac{\sin^2\theta}{B} & 2r\dot{\theta}^2 \\
\dfrac{\partial L}{\partial x^3} & = & 0 \\
\dfrac{\partial L}{\partial \dot{x}^3} & = & 2r^2\sin^2\theta\dot{\phi}
\end{array} \tag{182}$$

Consequently putting $a = 0$ in the variational equation (178) gives

$$0 + \frac{d}{dv}\left(2A(r)\dot{t}\right) = 0$$

where $r = r(v)$, $t = t(v)$, so $A(r) = A(r(v))$ and $\dot{t} = \dot{t}(v)$. Thus

$$2A(r)\ddot{t} + 2A'(r)\dot{r}\dot{t} = 0 \quad \Leftrightarrow \quad \ddot{t} + \frac{A'(r)}{A(r)}\dot{r}\dot{t} = 0.$$

Comparing with the geodesic equation (172) for $k = 0$, which is,

$$\ddot{x}^0 + \Gamma^0{}_{ij}\dot{x}^i\dot{x}^j = 0$$

we see that

$$\Gamma^0{}_{10} = \frac{1}{2}\frac{A'(r)}{A(r)} = \Gamma^0{}_{01}, \quad \Gamma^0{}_{ab} = 0 \ \ otherwise.$$

Similarly putting $k = 1$, the non-zero $\Gamma^1{}_{bc}$ are

$$\Gamma^1{}_{00} = \frac{1}{2}\frac{A'}{B}, \quad \Gamma^1{}_{11} = \frac{1}{2}\frac{A'}{A}, \quad \Gamma^1{}_{22} = -\frac{r}{B}, \quad \Gamma^1{}_{33} = -\frac{r\sin^2\theta}{B}$$

putting $k = 2$, the non-zero $\Gamma^2{}_{bc}$ are

$$\Gamma^2{}_{12} = \Gamma^2{}_{21} = \frac{1}{r}, \quad \Gamma^2{}_{33} = -\sin\theta\cos\theta$$

and putting $k = 3$, the non-zero $\Gamma^3{}_{bc}$ are

$$\Gamma^3{}_{13} = \Gamma^3{}_{31} = \frac{1}{r}, \quad \Gamma^3{}_{23} = \Gamma^3{}_{23} = \frac{\cos\theta}{\sin\theta}.$$

6.7 Gravity

The discussion so far has not included gravity. How do we do this? The fundamental insight of Einstein was to realise that gravity is unlike every other force, so the minimal coupling idea completely fails in this case. Ultimately this is the reason why gravity is best described by a curved space-time structure.

The point is that when one allows arbitrary choice of coordinates and reference frames, there is no vector G^a describing the gravitational force in a way analogous to the way a Newtonian gravitational force is described by a 3-vector g^α. The reason for this is that one can transform gravity away by change of reference frame to a freely falling frame; and conversely, one can generate an apparent gravitational field by changing from a freely falling frame to an accelerating frame. This is the burden of Einstein's famous 'thought experiments' concerning an observer in a lift. Because all objects accelerate at the same rate in the earth's gravitational field (the Tower of Pisa experiment by Galileo), an observer isolated in a lift cannot distinguish between the earth's gravitational field and a uniform acceleration, through any experiments he carries out in the lift. Furthermore the gravitational field can be transformed to zero by letting the lift fall freely (without friction), for then if the observer drops a weight it will apparently float in the air alongside him - for it will accelerate down the liftshaft at exactly the same rate as the observer is accelerating (see e.g. section 5.2 of *Flat and Curved Space-Times*). This is not just theory: indeed it is now commonplace to see films of astronauts floating in their spacecraft as if gravity had been abolished.

Einstein formalised this as the *principle of equivalence*:

> *gravity and inertia are equivalent, as far as local physical experiments are concerned. They cannot be distinguished from each other experimentally.*

(see section 5.2 of *Flat and Curved Space-Times* for further discussion). Furthermore, both depend on the frame of reference adopted, and their combined effect can be transformed to zero by appropriate choice of reference frame. Thus *the gravitational force is not a tensor quantity*, for such a quantity vanishes in all frames if it vanishes in one (section 2.3). Gravity does not have this property. For this reason, the minimal coupling idea simply does not work in the case of gravity.

How then do we represent gravity? The key proposal is that

> *a particle moves on a space-time geodesic if in free fall, that is, if it moves under gravity and inertia alone.*

This is a slightly extended interpretation of equation (145). Previously we implicitly regarded this as the equation of motion of a body moving under inertia alone; we now interpret it as representing any freely falling object, moving under the combined effects of gravity and inertia, but with no other forces acting (this is discussed further insection 5.3 of *Flat and curved space-times*).

The point then is that in flat space time, inertial forces are represented by the Γ's in (98,100). We can see this for example by transforming from Minkowski coordinates to (a) rotating coordinates, and (b) uniformly accelerated coordinates. The equivalence principle however says that we cannot locally distinguish gravity from inertia. Thus we conclude that *both gravitational and inertial forces are incorporated in the* $\Gamma^a{}_{bc}$ in (143) and (145); this is consistent because (like the gravitational-inertial force) they can locally be set to zero by a change of coordinates (cf. the discussion of normal coordinates above). We thus arrive at the astonishing conclusion: there is no need to define a special vector to describe the gravitational force, it is already incorporated in the force law (143) via the $\Gamma^a{}_{bc}$. When we adopt (145) to describe free motion of a particle, this will (in a curved space-time) automatically include a description of the effect of the gravitational field on its motion. However it would be incorrect to say that these quantities describe gravity only; they describe ‘gravity plus inertia’ together, which cannot be intrinsically distinguished from each other by any local physical experiments.

6.8 Physical meaning of parallel transfer

We have identified timelike geodesics as representing the motion of free particles in space-time, and null geodesics as representing the paths of light rays. These represent parallel transfer of the tangentvector to a curve, alongthat curve. One final interesting question remains: what is the physical meaning of parallel transfer of spacelike vectors along a (timelike) world-line?

Consider a timelike geodesic. This can be taken as representing the motion of a freely falling observer. The issue then is, what physical meaning can we give to the question, what spatial direction at time t_2 in his history is parallel to a given direction at a time t_1? At first glance it might seem impossible to compare two directions ‘at the same place but at different times’, but this is what we do whenever we set up a physically non-rotating reference frame. There are many ways to do so; essentially, they are equivalent to using perfect gyroscopes to maintain a fixed spatial direction at different times. Geometrically, this procedure gives a rule for transferring a spatial vector along a timelike curve.

The obvious suggestion (cf. section 5.7 of *Flat and Curved Space- time*) is that this must correspond to parallel transport of a direction along that curve, as determined by the covariant derivative operator (see (98,99)); and indeed this turns out to be correct: parallel transfer of a vector along a geodesic corresponds to physically invariant direction, as determined by a Foucault pendulum or perfect gyroscopes (see *General Relativity*, J L Synge, for detailed discussion). Now such parallel transport is determined by the Γ’s and so by the derivatives of the metric. Thus whatever equations we find to determine the metric tensor will also determine the nature of physically non-rotating reference frames in space-time.

6.8.1 Gravity as geometry

In later chapters we will see that this proposal gives an accurate description of the observed motions of particles under the influence of gravity. However before we reach that point, there is a major issue to be understood first: What are the equations that determine the combined effects of gravity and inertia? They are determined by the Γ’s; and through the Christoffel relations, the Γ’s are given by the first derivatives of metric. Thus *the gravitational field equations must be equations for the space-time metric.*

Now as always we want tensor equations, but the first covariant derivative of the metric is zero, so we cannot obtain the required equations simply by taking derivatives of the metric. To get the required equations, we have to turn to the concept of space-time curvature, the subject of the following chapter. However for the moment one thing is clear: whatever these equations are, *the effect of gravity is encoded in the space-time metric, that is, it is determined by the space-time geometry.* Thus the theory of gravity has to be a theory for this geometry, making it quite different from every other force we know.

Inter alia this means that we can no longer take space-time as a fixed background on which physics occurs, but must regard it as an active participant in the flow of physics; and furthermore ways to measure and determine the geometry of space-time become an important part of physics. This is the basic view of general relativity: the Einstein gravitational field equations (to be discussed below, once space-time curvature has been defined) determine the space-time metric, which then forms the background in which all other physics takes place.

Examples 5: Covariant derivative applications

1. Assume ξ_a satisfies the Killing's equation $\xi_{(a;b)} = 0$.
(i) Show that $\xi_a k^a$ is constant along any geodesic vector k^a.
(ii) Show that $P^a = T^{ab}\xi_b$ has vanishing divergence if T^{ab} is the energy momentum tensor.

2. (i) Show that the Maxwell equations imply $J^\alpha{}_{;\alpha} = 0$. Interpret this as implying conservation of current. [Hint: use (161), the antisymmetry of F_{ab}].
(ii) Show that the Maxwell equations imply

$$T^{\alpha\beta}{}_{;\beta} = \frac{1}{4\pi} F^{\beta\alpha} J_\beta$$

(the rate of change of energy-momentum of the field is due to interaction with the sources), where $T^{\alpha\beta}$ is the electromagnetic stress tensor $T^{\alpha\beta} = \frac{1}{4\pi}(F^{\alpha\gamma}F^\beta{}_\gamma - \frac{1}{4}g^{\alpha\beta}F_{\gamma\delta}F^{\gamma\delta})$. [Hint: you have to use both sets of Maxwell equations].

3.(i) Show that, in a Robertson Walker universe, $u_{a;b} = \frac{1}{3}\Theta h_{ab}$, $\Theta = 3\dot{S}/S$, $\dot{u}_a = u_{a;b}u^b = 0$, where u^a is the preferred velocity $u^a = \delta^a_0$, and $h_{ab} = g_{ab} + u_a u_b$ (NB: this is the metric of the 3-surface orthogonal to u^a).
(ii) Using the perfect fluid form of T^{ab}, determine the form of the energy and momentum conservation equations in a Robertson-Walker universe. Show that one is identically satisfied and the other leads to non-trivial relations.

(iii) Deduce the form of Maxwell's equations found by the preferred observers in a Robertson Walker universe. Remember that $F_{ab} = -2E_{[a}u_{b]} + \eta_{ab}{}^{cd}H_c u_d$.

4. Show that if

$$ds^2 = (dx^1)^2 + g_{ab}dx^a dx^b \qquad a, b = 2, ...n$$

then the curve $\quad x^2 = const, x^3 = const, ...x^n = const \quad$ generated by $\frac{\partial}{\partial x^1}$ is a geodesic with x^1 an affine parameter. Hence show that each great circle on a 2-sphere is a geodesic.

5. Find the Christoffel symbols and the geodesic equation for the Friedman Robertson Walker spacetime with metric

$$ds^2 = -dt^2 + S^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + sin^2\theta d\varphi^2)\right] \qquad .$$

---

7 Curvature

The question now is what distinguishes a flat space-time from a curved one [3]. This is far from obvious. If we are given a space-time with coordinates such that the metric takes the form (22), we know it is flat; but if we now make arbitrary coordinate transformations, we can change this form to infinitely many other ones in which it is unclear that we are in fact dealing with flat space-time. Indeed there are many cases where published solutions of the Einstein equations have turned out to be flat space-time in complicated coordinates.

Is is clear, then, that there will be no simple criterion for flatness based directly on the metric tensor itself. The best approach is to consider the idea of curvature, either through considering the way surfaces of lower dimension are imbedded in surfaces of higher dimension (the way the idea was initially developed), or intrinsically through considering when parallel transfer along a curve is and is not integrable (see e.g. *Flat and Curved space-times*, sections 5.1 and 5.7). Taking the latter approach, it turns out that the central equation determining curvature is the Ricci identity, which we now consider.

7.1 The curvature tensor

We consider the second covariant derivatives $X^a{}_{;bc} := (X^a{}_{;b})_{;c}$ of an arbitrary vector field X^a, where this quantity has the obvious meaning: from X^a we form the tensor field $Y^a{}_b := X^a{}_{;b}$ by taking a covariant derivative; then we take another covariant derivative, to get $Y^a{}_{b;c} = X^a{}_{;bc}$ (the covariant derivatives of X^a and $Y^a{}_b$ being defined as in Chapter 3). Now the question is, what happens if we commute the derivatives, that is, if we interchange the 'b' and 'c'?

Calculating $X^a{}_{;bc}$ in the way just outlined (and allowing for a general basis), we find

$$(X^a{}_{;b})_{;c} = (X^a{}_{;b})_{,c} + X^e{}_{;b}\Gamma^a{}_{ce} - X^a{}_{;e}\Gamma^e{}_{cb}$$

Interchanging b and c, we obtain

$$(X^a{}_{;c})_{;b} = (X^a{}_{;c})_{,b} + X^e{}_{;c}\Gamma^a{}_{be} - X^a{}_{;e}\Gamma^e{}_{bc}$$

Subtracting and using (128) gives

$$X^a{}_{;bc} - X^a{}_{;cb} = (X^a{}_{;b})_{,c} - (X^a{}_{;c})_{,b} + X^e{}_{;b}\Gamma^a{}_{ce} - X^e{}_{;c}\Gamma^a{}_{be} - X^a{}_{;e}\gamma^e{}_{cb}$$

where $\gamma^c{}_{ab}$ are the commutation coefficients, zero iff a coordinate basis is used.

[3] most of what follows is true also in any curved space, however because of the application to relativity theory this presentation will focus on space-times.

Using (109) the right hand side is

$$(X^a{}_{,b} + X^f\Gamma^a{}_{bf})_{,c} - (X^a{}_{,c} + X^f\Gamma^a{}_{cf})_{,b} + (X^e{}_{,b} + X^f\Gamma^e{}_{bf})\Gamma^a{}_{ce} -$$

$$-(X^e{}_{,c} + X^f\Gamma^e{}_{cf})\Gamma^a{}_{be} - (X^a{}_{,e} + X^f\Gamma^a{}_{ef})\gamma^e{}_{cb}\,.$$

Carrying out the partial derivatives, this is

$$X^a{}_{,bc} + X^f{}_{,c}\Gamma^a{}_{bf} + X^f\Gamma^a{}_{bf,c} - X^a{}_{,cb} - X^f{}_{,b}\Gamma^a{}_{cf} - X^f\Gamma^a{}_{cf,b} + X^e{}_{,b}\Gamma^a{}_{ce} +$$

$$+X^f\Gamma^e{}_{bf}\Gamma^a{}_{ce} - X^e{}_{,c}\Gamma^a{}_{be} - X^f\Gamma^e{}_{cf}\Gamma^a{}_{be} - X^a{}_{,e}\gamma^e{}_{cb} - X^f\Gamma^a{}_{ef}\gamma^e{}_{cb}$$

Now it is useful to recall here that in the case of a function, the second covariant derivatives are the same:

$$f_{;bc} = f_{;cb} \quad \Leftrightarrow \quad f_{,bc} - f_{,cb} = \gamma^c{}_{ab} f_{,c} \tag{183}$$

(see (125), (129)); accordingly the first and fourth terms cancel with the second last one (on applying (183) to the function X^f; note that the components of a vector or tensor are indeed just functions, and as far as *partial derivatives* are concerned, the usual rules apply). Also the second and the fourth from last terms cancel, as do the fifth and seventh terms. We are left with just

$$X^f(\Gamma^a{}_{bf,c} - \Gamma^a{}_{cf,b} + \Gamma^e{}_{bf}\Gamma^a{}_{ce} - \Gamma^e{}_{cf}\Gamma^a{}_{be} - \Gamma^a{}_{ef}\gamma^e{}_{cb})\,.$$

Thus finally we obtain the *Ricci identity*

$$X^a{}_{;bc} - X^a{}_{;cb} = X^f R_f{}^a{}_{bc} \tag{184}$$

where we have defined

$$R_f{}^a{}_{bc} := \Gamma^a{}_{bf,c} - \Gamma^a{}_{cf,b} + \Gamma^e{}_{bf}\Gamma^a{}_{ce} - \Gamma^e{}_{cf}\Gamma^a{}_{be} - \Gamma^a{}_{ef}\gamma^e{}_{cb}\,; \tag{185}$$

this is the (Riemann) *Curvature tensor* for the space-time.

Now it is not obvious from (185) that $R_f{}^a{}_{bc}$ are components of a tensor. However if we remember the tensor detection theorem (section 2.2.1), then it is clear from (184) that this is indeed so (the left hand side is always a tensor, by construction, for arbitrary vectors X^f, so the appropriate version of that theorem applies). This tensor contains all the information on space-time curvature, as we will see in the following; in particular it vanishes if and only if space-time is locally flat. Before developing the geometry of the curvature tensor, we obtain the Ricci identity for general tensors, and determine its symmetries.

7.1.1 Ricci identity for tensors

Equation (184) holds for any contravariant vector (index upstairs). If we lower the index, because the metric can be take into the covariant derivative with impunity (cf. (120)), the same equation holds also for covariant vectors (index downstairs):

$$X_{a;bc} - X_{a;cb} = X^f R_{fabc} = X_f R^f{}_{abc} \tag{186}$$

Now one could obtain the corresponding equation for a tensor with two indices by proceeding from first principles (as in the previous section), but it is more convenient to deduce the result from what we already know. Consider a tensor T_{ab} that is a product of vectors: $T_{ab} = W_a Y_b$. Then

$$T_{ab;cd} = (W_{a;c}Y_b + W_a Y_{b;c})_{;d} = W_{a;cd}Y_b + W_{a;c}Y_{b;d} + W_{a;d}Y_{b;c} + W_a Y_{b;cd}$$

Hence, subtracting the corresponding equation with c and d interchanged,

$$T_{ab;cd} - T_{ab;dc} = (W_{a;cd} - W_{a;dc})Y_b + W_{a;c}Y_{b;d} - W_{a;d}Y_{b;c} +$$

$$+ W_{a;d}Y_{b;c} - W_{a;c}Y_{b;d} + W_a(Y_{b;cd} - Y_{b;dc})$$

Now the first and last terms can be re-expressed in terms of the curvature tensor, on using (186), and the other terms cancel in pairs: so we get

$$T_{ab;cd} - T_{ab;dc} = W_f R^f{}_{acd}\, Y_b + W_a\, Y_f R^f{}_{bcd} = T_{fb} R^f{}_{acd} + T_{af} R^f{}_{bcd}$$

Because of the linearity of this equation, it will be true not only for tensors built of products of vectors, but also for all linear combinations of such tensors, and so it will hold for all tensors with two indices down: so generally,

$$T_{ab;cd} - T_{ab;dc} = T_{fb} R^f{}_{acd} + T_{af} R^f{}_{bcd} \tag{187}$$

Looking at (183), (184), (186), (187), the general rule is clear: on commuting second covariant derivatives for any tensor, we get a curvature-tensor term for each index (thus dealing in turn with the effect of the change of order of derivatives on each tensor index; this is analogous to what happens in the formulae for the covariant derivative, see (115); however here the sign is always positive). In the term for a particular index, that index position is summed onto the curvature tensor (first position), the index itself being transferred to the curvature tensor (second position); while the derivative terms always occur in the third and fourth positions, the indices being ordered as in the positive term on the left-hand side.

7.1.2 Symmetries

There are four important symmetries of the curvature tensor.

1] From (184) or (185), the curvature tensor is skew in the last pair of indices:

$$R^f{}_{abc} = -R^f{}_{acb} \quad \Leftrightarrow \quad R^f{}_{abc} = R^f{}_{a[bc]} \tag{188}$$

2] The curvature tensor is also skew in the first pair of indices. To see this, apply (187) to the metric tensor g_{ab}. On the left, we find $g_{ab;cd} - g_{ab;dc} = 0$ (because $g_{ab;c} = 0$). On the right, we then find $0 = g_{fb}R^f{}_{acd} + g_{af}R^f{}_{bcd}$. The metric tensor lowers the indices on the curvature tensors to give

$$R_{abcd} = -R_{bacd} \quad \Leftrightarrow \quad R_{abcd} = R_{[ab]cd} \tag{189}$$

3] The curvature tensor obeys a cyclic identity like (162). To see this, apply (187) to the case when X_a is the gradient of some function ϕ. Then

$$X_a = \phi_{,a} \;\Rightarrow\; X_{[a;b]} = 0 \;\Rightarrow\; X_{[a;b]c} \equiv X_{[a;b];c} = 0 \;\Rightarrow\; X_{[a;bc]} = 0$$

Consequently the completely skew part of (186) shows

$$X_{[a;bc]} - X_{[a;cb]} = 0 = \phi_{;f} R^f{}_{[abc]}$$

for every function $\phi(x^j)$. It follows (choose $\phi = x^k \Rightarrow \phi_{;f} = \phi_{,f} = \delta^k_f$) that for all k,

$$R^k{}_{[abc]} = 0 \quad \Leftrightarrow \quad R^k{}_{abc} + R^k{}_{cab} + R^k{}_{bca} = 0 \tag{190}$$

the skew symmetry (188) leading to the second form from the first in the same way that (162) led to (165). Using the identity (190) in the Ricci identities (186) shows that every vector field satisfies the equations $X_{[a;bc]} = 0$.

4] Finally, given the symmetries (188-190), it follows that the curvature tensor is symmetric under interchange of the pairs of skew indices. This is shown by the following chain of argument:

$$R_{abcd} = -R_{adbc} - R_{acdb} = +R_{dabc} + R_{cadb} =$$

$$= -(R_{dcab} + R_{dbca}) - (R_{cbad} + R_{cdba}) = R_{cdab} + R_{bdca} + R_{bcad} + R_{cdab} =$$

$$= 2R_{cdab} - R_{badc} = 2R_{cdab} - R_{abcd}$$

where we have successively used: the cyclic symmetry, skew symmetry twice, cyclic symmetry twice, skew symmetry (in each term), cyclic symmetry (second and third terms), and skew symmetries (last term). In the final form, the right hand term is (minus) the same as the one we started with; so finally we get

$$R_{abcd} = R_{cdab} \tag{191}$$

the final Riemann tensor symmetry.

7.1.3 The contractions of the curvature tensor

In formulating the Einstein Field Equations, it will be important to note the following: because of the symmetries just described, there is only one non-trivial contraction of the curvature tensor. This is the *Ricci tensor*, defined by

$$R_{bd} := R^a{}_{bad} \quad \Rightarrow \quad R_{bd} = R_{db} \tag{192}$$

where the symmetry follows from (191) above. Note that the contraction here is on the first and third indices; because of the symmetries (188), (189), any other contraction of the Riemann tensor is either zero (e.g. $R^a_{acd} = 0$) or the same as the Ricci tensor (up to a sign), e.g. $R^a{}_{bca} = -R^a{}_{bac} = -R_{bc}$. For later use, it is convenient to note the expression for the Ricci tensor in terms of the connection coefficients: from (192) and (185),

$$R_{ab} = R^s{}_{asb} = \Gamma^s{}_{ba,s} - \Gamma^s{}_{sa,b} + \Gamma^e{}_{ba}\Gamma^s{}_{se} - \Gamma^e{}_{sa}\Gamma^s{}_{be} - \Gamma^s{}_{ea}\gamma^e{}_{sb}\,. \tag{193}$$

The *Ricci scalar* is the scalar formed by contracting the Ricci tensor: that is

$$R := R^a{}_a = g^{ab}R_{ab} = R^{ab}{}_{ab} \tag{194}$$

This is also known as the curvature scalar.

7.1.4 Dimensionality of the curvature tensor

Because of these symmetries, the curvature tensor in four dimensions has many less independent components than the $4^4 = 256$ components of a general 4-index tensor. If we label each skew index pair of R_{abcd} by a capital letter ($ab \to A$, $cd \to B$) then these new indices take 6 values (A, B run from 1 to 6, where for example we use the correspondence $12 \to 3$, $23 \to 1$, $31 \to 2$, $01 \to 4$, $02 \to 5$, $03 \to 6$). Thus we can represent the curvature tensor by a 6×6 matrix R_{AB}. Now this matrix is symmetric (by (191)), so symmetries (188), (189), (191) reduce the number of independent components of $R^a{}_{bcd}$ to 21 (corresponding to a 6 dimensional symmetric matrix). Finally (190) entails one more relation between the curvature tensor components, not contained in the symmetries (188,189,191), namely the relation $R_{dabc} + R_{dcab} + R_{dbca} = 0$ where each of a, b, c, d take separate values. Thus in 4 dimensions the curvature tensor has 20 independent components.

Now in 4 dimensions the Ricci tensor, being symmetric, has 10 independent components; and of course the Ricci scalar, only 1. If we consider the dimensions up to 4, the number of independent components are as follows:

dimension	R_{abcd}	R_{ab}	R
n = 2	1	1	1
n = 3	6	6	1
n = 4	20	10	1

Thus in two dimensions, the curvature is completely determined by the curvature scalar, while in three dimensions, it is completely determined by the Ricci tensor. Thus four is the smallest number of dimensions where the curvature tensor is not fully determined by the Ricci tensor. Thus has profound implications for the nature of gravity in different dimensions, as we will see in the next chapter.

7.2 Geometry of curvature

There are two main ways of seeing the meaning of curvature: through considering parallel transport round closed curves, or through geodesic deviation. We look at these in turn.

7.2.1 Parallel transfer

One manifestation of curvature is that parallel transfer of a vector round a closed loop is not integrable. A straightforward but somewhat tedious calculation (somewhat like that leading to equation (133)) shows the following: if a vector X^a is transported parallely round a surface element $\delta F^{cd} := V^{[c}W^{d]}$ defined by vectors V^a and W^b, then when arriving back at the initial point X^a will differ from its original value there by the amount

$$\delta X^a = R^a{}_{bcd} X^b \delta F^{cd} \tag{195}$$

This equation can be integrated round any closed loop, to show that the parallel transfer is not integrable by the amount of curvature contained in the loop (see e.g. Synge and Schild for details). We will not pursue this further here, except to note its application in flat space.

7.2.2 Flat space time

Suppose a flat space-time is given in the coordinates (12). Then (using the associated coordinate basis), because the metric components are constant, the Christoffel symbols vanish (from (131)) and so the curvature tensor is zero in these coordinates (from (185)). Thus the curvature tensor will vanish in all coordinates in such a space-time (because it is a tensor: if it is zero in one coordinate system, it is zero in all).

Conversely, suppose the curvature tensor vanishes everywhere in a space-time. By (195), parallel transfer will be integrable round all closed loops. Thus provided the space-time is simply connected (that is, all closed loops can be smoothly shrunk to zero), we can construct a global Minkowski coordinate system from a set of orthonormal vectors at one event. Essentially, parallel transferring them to every point along any curve, we establish a global set of orthonormal vectors; no inconsistency arises in so doing, no matter what path is chosen to move the vectors from one point to another, because parallel transfer is integrable in this case. Then along each curve with tangent vector X^a, we have $X^b\Gamma^c{}_{ba} = 0$ (equation (105)); as this is true for all X^a at all points, the $\Gamma^a{}_{bc}$ are all zero everywhere, hence so are the $\gamma^a{}_{bc}$ (see (128)), so the basis so established is a coordinate basis. Since any curve may be used for this parallel transfer, we may as well use geodesics for the purpose; then in effect the normal coordinate relations (177) have been extended to the whole space-time, showing the global existence of a coordinate system (12).

The conclusion is that, provided the space-time is simply connected, it is flat (and allows a Minkowski coordinate system) if and only if the curvature tensor vanishes everywhere. If it is not simply connected, so one cannot shrink some family of loops to zero, then a space-time with vanishing curvature tensor will be locally but not globally flat. This will be the case, for example, if there are conical-like singularities (such as are sometimes used to model infinitely thin cosmic strings) in an otherwise flat space-time.

The vanishing of space-time curvature is an example of an integrability condition: it is the locally necessary and sufficient condition for the existence of a family of parallel vector fields (they are parallel to each other, as defined by parallel transport along a curve, no matter what curve you consider between two points). However this is a very special situation: if there is any curvature, 'parallel vector fields' cannot be defined, for 'parallel at a distance' has no unique meaning; it is a path-dependent concept.

7.2.3 Geodesic deviation

Consider a 1-parameter family of geodesics $x^a(v,w)$ where w labels the geodesics, and v is an affine parameter along each of them. Then $V^a = dx^a/dv$ is a geodesic vector field:

$$V^a{}_{;b}V^b = 0 \tag{196}$$

The vector $W^a = dx^a/dw$ is tangent to the curves $\{v = const\}$ joining the geodesics, and commutes with their tangent vector V^a; that is (cf. (133,134))

$$[V,W] = 0 \quad \Leftrightarrow \quad V^a{}_{;b}W^b = W^a{}_{;b}V^b \tag{197}$$

Thus the two vector fields exactly fit together (W^a being dragged into itself by V^a). We call W^a a *geodesic deviation vector*, as it represents the displacement between neighbouring geodesics in the family. From these equations it follows that $(V^aW_a)_{;b}V^b = V^aW_{a;b}V^b = V^aV_{a;b}W^b = \frac{1}{2}(V^aV_a)_{;b}W^b$. Thus if the geodesics are normalised so that their magnitude is everywhere constant (as will be the case if proper time or proper distance is used as the affine parameter), then V^aW_a is constant along each geodesic. The natural choice is to set this constant to zero (e.g. choose the vectors initially orthogonal); then W^a is an orthogonal connecting vector for the normalised family of geodesics generated by V^a, and its magnitude gives the distance between neighbouring geodesics.

Now using the notation $\delta Y^a/\delta v = Y^a{}_{;b}V^b$ (cf. (108)), we have

$$\begin{aligned}
\frac{\delta^2 W^a}{\delta v^2} &= (W^a{}_{;b}V^b)_{;c}V^c = (V^a{}_{;b}W^b)_{;c}V^c \\
&= V^a{}_{;bc}W^bV^c + V^a{}_{;b}W^b{}_{;c}V^c \\
&= (V^a{}_{;cb} + V^dR_d{}^a{}_{bc})W^bV^c + V^a{}_{;b}W^b{}_{;c}V^c \\
&= (V^a{}_{;c}V^c)_{;b}W^b - V^a{}_{;c}V^c{}_{;b}W^b + V^dR_d{}^a{}_{bc}W^bV^c + V^a{}_{;c}W^c{}_{;b}V^b \\
&= V^dR_d{}^a{}_{bc}W^bV^c - V^a{}_{;c}(V^c{}_{;b}W^b - W^c{}_{;b}V^b)
\end{aligned}$$

where the first step develops the definition; the second uses (197); the third uses Leibniz rule for the derivative; the fourth uses the Ricci identity (184); the fifth uses Leibniz rule, and relabels the dummy indices in the last term (for later convenience); and the sixth uses (196) and collects terms together, so that (197) shows the last bracket vanishes. So finally

$$\frac{\delta^2 W^a}{\delta v^2} = V^dR_d{}^a{}_{bc}W^bV^c \tag{198}$$

which is the geodesic deviation equation: curvature directly controls the way that geodesics diverge from each other in space-time.

To get a feel for the meaning of this equation, consider flat space time. There

$$R_{abcd} = 0 \quad \Rightarrow \quad \delta^2 W^a/\delta v^2 = 0 \quad \Rightarrow \quad W^a = C^a v + K^a$$

where C^a and K^a are constant along the geodesics ($\delta C^a/\delta v = 0 = \delta K^a/\delta v$.) If $K^a = 0$ the geodesics all pass through the same point, diverging linearly with distance from it; if $C^a = 0$, they are parallel and forever remain at the same distance apart. Thus equation (198) is both at the base of the famous Euclid parallel axiom in flat space, and its failure in curved spaces and space-times (when a non-zero curvature term in (198) will show that this 'axiom' is no longer true).

In a curved space-time, this equation enables us to identify the curvature tensor $R^a{}_{bcd}$ as representing a tidal gravitational field, for it shows that space-time curvature directly influences the way that freely falling particles (moving on timelike geodesics) diverge from each other (just as a gravitational tidal force does). Furthermore as the equation also applies to null geodesics, it shows how light rays are focused and lensed by space-time curvature. These themes will be developed in the following chapter.

7.3 Integrability conditions

There are two important integrability conditions that are valid in every space-time.

7.3.1 The Jacobi identities

Consider arbitrary vector fields $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$. From (72), the following cyclic identity is always true:

$$[\mathbf{X},[\mathbf{Y},\mathbf{Z}]]+[\mathbf{Z},[\mathbf{X},\mathbf{Y}]]+[\mathbf{Y},[\mathbf{Z},\mathbf{X}]]=0 \tag{199}$$

One sees this by just developing the definitions: expanding the inner brackets first according to (72), the left hand side is

$$[\mathbf{X},\mathbf{YZ}-\mathbf{ZY}]+[\mathbf{Z},\mathbf{XY}-\mathbf{YX}]+[\mathbf{Y},\mathbf{ZX}-\mathbf{XZ}]$$

Now expanding the remaining brackets, this is

$$\mathbf{X}(\mathbf{YZ})-\mathbf{X}(\mathbf{ZY})-(\mathbf{YZ})\mathbf{X}+(\mathbf{ZY})\mathbf{X}+\mathbf{Z}(\mathbf{XY})-\mathbf{Z}(\mathbf{YX})-$$

$$-(\mathbf{XY})\mathbf{Z}+(\mathbf{YX})\mathbf{Z}+\mathbf{Y}(\mathbf{ZX})-\mathbf{Y}(\mathbf{XZ})-(\mathbf{ZX})\mathbf{Y}+(\mathbf{XZ})\mathbf{Y}\,;$$

the terms cancel in pairs (e.g. the first cancels with the seventh), giving (199), which is the *Jacobi identity.*

The useful application for us now is when we choose the vectors in (199) to be basis vectors; thus we set $\mathbf{X}=\mathbf{e}_a$, $\mathbf{Y}=\mathbf{e}_b$, $\mathbf{Z}=\mathbf{e}_c$, obtaining

$$[\mathbf{e}_a,[\mathbf{e}_b,\mathbf{e}_c]]+[\mathbf{e}_c,[\mathbf{e}_a,\mathbf{e}_b]]+[\mathbf{e}_b,[\mathbf{e}_c,\mathbf{e}_a]]=0\,. \tag{200}$$

From (129), this gives us

$$[\mathbf{e}_a,\gamma^d{}_{bc}\mathbf{e}_d]+[\mathbf{e}_c,\gamma^d{}_{ab}\mathbf{e}_d]+[\mathbf{e}_b,\gamma^d{}_{ca}\mathbf{e}_d]=0 \tag{201}$$

Now from (72) we find a useful general relation: for any vector fields $\mathbf{X}$, $\mathbf{Y}$ and function f,

$$[\mathbf{X},f\mathbf{Y}]=X(f)\mathbf{Y}+f[\mathbf{X},\mathbf{Y}]$$

(to see this, apply the left-hand bracket to an arbitrary function g). As the $\gamma^a{}_{bc}$ are just a set of functions, we can use this relation in (201). Remembering the notation $f_{,a}=e_a(f)$ (section (3.1)), we find

$$\gamma^d{}_{bc,a}\mathbf{e}_d+\gamma^d{}_{bc}[\mathbf{e}_a,\mathbf{e}_d]+\gamma^d{}_{ab,c}\mathbf{e}_d+\gamma^d{}_{ab}[\mathbf{e}_c,\mathbf{e}_d]+\gamma^d{}_{ca,b}\mathbf{e}_d+\gamma^d{}_{ca}[\mathbf{e}_b,\mathbf{e}_d]=0$$

Now using (72) again, and relabelling the dummy suffix 'd' in the derivative terms to 'f', we get

$$(\gamma^f{}_{bc,a}+\gamma^d{}_{bc}\gamma^f{}_{ad}+\gamma^f{}_{ab,c}+\gamma^d{}_{ab}\gamma^f{}_{cd}+\gamma^f{}_{ca,b}+\gamma^d{}_{ca}\gamma^f{}_{bd})\mathbf{e}_f=0$$

Consequently finally we have

$$\gamma^f{}_{bc,a}+\gamma^f{}_{ab,c}+\gamma^f{}_{ca,b}+\gamma^d{}_{bc}\gamma^f{}_{ad}+\gamma^d{}_{ab}\gamma^f{}_{cd}+\gamma^d{}_{ca}\gamma^f{}_{bd}=0 \tag{202}$$

which can be written in the shorter form

$$\gamma^f{}_{[bc,a]} + \gamma^d{}_{[bc}\gamma^f{}_{a]d} = 0 \tag{203}$$

A special case of this equation will be familiar to all those who have worked with Lie Algebras: when the γ's are constants, they are the structure constants of a Lie algebra, and obey the identity resulting from (203):

$$\gamma^f{}_{bc,a} = 0 \quad \Rightarrow \quad \gamma^d{}_{bc} = C^d{}_{bc}, \quad C^d{}_{[bc}C^f{}_{a]d} = 0$$

In general these quantities are non-constant. Then it is (203) which ensures that (using a general basis) the curvature tensor as represented by (185) satisfies the cyclic identities (190), and so also satisfies (191) (thus ensuring (193) is symmetric). Thus from this viewpoint, (190) are a form of the integrability conditions (200) (or equivalently, (203)) for the existence of the basis vectors $\mathbf{e}_a$.

7.3.2 The Bianchi identities

The curvature tensor in turn must satisfy integrability conditions, the *Bianchi Identities*. One convenient way to obtain them is to apply the Ricci identity in two different ways to the tensor $V_{ab} = X_{a;b}$ where the vector field X_a is arbitrary. From (186), taking a covariant derivative and then skew-symmetrising the resulting relation $(X_{a;bc} - X_{a;cb})_{;e} = (X_d R^d{}_{abc})_{;e}$, we find

$$2X_{a;[bce]} = X_{d;[e}R^d{}_{|a|bc]} + X_d R^d{}_{a[bc;e]} \tag{204}$$

where the vertical bars indicate that the index 'a' is omitted from the skew-symmetrisation. On the other hand, skew symmetrising equation (187) applied to V_{ab}: $(V_{ab})_{;ce} - (V_{ab})_{;ec} = V_{sb}R^s{}_{ace} + V_{as}R^s{}_{bce}$, we obtain

$$2V_{a[b;ce]} = V_{s[b}R^s{}_{|a|ce]} + V_{as}R^s{}_{[bce]} \tag{205}$$

Now by the definition of V_{ab}, the left hand sides of equations (204), (205) are the same. Equating the two right hand sides, the first term on the right of (204) cancels with the first on the right of (205) (they are even cyclic permutations of a totally skew quantity, because $ebc \rightarrow bec \rightarrow bce$ involves two swaps of adjacent indices). The second term on the right of (205) vanishes by (190). Hence the second term on the right of (204) vanishes for all X_a; consequently

$$R^d{}_{a[bc;e]} = 0 \quad \Leftrightarrow \quad R^d{}_{abc;e} + R^d{}_{aeb;c} + R^d{}_{ace;b} = 0 \tag{206}$$

which are the required identities. Their similarity to (162) is the basis for some of the analogies between gravitational theory and electromagnetism. Similarly to the transition from (162) to (164), we can define the (right) dual of the Riemann tensor: $R^*{}_{daef} = \frac{1}{2}R_{dabc}\eta^{bc}{}_{ef}$ (cf. (54)), and then rewrite (206) in the form $R^{*dafe}{}_{;e} = 0$ (similar to (164)).

7.3.3 Contracted Bianchi identities

The structure of the Einstein equations, considered in the next chapter, is fundamentally shaped by the *contracted Bianchi identities*, obtained from (206) on contracting twice. In detail, set $d = b$ and multiply by g^{ac} to obtain (on using (193), (194), and the skew symmetries (188), (189))

$$R^{bc}{}_{bc;e} + R^{bc}{}_{eb;c} + R^{bc}{}_{ce;b} = 0 \quad \Leftrightarrow \quad R_{;e} - R^{bc}{}_{be;c} - R^{cb}{}_{ce;b} = 0\,,$$

that is: $R_{;e} - R^{c}{}_{e;c} - R^{b}{}_{e;b} = 0$, and so

$$R_{;e} - 2R^{c}{}_{e;c} = 0 \quad \Leftrightarrow \quad (R^{c}{}_{e} - \frac{1}{2}Rg^{c}{}_{e})_{;c} = 0 \tag{207}$$

Note that this is valid for all curved spaces and space-times, independent of their dimensions.

7.4 Spaces/spacetimes of constant curvature

The simplest curved spaces and space-times are spaces of constant curvature. By definition, these are those spaces in which the curvature tensor takes the form

$$R_{abcd} = K(g_{ac}g_{bd} - g_{ad}g_{bc}) \tag{208}$$

which implies on contraction that

$$R_{bd} = (n-1)Kg_{bd}, \quad R = n(n-1)K \tag{209}$$

when the space has dimension n (so $g^{a}{}_{a} = \delta^{a}_{a} = n$). The right hand side of (208) has all the Riemann tensor symmetries, as required for this equation to make sense; and indeed this is the only combination of metric tensor terms that has all these symmetries.

A space or spacetime is a 'space of constant curvature' if it obeys (208) and $K = constant$. Now if we put (208) into the contracted Bianchi identities (207), we find $(n-1)(n-2)K_{;e} = 0$; consequently (208) implies $K = constant$ whenever $n \geq 3$, so such spaces are necessarily spaces of constant curvature. Two dimensions is the exceptional case[4]: then the curvature of necessity always has the form (208), but K can be non-constant.

Spaces of constant curvature have particularly simple geometry. If we put (208) into (195), we find that

$$\delta X^{a} = K(g^{a}{}_{c}g_{bd} - g^{a}{}_{d}g_{bc})X^{b}F^{cd} = 2KF^{a}{}_{b}X^{b}.$$

[4]Excluding 1-dimensions, where the curvature tensor is necessarily zero.

Then if $F_{ab}X^b = 0$ (the vector is perpendicular to the loop) $\delta X^a = 0$: there is no change on parallel transfer around it. However if X^a lies in the plane of the loop, it is rotated in that plane by an amount proportional to K and to the size of the loop.

If we put (208) into the geodesic deviation equation (198), we find

$$\frac{\delta^2 W^a}{\delta v^2} = V^d K(g_{db}g_c{}^a - g_{dc}g_b{}^a)V^c W^b = K(V^a(W^b V_b) - W^a(V^b V_b)).$$

Choosing normalized geodesic tangent vectors ($V^a V_a = \epsilon$) and orthogonal connecting vectors ($V^a W_a = 0$) as discussed above, we find

$$\frac{\delta^2 W^a}{\delta v^2} = -\epsilon K\, W^a, \quad \epsilon, K \; const \tag{210}$$

On using a parallel basis along the geodesics, the covariant derivatives become ordinary derivatives, and the equation has the solution

$$W^a = C^a f_1(v) + K^a f_2(v) \tag{211}$$

where C^a and K^a are covariantly constant along the geodesics ($C^a{}_{;b}V^b = 0 = K^a{}_{;b}V^b$) and the $f(v)$ are independent solutions of the simple harmonic equation $d^2 f(v)/dv^2 = -\epsilon K f(v)$. It is convenient to choose $f_1(v)$ to be $\sin \alpha v$, αv or $\sinh \alpha v$, where $\alpha^2 = \epsilon K$ is respectively positive, zero, and negative, with $\alpha^2 = \epsilon K|$, and $f_2(v)$ to be $\cos \alpha v$, $const$, or $\cosh \alpha v$. This shows how geodesics diverge or converge if $K \neq 0$ in spaces of constant curvature.

On using geodesic coordinates (176), equations (210,211) lead to the characteristic metric tensor form

$$ds^2 = \epsilon dr^2 + f^2(r) d\theta^2 \tag{212}$$

for the 2-dimensional metrics of constant curvature. To see this note that $W^a = \delta^a_2$ is a deviation vector for the geodesics generated by $V^a = \delta^a_1$ where $(x^1, x^2) = (r, \theta)$. Now either (a) evaluate the geodesic deviation equation in the obvious coordinate basis, working out the needed Christoffel symbols, or (b) introduce the parallel propagated orthonormal basis $e_1 = \frac{\partial}{\partial r}$, $e_2 = \frac{1}{f(r)}\frac{\partial}{\partial \theta}$ and evaluate the geodesic deviation equation in this frame, where $V^a = \delta^a_1$, $W^a = f(r)\delta^a_2$, and covariant derivatives along the geodesics are equivalent to ordinary derivtaives (as the basis is parallel along them), or (c) use the covariant results (211) to deduce that f in (212) takes the form of f_i in (211) (evaluate (211) in the parallel frame just mentioned).

Correspondingly

$$ds^2 = \epsilon dr^2 + f^2(r)(d\theta^2 + \sin^2\theta\, d\phi^2) \tag{213}$$

for the 3-dimensional metrics of constant curvature, where we usually choose $f(r)$ as the solution $f_1(r)$ which is zero when $R = 0$. The 4-dimensional space-times of constant curvature are the de Sitter and Anti-de Sitter spacetimes, and Minkowski spacetime.

7.5 Some consequences

We finally briefly note some consequences of the curvature equations and integrability conditions discussed in this chapter.

7.5.1 Maxwell equations revisited

In any space-time,

1] From the Ricci identities, identically

$$F^{ab}{}_{;ab} = 0 \quad \Rightarrow \quad J^a{}_{;a} = 0$$

i.e. charge conservation follows from the Maxwell equations.

2] In terms of the 4-potential (166), Maxwell's equations (161) are

$$\Phi^{a;b}{}_{;b} - \Phi^{b;a}{}_{;b} = J^a \quad \Leftrightarrow \quad \Box\Phi^a - (\Phi^b{}_{;b})^{;a} - \Phi^d R_d{}^b{}_{ba} = J^a$$

The Gauge condition $\Phi^b{}_{;b} = 0$ then gives the 'wave equation' form

$$\Box\Phi^a + \Phi^d R^a{}_d = J^a \tag{214}$$

which is not of itself of a minimal coupling form (for the curvature tensor couples directly to the potential), but comes from the minimal coupling form of the Maxwell equations (161) discussed above. This shows that sometimes we have to choose one particular form of the physical equations from several possibilities, before using the minimal coupling assumption. In principle this leads to some ambiguity, although in practice it does not seem a serious problem.

7.5.2 Killings equations revisited

1] The Jacobi identities (199) for arbitrary vectors implies the Lie derivative property:

$$L_{\mathbf{X}} L_{\mathbf{Y}} \mathbf{T} - L_{\mathbf{Y}} L_{\mathbf{X}} \mathbf{T} = L_{[\mathbf{X},\mathbf{Y}]} \mathbf{T}$$

holds for any tensor $\mathbf{T}$ (because this is true for functions and vectors). This is the reason that symmetry groups form a Lie algebra: for it shows that

$$L_{\mathbf{X}} \mathbf{T} = 0, \quad L_{\mathbf{Y}} \mathbf{T} = 0 \Rightarrow \quad L_{[\mathbf{X},\mathbf{Y}]} \mathbf{T} = 0$$

2] From the Ricci identity applied to a Killing vector, every Killing vector obeys the equation

$$\xi_{c;ba} = \xi^d R_{dabc} \tag{215}$$

[To see this, add 3 versions of (137), then apply the cyclic symmetry on page 87.] Together with Killing's equations, this shows that every Killing vector field is determined by the values of ξ_a and $\xi_{a;b} = \xi_{[a;b]}$ at any one point p in a space (taking covariant derivatives of (215) gives all higher derivatives at that point in terms of these quantities).

This in turn shows that the maximum number of Killing vectors possible in an n-dimensional space or space-time is $n + \frac{1}{2}n(n-1) = \frac{1}{2}n(n+1)$, n being the maximal dimension for the translations of p (generated by the Killing vectors that are non-zero at p) and $\frac{1}{2}n(n-1)$ the maximal dimension of the isotropy group of p (generated by the Killing vectors that are zero at p, but have non-zero first derivatives there; as these derivatives are skew, they generate Lorentz transformations there, cf. (94-97)). The space or space-time only has this maximal number of Killing vectors if it is a space of constant curvature.

Together these results show that the symmetry group of a space-time is a finite dimensional Lie group of dimension $r \leq \frac{1}{2}n(n+1)$, generated by the Lie algebra of Killing vectors.

3] From (137) it follows that every Killing vector commutes with geodesic vectors [space-time symmetries drag geodesics into themselves], and so from (215) that every Killing vector obeys the geodesic deviation equation. Consequently if we choose a deviation vector that is the same as a Killing vector at a point p and has the same first derivative, then since they both obey the same propagation equation, the Killing vector *is* the deviation vector for a particular family of geodesics. These are the geodesics formed by choosing an initial geodesic γ and then dragging it into a 1-parameter family of geodesics by means of the isometries generated by the Killing vector.

7.5.3 Normal coordinates

When one uses the geodesic deviation equation or Ricci identity to investigate normal coordinates (176), one finds that now the curvature tensor enters explicitly at quadratic order (see e.g. Stephani, pp. 22-23 and 54-56). Thus spacetime is locally Minkowskian at any point (we can choose an orthonormal basis there) and to first order in distance from the point (as the Christoffel symbols vanish there, in normal coordinates); the curvature enters at second order, in terms of this distance.

Examples 6: CURVATURE

1.(i) Show that $X^a{}_{;bc} - X^a{}_{;cb} = X^e R_e{}^a{}_{bc} \Rightarrow X_{a;bc} - X_{a;cb} = X_e R^e{}_{abc}$. What is the corresponding result for tensors of type 2 (e.g. T^{ab}) ?

2. Suppose $R_{abcd} = K(g_{ac}g_{db} - g_{ad}g_{bc})$, where K is some function. Check that the right hand side has the correct symmetries for a curvature tensor. Find R_{ac} and R, and prove that K is a constant if $n > 2$, where n is the dimension of the space.

3. Coordinates such that $ds^2 = dr^2 + f^2(r)(d\theta^2 + sin^2\theta d\phi^2)$ are used in the space considered in problem 2, where $\lim_{r\to 0} f(r) = r$. Show that $W = \frac{\partial}{\partial\theta}$ (i.e. $W^a = \delta^a_\theta$) is a geodesic deviation vector [You may use the result of problem 4 of the previous example set]. Find $f(r)$ by the use of the geodesic deviation equation (198). [Hint: this is easiest to do by using a parallely propagated basis e_α. Substitute from Problem 2 for the form of R_{abcd}].

4*. (i) Prove from an expression for the components $R^a{}_{bcd}$ of the curvature tensor in terms of $\Gamma^a{}_{bc}$ and $\gamma^a{}_{bc}$ when a general basis is used, that the Ricci tensor is symmetric. [Hint: Start with (202), and use (128), (193), (123), and one other equation. You don't need to use (131).]

---

8 The Einstein Field Equations

The final fundamental issue is what are the equations that determine the space-time geometry. Einstein's basic idea was that *matter determines the space-time structure*, as it should determine the metric tensor (cf. section 5.3.2) . The question is, what are the equations that describe how this happens.

Now we would expect that when we use a coordinate description of space time (and associated coordinate basis), we would find some kind of second order equations for the metric components that encode the gravitational field (in order that we obtain the second order equations of the Newtonian limit). Through the Christoffel relations (131)[5]

$$\Gamma^a{}_{ed} = g^{ac}\Gamma_{ced}, \quad \Gamma_{ced} = \frac{1}{2}\{g_{cd,e} + g_{ec,d} - g_{de,c}\} \tag{216}$$

and the definition of the curvature tensor (6.3)

$$R_f{}^a{}_{bc} = \Gamma^a{}_{bf,c} - \Gamma^a{}_{cf,b} + \Gamma^e{}_{bf}\Gamma^a{}_{ce} - \Gamma^e{}_{cf}\Gamma^a{}_{be} \tag{217}$$

it is clear that the curvature tensor components are second-order in terms of the metric tensor components. Hence it will be consistent with the idea of representing gravity geometrically to suppose (following Einstein) that *matter causes space-time curvature*. This is made even more plausible when one considers the application of the geodesic deviation equation (198) to timelike geodesics representing freely falling objects. It is clear that the distance apart of these objects (represented by an orthogonal geodesic deviation vector W^a), regarded as a function of proper time along the particle world lines, is controlled by a second-order equation (namely (198)), just as the corresponding Newtonian tidal force is. Thus curvature represents the tidal effects of gravitation; and so should be determined by the distribution of matter in space-time through suitable equations. The problem then is to elucidate what representation to use for matter on the one hand, and the space-time curvature on the other.

Now there is only one tensor characteristic of all types of matter, irrespective of what form it takes; and that is its energy-momentum tensor T_{ab} (see section 5.1.2), which is symmetric (section 3.3.4) and conserved (by 153)):

$$T_{ab} = T_{(ab)}, \quad T^{ab}{}_{;b} = 0 \tag{218}$$

This does not easily fit with the full curvature tensor (217), for that has 4 indices. However the Ricci tensor (193)

$$R_{ab} = R^s{}_{asb} = \Gamma^s{}_{ba,s} - \Gamma^s{}_{sa,b} + \Gamma^e{}_{ba}\Gamma^s{}_{se} - \Gamma^e{}_{sa}\Gamma^s{}_{be} \tag{219}$$

[5] The γ's vanish because we are using a coordinate basis

is promising because it also represents the space-time curvature but like T_{ab} is a symmetric 2-index tensor ($R_{ab} = R_{(ab)}$, see (192)).

Thus one proposal for the gravitational equations might be

$$R_{ab} = \kappa T_{ab} \tag{220}$$

where κ is a coupling constant. This was indeed proposed by Einstein initially. However it will not work because of the contracted Bianchi identities (207). Taking the divergence of (220), by (218)

$$R^{ab}{}_{;b} = 0 \;\; \Rightarrow \;\; R_{,a} = 0 \;\Leftrightarrow\; R = const,$$

the implication following from (207). Because (220) implies $R = \kappa T^a{}_a$, the only solutions allowed will be those where the matter tensor has a constant trace $T := T^a{}_a$; and this will be far too restrictive to represent ordinary physical situations (for example, in the perfect fluid case (154) this will imply that $3p - \mu$ is always a constant, clearly unrealistic).

The problem has arisen from the identities (207), and its elegant solution also arises from those equations. Define the *Einstein tensor* by

$$G_{ab} = R_{ab} - \frac{1}{2} R g_{ab} \;\; \Rightarrow \;\; G_{ab} = G_{(ab)}, \tag{221}$$

then (207) is just the statement

$$G^{ab}{}_{;b} = 0 \tag{222}$$

Thus if we replace (220) by

$$G_{ab} = \kappa T_{ab} \tag{223}$$

we now find that the Einstein tensor symmetry (see (221)) guarantees the symmetry of T_{ab}, and its vanishing divergence (222) guarantees energy-momentum conservation; that is, the field equations (223) ensure equations (218) are identically true. Hence (223) do not suffer from the problems afflicting (220), and these equations, the *Einstein field equations*, are a satisfactory candidate for the equations representing how gravitational effects arise through matter causing space-time curvature. Indeed they are fully consistent and also are in accord with all experimental tests that have been carried out to date.

There is however one complication. When looking for static cosmological solutions, Einstein discovered a generalised set of equations that share the desirable properties of (223); namely

$$G_{ab} + \Lambda g_{ab} = \kappa T_{ab} \tag{224}$$

where Λ is called the *cosmological constant*. Provided Λ is indeed constant, $(\Lambda g^{ab})_{;b} = 0$ (by (119)), so equally (224) implies that (218) is true.

There has been a long historical controversy over whether Λ is zero or not. Physically it is equivalent to a long range repulsive force proportional to distance and acting on all matter equally. Its value must be small in order that we do not detect its effect in the solar system, but it could still be important in cosmology. In modern terms it gives the effect of the vacuum on space-time curvature, for in the vacuum case $T_{ab} = 0$, (224) shows that $G_{ab} = -\Lambda g_{ab}$. If we take the Λ-term to the right hand side of the field equations in this way to give an effective matter term, and then compare it with the perfect fluid stress tensor (154), we can regard it as equivalent to a perfect fluid with $\kappa p = -\Lambda = const$, $\mu + p = 0$ (and with arbitrary 4-velocity).

It is possible that Λ is non-zero. However it is consistent with all solar system, astrophysical, and cosmological observations that it is zero [6]. In the rest of these lectures we therefore follow the Occam's razor (simplest theory) path of assuming Λ is zero. The resulting equations are then the standard form of the Einstein Field Equations (EFE):

$$R_{ab} - \frac{1}{2} R g_{ab} = \kappa T_{ab} \tag{225}$$

Because space-time is 4-dimensional, contracting (10) shows $R - \frac{1}{2} 4R = \kappa T \Rightarrow -R = \kappa T$, $T := T^a_a$, so (225) can equally be written in the form

$$R_{ab} = \kappa (T_{ab} - \frac{1}{2} T g_{ab}) \tag{226}$$

which is sometimes more convenient in practical calculations. In particular this form easily gives the *vacuum Einstein equations* : in an empty space-time,

$$T_{ab} = 0 \;\;\Rightarrow\;\; T = 0 \;\;\Rightarrow\;\; R_{ab} = 0 \tag{11a}$$

so vacuum solutions are characterised by the vanishing of the Ricci tensor.

The *geometrodynamics* school of thought maintains the vacuum form of the field equations is better founded and more fundamental than the equations (226), and examines the way that vacuum solutions can have properties one would expect to be associated with matter. However more the general standpoint is that (225,226) are the valid classical (i.e. non-quantum) gravitational field equations. As remarked above, not only are these equations consistent with

[6] with one possible exception: recent (distance,redshift) measurements for galaxies, based on observations of distant supernovae, may suggest a positive cosmological constant. However this is a marginal claim, not yet on an adequate observational basis.

the energy-momentum relations (218), they actually demand them. As the tensors in this equations are all symmetric, there are 10 independent components that must be satisfied ($R_{00} - \frac{1}{2}Rg_{00} = \kappa T_{00}$, $R_{01} - \frac{1}{2}Rg_{01} = \kappa T_{01}$, etc.), in contrast to the single equation ((231) below) of Newtonian Gravitational Theory.

From now on we accept the hypothesis that (225,226) are indeed the equations by which matter causes space-time curvature, and examine the resulting theory.

8.1 The Newtonian limit

Given that we know Newtonian Gravitational Theory (NGT) works very well on the scale of the solar system, it is important that we are able to derive NGT as an appropriate limit of the EFE. It is far from obvious that this can be done; but if we could not do so, the EFE would be disproved by the well-tested observations of motion of the planets in the Solar System. We will now obtain the Newtonian limit in the linearised (weak-field) static case.

Consider motion of matter at slow speeds in a weak static gravitational field. Then firstly, there exist coordinates in which the space-time metric nearly has the Minkowski form (12):

$$g_{ab} = \eta_{ab} + h_{ab}, \quad |h_{ab}| << 1, \tag{227}$$

where η_{ab} is defined by (11) and we set the speed of light $c = 1$ by appropriate choice of units of distance (for otherwise the 'smallness' of the metric perturbation h_{ab} has no meaning independent of the units chosen); so in fact $\eta_{ab} = diag(-1, 1, 1, 1)$. Secondly, in these coordinates matter moves slowly compared with the speed of light[7]:

$$|v^\alpha/c| = |dx^\alpha/dt| << 1 \tag{228}$$

Thirdly, all the metric derivatives are small in this coordinate system, and the gravitational field is taken as static:

$$|h_{ab,c}| << 1, \quad h_{ab,0} = 0 \quad \Rightarrow \quad |\Gamma^a{}_{bc}| << 1 \tag{229}$$

(the second condition is a sharpening of the first in the case of time derivatives). All three conditions are satisfied to a high accuracy in the solar system.

Now we have to check two separate things. Firstly that under these conditions we can express particle motion in terms of a gravitational potential Φ, as we can in NGT, in the form

$$d^2x^\alpha/dt^2 = -\delta^{\alpha\beta}\Phi_{,\beta}\,. \tag{230}$$

[7]Here and in what follows, Greek letters are spatial indices running from 1 to 3, while Latin letters are spacetime indices running from 0 to 3

Secondly, that the gravitational equations give the right form of the potential from a given matter distribution, i.e. that Φ satisfies the Poisson equation

$$\nabla^2\Phi = \Phi_{,\alpha\beta}\delta^{\alpha\beta} = 4\pi G\rho \tag{231}$$

where ρ is the matter density. We consider these in turn, systematically at all stages approximating by keeping only the largest terms in each equation.

8.1.1 The equations of motion

Consider a test particle moving in a gravitational field satisfying (227-229), under gravity and inertia alone. Then (section 5.3) it moves on a space-time geodesic, thus (section 5.2) obeying the equation

$$\frac{d^2x^a}{d\tau^2} + \Gamma^a{}_{bc}\frac{dx^b}{d\tau}\frac{dx^c}{d\tau} = 0 \tag{232}$$

where τ is proper time along the curve. Separating out the space and time terms in the summation,

$$\frac{d^2x^a}{d\tau^2} + \Gamma^a{}_{00}\frac{dx^0}{d\tau}\frac{dx^0}{d\tau} + 2\Gamma^a{}_{0\gamma}\frac{dx^0}{d\tau}\frac{dx^\gamma}{d\tau} + \Gamma^a{}_{\beta\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} = 0\,.$$

By the slow-motion assumption (228), we can ignore the 3rd and 4th terms, obtaining

$$\frac{d^2x^a}{d\tau^2} + \Gamma^a_{00}\frac{dx^0}{d\tau}\frac{dx^0}{d\tau} = 0\,.$$

Now $\Gamma^a{}_{00}$ is given by (1) to be

$$\Gamma^a{}_{00} = g^{ac}\frac{1}{2}\{g_{c0,0} + g_{0c,0} - g_{00,c}\} = g^{ac}\frac{1}{2}\{-h_{00,c}\} = -\eta^{ac}\frac{1}{2}h_{00,c} \tag{17a}$$

to first order in h_{ab}, where we have used (229), (227) and the fact that $g^{ab} = \eta^{ab}$ to the required accuracy[8]. Thus the linearised equation of motion is

$$\frac{d^2x^a}{d\tau^2} - \eta^{ac}\frac{1}{2}h_{00,c}\frac{dx^0}{d\tau}\frac{dx^0}{d\tau} = 0\,. \tag{233}$$

Putting $a = 0$ in this equation we find from (229)

$$\frac{d^2x^0}{d\tau^2} = \frac{d^2t}{d\tau^2} = 0 \quad\Rightarrow\quad \frac{dt}{d\tau} = \gamma = const \simeq 1 \tag{18a}$$

[8]In order to avoid many problems, one should insist that indices are raised and lowered with the full metric g^{ab}, g_{ab}, rather than the flat space metric; suitable approximations can be made as required after doing so. This means that $\eta^{ab} := g^{ac}g^{bd}\eta_{cd}$ is approximately, but not exactly, equal to η_{ab}; however in the present calculation, to the required accuracy the difference does not matter, so we can take $\eta^{\alpha\beta} = \delta^{\alpha\beta}$, $\eta^{0\alpha} = 0$, $\eta^{00} = -1$ here.

where in the last relation we have used the definition of γ (15) and the slow motion condition (228). This is just the result that, for the slow motion considered, proper time for the particle is the same as the (almost-Minkowski) coordinate time, to high accuracy. Using this result in the $a = \alpha$ component of (233),

$$\frac{d^2x^\alpha}{dt^2} = \delta^{\alpha\gamma}\frac{1}{2}h_{00,\gamma}\,. \tag{234}$$

Comparing this result with (230), we see that we have the same equation of motion as in NGT if we identify

$$\Phi = -\frac{1}{2}h_{00} \quad \Leftrightarrow \quad g_{00} = -(1+2\Phi)\,. \tag{235}$$

8.1.2 The gravitational equations

To obtain the gravitational field equation (231), consider now equation (226) with $a = 0$, $b = 0$:

$$R_{00} = \kappa(T_{00} - \frac{1}{2}Tg_{00}) \tag{236}$$

For the right hand side, we use the perfect fluid form (154) with vanishing pressure: using (233a),

$$T_{ab} = \rho u_a u_b,\ u^a \simeq (1, v^\alpha) \quad \Rightarrow \quad T = T^a{}_a = -\rho,\ u_a = g_{ab}u^b \simeq (-1, v^\alpha)$$

so (236) becomes

$$R_{00} \simeq \kappa(\rho u_0 u_0 + \frac{1}{2}\rho\eta_{00}) = \frac{1}{2}\kappa\rho\,, \tag{237}$$

where as usual in this calculation, we have kept only the highest order terms. The left-hand side comes from (219):

$$R_{00} = \Gamma^s{}_{00,s} - \Gamma^s{}_{s0,0} + \Gamma^e{}_{00}\Gamma^s{}_{se} - \Gamma^e{}_{s0}\Gamma^e{}_{b0}$$

Now the second term here is zero (as the solution is static) and the third and fourth are second order (being products of first-order terms), so we have

$$R_{00} = \Gamma^0{}_{00,0} + \Gamma^\kappa{}_{00,\kappa}$$

Using (232a) and (229) in this equation,

$$R_{00} = -(\delta^{\kappa\beta}\frac{1}{2}h_{00,\beta})_{,\kappa}\,.$$

Putting this in (237) and using $h_{00} = -2\Phi$ (from (235)),

$$\delta^{\kappa\beta}\Phi_{,\beta\kappa} = \frac{1}{2}\kappa\rho\,. \tag{238}$$

This is the same as (231) provided we make the identification

$$\kappa = 8\pi G\,. \tag{239}$$

Thus we have obtained the Newtonian equation (231) from (226), as desired, and also identified the constant κ in (226) in terms of the Newtonian gravitational constant G. This has been done using units such that the speed of light is 1; in general units, dimensional analysis shows that relation (239) takes the form $\kappa = 8\pi G/c^4$.

This result has come from the "(00)" EFE. What we have not done here is show that in the Newtonian limit, the other nine field equations are necessarily satisfied to the appropriate order when we linearise those equations also. Here, we will simply assume that is true. If this were not so, the Newtonian limit would not work out as required, because we would have to satisfy other equations in addition to the Poisson equation (238). The theory would not reduce to one effective gravitational field equation.

8.1.3 Implications

This has shown that the Einstein field equations, in a static, weak field, and slow-motion limit, lead to the Newtonian equations of gravity and Newtonian particle motions. This is essential in order that the Einstein theory be a viable theory of gravity (for the Newtonian description is a very good one, within its limits of applicability). It is a remarkable feature of the equations that the ten Einstein equations give the simple NGT form in an appropriate limit, thereby ensuring that all the evidence for NGT is also evidence supporting the Einstein theory.

However we should note the correct logic of the situation. There is a tendency of some to (implicitly at least) see the Newtonian theory (which is a good model of our daily experience) as the basic or more fundamental theory, with the Einstein theory giving higher order perturbations to it in some circumstances. However seen from the relativistic viewpoint, this is the wrong way round. From this position, we assume that *the EFE are the fundamental equations of classical gravitation.* Then the reason the Newtonian equations (230), (231) are valid is precisely because they are a limiting form of the Einstein equations; they are valid as gravitational equations when and only when circumstances are such that appropriate approximations lead to (230), (231) from the true relativistic equations (232), (236).

That being said, it should be noted that the derivation above is on the face of it, more restricted than we in fact need in many practical cases. We use Newtonian theory in situations where the gravitational field is time-varying and non-linear (in consistent self-gravitating galactic models, for example, the particles making up the galaxy move in the gravitational field generated by their own cumulative gravitational effect, which therefore depends on their motion, and may additionally be time varying). Thus we use NGT in situations wider than those justified by the above derivation from the EFE. This is not to say that the derivation of NGT in non-linear situations cannot be provided, indeed the essential point is that one can have relatively large density perturbations while the metric perturbations remain quite small, so many of the approximations used above may remain true in those cases. The issue here is that without further analysis, it is not obvious the usual simple derivation of NGT from the EFE (outlined above), which is based on consistent linearisation throughout, gives a valid description when non-linearities or time dependence occur. Luckily the above analysis does indeed apply in the case of the Solar System, which is where we carry out our tests of the validity of NGT.

8.2 Properties of the field equations

Given that the EFE (225), (226) have a good Newtonian limit (at least in situations such as those in the Solar System), what are some of their more general properties, when we do not approximate the equations?

8.2.1 PDE's for the metric

Firstly, the EFE are a system of 10 coupled non-linear second order partial differential equations for the 10 metric components g_{ab}, as can be seen from (216) and (219). It should be noted here that the worst non-linearities come through the inversion of the metric: to calculate the Christoffel symbols through equations (216), we have to invert the metric g_{ab} to get the components g^{ab} from the relations (33): $g^{ab}g_{bc} = \delta^a_c$. If the metric is non- diagonal, this inversion will in general be a non-trivial operation. The complexity then comes because in (219) we take both derivatives and products of the g^{ab}.

Physically, the field equations are non-linear because matter curves space-time which then determines how the matter (which is curving the space-time) moves. We do not have any fixed background space-time within which this all takes place: the space-time which is the arena of physics and which shapes the form of the physical equations (by specifying the metric and the covariant derivatives) is itself changing in response to the outcome of those equations.

Given this non-linearity, it is perhaps surprising there are any solutions at all! The reason generic solutions exist is due to the combination of the form (225) of the EFE and the existence of the identities (218), (222). These show the field equations are consistent in the following sense.

8.2.2 Evolution from initial values

We suppose the EFE equations are true on an initial surface S given by $t = t_0$, and consider under what conditions they will remain true off this surface. Now the structure of the equations (hyperbolic equations if written in suitable coordinates) is that their characteristics are null rays; in physical terms, gravitational waves travel at the speed of light. Therefore we obtain a unique solution to the future of S from initial data on S only in the space-time region $D^+(S)$ called[9] the *Future Cauchy Development of* S, which is the region such that all past-directed timelike and null curves from each point in $D^+(S)$ intersect S (in brief: it is the future region of space-time such that all information arriving there at less than or equal to the speed of light has had to cross S, so conditions there are completely determined by data on S). The *past Cauchy development* $D^-(S)$ is similarly defined. The *Cauchy development* $D(S)$ of S is the union of these two regions and S, and so is the complete region of space-time that is determined by data on S.

Define the tensor $A^{ab} = G^{ab} - \kappa T^{ab}$. Then the EFE (225) are equivalent to the equations $A^{ab} = 0$. From (218), (222), A_{ab} has vanishing divergence:

$$A^{ab}{}_{;b} = 0 \quad \Leftrightarrow \quad A^{ab}{}_{,b} + A^{sb}\Gamma^a{}_{bs} + A^{as}\Gamma^b{}_{bs} = 0\,.$$

Separating out the time and space summations, this is

$$A^{a0}{}_{,0} + A^{a\beta}{}_{,\beta} + A^{s0}\Gamma^a{}_{0s} + A^{0\beta}\Gamma^a{}_{\beta 0} + A^{\nu\beta}\Gamma^a{}_{\beta\nu} + A^{a0}\Gamma^b{}_{b0} + A^{a\beta}\Gamma^b{}_{b\beta} = 0\,. \quad (240)$$

Now suppose that the 6 equations $A^{\alpha\beta} = 0$ are true in a space-time region V containing S and lying in the Cauchy development of S, while the 4 equations $A^{0a} = 0$ are true on the initial surface S. The form of this first-order linear set of differential equations (which give the time development of A^{a0} off the surface) guarantees a unique solution locally in V from given initial data for A^{a0} on S. However there is a solution of these equations given by $A^{a0} = 0$ in V, which of course implies $A^{\alpha 0} = 0$ on S. Thus given the initial conditions $A^{a0} = 0$ on S, the unique solution of these equations is this solution: $A^{a0} = 0$ in V. Hence we have shown the following: *if the 4 initial value equations* $A^{a0} = 0$ *are true on* S *and the 6 propagation equations* $A^{\alpha\beta} = 0$ *are true in* V, *then* $A^{a0} = 0$ *will hold in* V. Thus the latter 4 Einstein equations are first integrals of the other 6 equations (provided the energy-momentum conservation equations (218) are satisfied). This structure is helpful in actually solving the equations.

To see what initial data for the field equations is like, it is convenient to use geodesic normal coordinates based on S, leading to the metric form (176). The propagation equations $A^{\alpha\beta}$ turn out to be equations for $d^2h_{\alpha\beta}/dt^2$ in terms

[9]See the article by Tipler Clarke and Ellis in *General Relativity and Gravitation*, ed. A. Held, Plenum Press, for a discussion.

of the metric $h_{\alpha\beta}$ and its first derivatives $dh_{\alpha\beta}/dt$, $h_{\alpha\beta,\gamma}$, while the constraint equations turn out to be equations involving only $h_{\alpha\beta}$ and its first derivatives. Thus initial data for the EFE on S is

(a) the *first fundamental form* $h_{\alpha\beta}$ (just the intrinsic metric of that surface), plus

(b) the *second fundamental form* $\Theta_{\alpha\beta} = dh_{\alpha\beta}/dt$ (the derivative of the metric with respect to proper time measured along the orthogonal geodesics, which characterizes how the 3-space is imbedded in the 4-space), together with

(c) initial data for whatever matter fields may be present.

These data must be chosen to satisfy the constraint equations $A^{0a} = 0$; then the propagation equations $A_{\alpha\beta} = 0$ together with the evolution equations for the matter will determine the solution off S. The constraint equations will remain true off S if they are true on S, as just proved above. The solution will be determined by this data within the Cauchy development $D(S)$ of S, but not outside this space-time region.

8.2.3 The Free Gravitational field

The above section has considered the field equations as second order partial differential equations for the metric tensor components. However there is quite different way of looking at them.

In this view, we focus on the curvature tensor R_{abcd}. By (192), (194) the EFE (225,226) are just *algebraic* equations for linear combinations of this tensor. Now an interesting issue arises. There are 20 independent components of the curvature tensor R_{abcd} (see section 6.1.4). The EFE (226), algebraically specifying the Ricci tensor at each point (with 10 independent components), effectively determine 10 of the 20 Riemann tensor components at each space-time point. Where does the other curvature information reside?

The answer is that it is contained in the *Weyl tensor*, which we can regard physically as the free gravitational field at each point. It is defined by the following equation[10]:

$$R_{abcd} = C_{abcd} - \frac{1}{2}(R_{ac}g_{bd} - R_{ad}g_{bc} + R_{bd}g_{ac} - R_{bc}g_{ad}) + \frac{1}{6}R(g_{ac}g_{bd} - g_{ad}g_{bc}) \tag{241}$$

This equation expresses the curvature tensor in terms of a scalar part (the last bracket, of the constant curvature form (208)); a Ricci tensor part (the central bracket, containing the correct combination of Ricci tensor and metric parts to have all the Riemann tensor symmetries); and the Weyl tensor term C_{abcd}. As all the other terms in the equation have all the Riemann tensor symmetries,

[10]This form holds in 4 dimensions. In an n-dimensional space, replace 1/2 by $1/(n-2)$ and 1/6 by $1/(n-1)(n-2)$ in this equation.

so does the Weyl tensor; however additionally it has zero trace, i.e. it is the 'Ricci-free' part of the curvature tensor. To see this contract (241) to get

$$\begin{aligned} R_{bd} &= g^{ac}C_{abcd} + \frac{1}{2}(Rg_{bd} - R_{bd} + 4R_{bd} - R_{bd}) + \frac{1}{6}R(4g_{bd} - g_{bd}) \\ &= C^{a}{}_{abcd} + \frac{1}{2}(2R_{bd}) - \frac{1}{2}Rg_{bd} + \frac{1}{6}3Rg_{bd}\,. \end{aligned} \tag{242}$$

The Ricci tensor terms cancel, as do the scalar terms, so the trace of the Weyl tensor is zero. Thus (241) implies

$$C^{a}{}_{bad} = 0, \quad C_{abcd} = C_{[ab][cd]}, \quad C_{abcd} = C_{cdab}, \quad C_{a[bcd]} = 0 \tag{243}$$

This shows the Weyl tensor has 10 independent components, and we can regard the curvature tensor as made up algebraically of the Ricci tensor and the Weyl tensor through (26). In the vacuum case, the Weyl tensor and Ricci tensors are identical:

$$T_{ab} = 0 \quad \Rightarrow \quad R_{ab} = 0 \quad \Rightarrow \quad R_{abcd} = C_{abcd}\,. \tag{244}$$

Now the EFE determine R_{ab} and R in (241) directly at each point in terms of the matter present there. What about the Weyl tensor? It is not determined algebraically by the matter at each point p, but rather represents the combined effect of distant matter on the curvature at p; that is, it represents the tidal force there. Now how does distant matter succeed in affecting the matter at p? The answer is that the Bianchi identities (206) serve as propagation equations for C_{abcd} with source terms determined by the matter at each point, determining how it propagates through space-time from the source of the gravitational field to the point where its effect is felt. This is what makes possible both gravitational radiation and tidal forces. For example the matter at the Moon locally generates a Weyl tensor field, which then propagates through the intervening empty space to the Earth, where it is felt as tidal gravitational field. The Weyl tensor exerts its tidal affect locally on matter through the geodesic deviation equation, as discussed at the beginning of this chapter[11] (in a vacuum, the source term in (198) is just the Weyl tensor).

In more detail, (206) can be written as ${}^{*}R^{*abcd}{}_{;d} = 0$ where ${}^{*}R^{*abcd} = \frac{1}{4}\eta^{abst}R_{stuv}\eta^{cduv}$ (the double-dual of the Riemann tensor). With (241), this shows (after considerable work) that

$$C^{abcd}{}_{;d} = J^{abc}, \tag{245}$$

where

$$J^{abc} := R^{c[a;b]} - \frac{1}{6}g^{c[a}R^{,b]} \quad = \kappa(T^{c[a;b]} - \frac{1}{3}g^{c[a}T^{,b]})$$

[11] see section 5.3 of *Flat and Curved Space-times* for more details.

and always

$$C^{abcd}{}_{;dc} = 0 \quad \Rightarrow \quad J^{abc}{}_{;c} = 0\,. \tag{246}$$

In the vacuum case,

$$T_{ab} = 0 \;\Rightarrow R_{ab} = 0 \;\Rightarrow\; J^{abc} = 0 \quad \Rightarrow \quad C^{abcd}{}_{;d} = 0\,. \tag{247}$$

This has a striking similarity to the Maxwell equations (161) and their consequent current conservation equation $J^a{}_{;a} = 0$ (see 6.5.1). One can take the analogy further by defining the Electric and Magnetic parts of the Weyl tensor measured by an observer moving with 4-velocity u^a as

$$E_{ab} = C_{acbd}u^c u^d, \;\; H_{ab} = {}^*C_{acbd}u^c u^d\,; \tag{248}$$

(cf. (160)); these then turn out to both be symmetric, trace-free tensors orthogonal to u^a, that satisfy equations very similar in form to the electromagnetic Maxwell equations (see RC for details).

As has been emphasised above, it is the fact that the Riemann tensor is not completely determined by the matter at a point that allows gravitational effects to propagate through space-time so that matter here can have a gravitational effect on matter elsewhere (and vice versa). It is the existence of the Weyl tensor that makes this possible, justifying regarding it as the 'free gravitational field'. It is important to note that *this feature is possible because space-time has 4 dimensions*, for the table in section 6.1.4 shows that in 2 or 3 dimensions, all the information about local space-time curvature (as expressed by the curvature tensor) will be determined locally by the EFE, indeed at each point it is determined by the matter at that point. In low-dimension gravitational theories there is a possibility of non-integrability of parallel transfer due to non-local effects (e.g. the occurrence of conical singularities), but there is no room for a free gravitational field at each point and consequent tidal forces. Thus any such theories are missing a vital aspect of real gravity.

8.2.4 Variational Principle

One can obtain the gravitational field equations from a variational principle with

$$I = \int (R + 2\kappa L)dV \tag{249}$$

where R is the curvature scalar and L the matter Lagrangian (see e.g. pp. 91-94 of Stephani's book).

However this is of little practical use for the following reason.
a) if we vary I above for a general metric, we get the EFE (225); if we now choose a particular metric form (e.g. a static spherically symmetric metric as in

section 2.5.4) the equations (225) specialise to the EFE for that specific metric. These are the equations we have to solve to obtain specific solutions of the EFE with metric of the chosen form.

b) On the other hand we can calculate the curvature scalar R direct from the specific metric, and then put this into it (249) to get a variational principle for that metric. Performing the variation, we get another set of equations for metrics of the chosen form.

Now the problem is that *in general these two sets of equations are not the same*: performing the variation first and then specialising the metric does not commute with specialising the metric and then performing the variation. The reason is to do with the boundary terms in the variation, which we normally assume will go to zero. When we use method (b), in general they will not do so. Hence method (b), the obvious way to use the variational principle to simplify derivation of the field equations, may give the wrong answer.

Thus in practice one should only use the variational principle to get the field equations for a specific metric, with the greatest caution, checking in detail whether the surface terms vanish or not; otherwise the answer obtained could be wrong. Indeed to verify that the variational principle gives the right answer, one should check the answer obtained by method (b) with that found by method (a); but then one may as well just use the result obtained by (a).

8.2.5 Obtaining solutions

Two initial points:

Firstly, one should clearly state what matter content one is assuming for space-time (vacuum, perfect fluid, electromagnetic field, scalar field, etc.) when obtaining solutions, and also specify any needed equations of state for the chosen form of matter (for example in the perfect fluid case, one should state what relations hold between the energy density μ and the pressure p). Until this has been done, the EFE do not describe a well-defined physical situation.

Secondly, it is important to note that when obtaining solutions of the EFE, *one must always make certain that all 10 equations are satisfied.* Unless this has been verified, one is not in a position to claim one has a correct solution of the field equations.

Given a suitable matter description, three methods are commonly used to obtain solutions.

Firstly, a coordinate approach assumes some specific metric form (often on the basis of assumed symmetries, perhaps described by Killing's equations), calculates the Christoffel symbols and consequently field equations; and then solves the field equations. This is the traditional approach; for example in this way one can examine the family of spherically symmetric vacuum or fluid solutions.

Secondly a tetrad approach uses a tetrad to describe the geometry, calculating the rotation coefficients from (139), (131), then obtaining the Ricci tensor and the EFE. For consistency, the commutation coefficients must satisfy the identities (203). This is often a good approach to use in examining the consistency of the field equations when particular assumptions have been made about space-time geometry.

Thirdly, a tetrad description can be used with explicit introduction of the Weyl tensor components (241) also, and with explicit use of the Bianchi identities (206) as equations for the Weyl tensor components. The Newman-Penrose formalism essentially takes this approach. This is a good method to use for example in examining algebraically special solution (that is, space-times with particular simple forms of the Weyl tensor), or studying gravitational radiation.

The most important two sets of exact solutions are (i) the spherically symmetric fluid and vacuum solutions that underlie our basic understanding of gravitational collapse to form a black hole, see the Lecture Notes on Black Holes, and (ii) the Friedmann-Lemnaitre family of cosmological solutions based on the Robertson-Walker metrics, which are the standard models of cosmology. These are examined in the Lecture Notes on Cosmology.

Finally a covariant analysis of the problem at hand, using the tensor methods introduced in these lectures, can often give a great deal of useful information before having to introduce any specific coordinates or tetrads. When this approach is possible, it is most powerful because the results are true for all coordinate systems and tetrad bases (see e.g. J Ehlers: *Gen Rel Grav* **25**:1225-1266 (1993), and the forthcoming book by G F R Ellis, M A H MacCallum, and R Maartens, summarised in the Cargese Lectures by G F R Ellis and H van Elst, available on the Web via the UCT web pages).

* * *

Contents

1 Introduction **2**

2 Spacetime **4**
2.1 Events and coordinates . . . 4
2.1.1 Events . . . 6
2.1.2 Curves . . . 6
2.1.3 Surfaces . . . 7
2.1.4 Curve-Surface relations . . . 7
2.1.5 Coordinates, Curves and Surfaces . . . 8
2.2 Distances, Angles and Times . . . 9
2.2.1 Curved Spaces . . . 11
2.2.2 Flat Spacetime . . . 12
2.2.3 Curved Spacetimes . . . 13

3 Tensors **15**
3.1 Change of coordinates . . . 15
3.1.1 Functions . . . 15
3.1.2 Curves and vectors . . . 16
3.1.3 Gradient of Function . . . 18
3.1.4 Invariants . . . 19
3.2 Tensors and Tensor transformations . . . 20
3.2.1 Quotient Law: Tensor detection . . . 21
3.3 Tensor Equations . . . 22
3.3.1 Tensor operations . . . 23
3.3.2 Symmetry properties . . . 25
3.3.3 The volume element . . . 27

4 General Bases **32**
4.1 General basis . . . 34
4.1.1 Coordinate bases . . . 36
4.2 Tetrad bases . . . 37
4.2.1 Orthonormal bases . . . 37
4.2.2 Lorentz transformations . . . 38
4.3 Other tetrads . . . 39
4.4 Physics in an orthonormal basis . . . 40
4.4.1 4 velocity of particle . . . 40
4.4.2 The 4-momentum . . . 41
4.4.3 Electromagnetism . . . 42
4.4.4 Energy-momentum . . . 43
4.5 Lorentz Transformations . . . 45

5 Covariant Differentiation **48**
5.1 Parallel transport 48
5.1.1 Flat space 48
5.1.2 The Parallel Transport Operator 51
5.2 Curved space 52
5.2.1 Curved space-time 52
5.3 Covariant Differentiation 53
5.3.1 Vector fields 53
5.3.2 General Tensors 53
5.4 Relation to the metric 55
5.4.1 The derivative of the metric 56
5.4.2 Vanishing torsion 58
5.4.3 The Christoffel relations 59
5.4.4 Euclidean space quantities 60
5.5 The Lie Derivative 60

6 Physics in a curved space-time **64**
6.1 Force law 65
6.2 Conservation Equations 66
6.2.1 Scalar conservation 66
6.3 Energy-momentum mass conservation 69
6.4 Maxwell's equations 71
6.5 Scalar Field 73
6.6 Geodesics 74
6.6.1 Parametrisation 74
6.6.2 Existence 75
6.6.3 Geodesic coordinates 75
6.6.4 Extremal properties 77
6.6.5 Calculation of Gammas 79
6.7 Gravity 80
6.8 Physical meaning of parallel transfer 82
6.8.1 Gravity as geometry 82

7 Curvature **85**
7.1 The curvature tensor 85
7.1.1 Ricci identity for tensors 87
7.1.2 Symmetries 87
7.1.3 The contractions of the curvature tensor 89
7.1.4 Dimensionality of the curvature tensor 89
7.2 Geometry of curvature 90
7.2.1 Parallel transfer 90
7.2.2 Flat space time 90
7.2.3 Geodesic deviation 91
7.3 Integrability conditions 92

7.3.1 The Jacobi identities 93
7.3.2 The Bianchi identities 94
7.3.3 Contracted Bianchi identities 95
7.4 Spaces/spacetimes of constant curvature 95
7.5 Some consequences 97
7.5.1 Maxwell equations revisited 97
7.5.2 Killings equations revisited 97
7.5.3 Normal coordinates 98

8 The Einstein Field Equations 100
8.1 The Newtonian limit 103
8.1.1 The equations of motion 104
8.1.2 The gravitational equations 105
8.1.3 Implications 106
8.2 Properties of the field equations 107
8.2.1 PDE's for the metric 107
8.2.2 Evolution from initial values 108
8.2.3 The Free Gravitational field 109
8.2.4 Variational Principle 111
8.2.5 Obtaining solutions 112