Imperial College 4th Year Physics UG, 2013-14

General Relativity
Lecture notes

Toby Wiseman

October 3, 2013

Toby Wiseman; Huxley 507, email: `t.wiseman@imperial.ac.uk`

Office hour: 10-11 Monday, 10-11 Thursday

Rapid feedback class: 1-2 Tuesday with Andrew Hickling (please hand in work by Monday lunchtime in the UG office for marking)

During the course 8 example sheets will be handed out - one per week, starting week 3 (ie. Oct 22nd). Some questions will be for rapid feedback, some more advanced questions to develop understanding. By the conclusion of the course students will have the technical these example sheets, and these are to be considered an *essential* part of the learning process.

Books

This course is not based directly on any one book. Recommended reading for the course is;

- Sean Carroll, "Spacetime and Geometry"
  (Pearson, Addison Wesley)
- James Hartle, "Gravity: An introduction to Einstein's General Relativity"
  (Pearson, Addison Wesley)
- Ray D'Inverno, "Introducing Einstein's Relativity"
  (Oxford Univerity Press)
- Robert Wald "General Relativity"
  (Univ. Chicago Press)
- Steven Weinberg, "Gravitation and Cosmology"
  (Wiley)

I have largely tried to use the notation and conventions of Wald, which I regard as an excellent book. In particular the book goes significantly beyond this course, and so familiarity with it can be very useful long term. Also of interest in terms of Differential Geometry;

- Nakahara, "Geometry, Topology and Physics"
  (IOP)

1 Classical geometry

1.1 Recap: 3-d Euclidean space

Consider 3-dimensional Euclidean space, $\mathbb{E}^3$, with its canonical Cartesian coordinates $\{x, y, z\}$.

The line element (metric) on this space is,

$$ds^2 = dx^2 + dy^2 + dz^2 \tag{1}$$

Writing the coordinates as $x^i = \{x, y, z\}$ with $i = 1, 2, 3$, this defines the metric matrix,

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

where,

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3}$$

Then a vector $\mathbf{v}$ is a d-component object with components $v^i = \{v^1, v^2, v^3\} = \{v^x, v^y, v^z\}$ (in a slight bending of notation). We may compute the dot product of two vectors $\mathbf{v}$ and $\mathbf{w}$ by 'contracting them with the metric';

$$\mathbf{v} \cdot \mathbf{w} = g_{ij} v^i w^j = v^1 w^1 + v^2 w^2 + v^3 w^3 \text{ or } v^x w^x + v^y w^y + v^z w^z \tag{4}$$

For Euclidean space the norm of a vector, $\mathbf{v}$, gives its length or magnitude, and is,

$$|\mathbf{v}|^2 = g_{ij} v^i v^j = v^i v^i = (v^x)^2 + (v^y)^2 + (v^z)^2 \tag{5}$$

although we note that for general geometries a vector doesn't have a notion of length (although it does have a norm). We compute the angle between two vectors as,

$$\cos\theta = \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{v}||\mathbf{w}|} \tag{6}$$

and note that the angle between two vectors retains its meaning in general geometries.

1.1.1 Spherical coordinates

We may also take spherical polar coordinates $x'^{i'} = \{r, \theta, \phi\}$ on $\mathbb{E}^3$, where;

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta\cos\phi \\ z &= r\sin\theta\sin\phi \end{aligned} \tag{7}$$

(8)

In these coordinates the same line element (metric) becomes,

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2 \tag{9}$$

with metric matrix,

$$ds^2 = g'_{i'j'} dx^{i'} dx^{j'} \tag{10}$$

where,

$$g'_{i'j'} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix} \tag{11}$$

In different coordinates the values in g_{ij} and $g'_{i'j'}$ are different. However it is a very important point that this describes the same geometric quantity as before- i.e. the line element length. Quantities such as g_{ij} that have coordinate indexes and transform from one coordinate system to another are known as tensors. We shall define them carefully later.

Now a vector has components $v'^{i'} = \{v'^r, v'^\theta, v'^\phi\}$. Like the metric, the components of a vector in one coordinate system differ from those in another. Here the spherical components are related to the Cartesian components $\{v^x, v^y, v^z\}$ as,

$$\begin{pmatrix} v^x \\ v^y \\ v^z \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \end{pmatrix} \cdot \begin{pmatrix} v^r \\ v^\theta \\ v^\phi \end{pmatrix} \tag{12}$$

An important point is that; $v^i w^j g_{ij} = v'^{i'} w'^{j'} g'_{i'j'}$ - the length of a vector doesn't mind which coordinates one uses. (Ex. check this!)

1.1.2 Curves, tangents and straight lines

To state the very obvious, a straight line in $\mathbb{E}^3$ take a simple form in Cartesian coordinates being parametrized as,

$$x^i(\lambda) \;=\; a^i + \lambda b^i \tag{13}$$

where the two vectors a^i and b^i have *constant* components (ie. independent of position x^i) and give the offset and direction, and the variable λ parameterizes the line.

In general we may define a curve parametrically as $x^i = x^i(\lambda)$. The tangent vector $\mathbf{v}$ or 'speed and direction' to the curve is then the rate of change with respect to the parameter, having components,

$$v^i(\lambda) \equiv \frac{dx^i(\lambda)}{d\lambda} \tag{14}$$

For example; for the straight line above in Cartesian coordinates the tangent vector is constant with $v^i = b^i$ at any point on the line.

An important point is that this is *true in any coordinate system.* It is an example of a *tensor equation.* The values of the components of a tangent vector for a curve will differ depending on the coordinates - eg. Cartesian or Spherical. But the equation defining the tangent vector will be the same. As we will see all quantities in geometry can be phrased in terms of such tensor equations.

We say two curves that pass through the point $x^i_{(0)}$ are tangent if their tangent vectors at that point are a constant multiple of each other ie. they have the same direction (although not necessarily the same speed).

The distance traversed along a section of a curve between parameter λ_0 and λ_1, which we denote $s(\lambda_0, \lambda_1)$, is given by integrating the norm of the tangent vector,

$$s(\lambda_0, \lambda_1) = \int_{\lambda_0}^{\lambda_1} d\lambda \sqrt{g_{ij}(x(\lambda)) \frac{dx^i(\lambda)}{d\lambda} \frac{dx^j(\lambda)}{d\lambda}} \tag{15}$$

1.2 Classical Geometry of surfaces

In order to describe interesting geometries we may take the classical approach of considering the geometry induced on a curved surface in $\mathbb{E}^3$. Suppose we describe a surface by the embedding,

$$z = f(x, y) \tag{16}$$

for some function f at least over some range of the x, y coordinates. The function f must encode the geometry of this curved surface.

For example, we may embed the upper hemisphere of a unit 2-sphere as $z = \sqrt{1 - (x^2 + y^2)}$.

Now consider a curve that lies within this surface. Again we may parametrically describe it by $x^i = x^i(\lambda)$, but we note it is constrained to obey $z(\lambda) = f(x(\lambda), y(\lambda))$, and hence,

$$\frac{dz}{d\lambda} = \left.\frac{\partial f}{\partial x}\right|_y \frac{dx(\lambda)}{d\lambda} + \left.\frac{\partial f}{\partial y}\right|_x \frac{dy(\lambda)}{d\lambda} \tag{17}$$

using the chain rule.

Thus the tangent vector to the curve is;

$$v^i = \left\{ \frac{dx}{d\lambda}, \frac{dy}{d\lambda}, \frac{\partial f}{\partial x}\frac{dx(\lambda)}{d\lambda} + \frac{\partial f}{\partial y}\frac{dy(\lambda)}{d\lambda} \right\} \tag{18}$$

1.2.1 Tangent space to a surface

Consider a tangent vector to a curve at some point. Since we have not chosen any particular curve through that point, then $\frac{dx}{d\lambda}, \frac{dy}{d\lambda}$ are free for us to choose. Hence we may say that at any point in the surface, the tangent to a curve through that point takes the form,

$$v^i = \left\{ v^x, v^y, \frac{\partial f}{\partial x} v^x + \frac{\partial f}{\partial y} v^y \right\} \tag{19}$$

Any vector that is tangent to a curve in this surface – a *tangent vector* – is of the form above.

We say the set of all tangent vectors at a point on the surface is the *tangent space* to the surface. This tangent space is a 2-dimensional 'vector space'.

If one is given two tangent vectors to the surface at some point, u^i and v^i, these can be linearly combined to define a third vector w^i as,

$$\begin{aligned} u^i &= \left\{ u^x, u^y, \frac{\partial f}{\partial x} u^x + \frac{\partial f}{\partial y} u^y \right\} \\ v^i &= \left\{ v^x, v^y, \frac{\partial f}{\partial x} v^x + \frac{\partial f}{\partial y} v^y \right\} \\ &\rightarrow \quad w^i = \left\{ a u^x + b v^x, a u^y + b v^y, \frac{\partial f}{\partial x}(a u^x + b v^x) + \frac{\partial f}{\partial y}(a u^y + b v^y) \right\} \end{aligned} \tag{20}$$

where a, b are constants which we note is also tangent to the surface. Hence we see that we may think of a vector in this tangent space of the surface as being simply given by the 2-dimensional vector $\{v^x, v^y\}$ by the map,

$$\{v^x, v^y\} \rightarrow \left\{ v^x, v^y, \frac{\partial f}{\partial x} v^x + \frac{\partial f}{\partial y} v^y \right\} \tag{21}$$

and that linearly combining such 2-dimensional vectors $\{u^x, u^y\}$ and $\{v^x, v^y\}$ at the same point on the surface gives a third $\{a u^x + b v^x, a u^y + b v^y\}$ tangent vector in the surface. Formally we say that the tangent space at a point is then given by the vector space $\mathbb{R}^2$.

1.2.2 Induced metric on a surface

Now consider the dot product of two vectors in the surface. Denoting these as 2-dimensional vectors $\{u^x, u^y\}$ and $\{v^x, v^y\}$ then their dot product in $\mathbb{E}^3$ is,

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= u^x v^x + u^y v^y + \left(\frac{\partial f}{\partial x} u^x + \frac{\partial f}{\partial y} u^y \right) \left(\frac{\partial f}{\partial x} v^x + \frac{\partial f}{\partial y} v^y \right) \\ &= u^x v^x \left(1 + \left(\frac{\partial f}{\partial x} \right)^2 \right) + u^y v^y \left(1 + \left(\frac{\partial f}{\partial y} \right)^2 \right) + (u^x v^y + u^y v^x) \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \\ &= (u^x, u^y) \cdot \begin{pmatrix} 1 + \left(\frac{\partial f}{\partial x} \right)^2 & \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} & 1 + \left(\frac{\partial f}{\partial y} \right)^2 \end{pmatrix} \cdot \begin{pmatrix} v^x \\ v^y \end{pmatrix} \end{aligned} \tag{22}$$

Thus thinking about two tangent vectors $\mathbf{u}, \mathbf{v}$ at a point on the surface as 2-dimensional vectors u^a, v^a, with $a = 1, 2$, then we learn that we may compute their dot product,

$$\mathbf{u} \cdot \mathbf{v} = g_{ab} u^a v^b \tag{23}$$

where the 2×2 matrix g_{ab} is the *induced metric* on the surface,

$$g_{ab} = \begin{pmatrix} 1 + \left(\frac{\partial f}{\partial x}\right)^2 & \frac{\partial f}{\partial x}\frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial x}\frac{\partial f}{\partial y} & 1 + \left(\frac{\partial f}{\partial y}\right)^2 \end{pmatrix} \tag{24}$$

and is a *symmetric* matrix defined as a function over x, y.

Having a metric, it is natural to define its inverse. Let us define the inverse matrix g^{ab} to g_{ab}, so that,

$$g^{ab} g_{bc} = \delta^a_c \tag{25}$$

Comment: If the surface is embedded trivially as f =constant then the surface is just a 'flat' plane and the induced metric is simply $g_{ab} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, ie. the 2-dimensional Euclidean metric. If the surface is a plane we may always perform a rotation of the Euclidean space so that the embedding is simply f =constant.

1.2.3 A surface is locally a plane

Consider a surface in $\mathbb{E}^3$ and choose any point on that surface. Without loss of generality we may translate the surface (so we don't change its geometry) so that the coordinates of that point are $x = y = 0$. Perhap less obvious is that by appropriate rotation about the x and y axes we may generally arrange that the embedding function for the surface has the property,

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0 \quad \text{at} \quad x = y = 0 \tag{26}$$

Let us focus on some region of the surface at this chosen point $x = y = 0$. We may Taylor expand the embedding function f in the neighbourhood of this region about $x = y = 0$ as,

$$f(x,y) = f(0,0) + \left.\frac{\partial f}{\partial x}\right|_0 x + \left.\frac{\partial f}{\partial y}\right|_0 y \tag{27}$$

$$+\frac{1}{2}\left.\frac{\partial^2 f}{\partial^2 x}\right|_0 x^2 + \frac{1}{2}\left.\frac{\partial^2 f}{\partial^2 y}\right|_0 y^2 + \left.\frac{\partial^2 f}{\partial x \partial y}\right|_0 xy + \ldots$$

$$= O\left(x^2, y^2, xy\right) \tag{28}$$

due to our chose of positioning and orientation of our surface, so that $0 = \frac{\partial f}{\partial x} = \frac{\partial f}{\partial y}$ at $x = y = 0$.

The implication of this is that the induced metric above for the surface near the chosen point $x = y = 0$ simply becomes,

$$g_{ab} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + O\left(x^2, y^2, xy\right) \tag{29}$$

and thus we see the induced metric appears to be that of $\mathbb{E}^2$ up to quadratic corrections in the displacement away from the chosen point.

Locally the quadratic deviation of the metric from being Euclidean describes the fact that the surface is 'curved'. It indicates that any quantity measuring 'curvature' at a point should involve two derivative of the metric.

This is a rather formal way of saying that if you look at a curved surface very closely (relative to the curvature scale!) then it looks 'flat'. This is a crucial observation in our development of GR later.

1.2.4 Geodesics of a surface

We define a geodesic to be a curve within the surface that has minimal length. In this sense it is the *most direct* or *straightest* path. We may parameterize the curve as $x^a = x^a(\lambda)$ so that $x^i = \{x^1(\lambda), x^2(\lambda), f(x_1(\lambda), x_2(\lambda))\}$.

We see from eqn (18) that the tangent vector at a point on the curve can be neatly written as the 2-vector v^a,

$$v^a = \frac{dx^a(\lambda)}{d\lambda} = \left\{\frac{dx}{d\lambda}, \frac{dy}{d\lambda}\right\} \tag{30}$$

The length of the path is obtained by integrating the norm of the tangent vector which we have seen we may write as,

$$s = \int_{\lambda_0}^{\lambda_1} d\lambda \sqrt{g_{ab} v^a v^b} = \int_{\lambda_0}^{\lambda_1} d\lambda \sqrt{g_{ab}(x) \frac{dx^a(\lambda)}{d\lambda} \frac{dx^b(\lambda)}{d\lambda}} \tag{31}$$

Now an important observation is that under a reparameterization of the path, $\lambda = \lambda(\tau)$, the path length is invariant. To check,

$$\begin{aligned} s &= \int d\lambda \sqrt{g_{ab} \frac{dx^a(\lambda)}{d\lambda} \frac{dx^b(\lambda)}{d\lambda}} \\ &= \int d\tau \frac{d\lambda}{d\tau} \sqrt{g_{ab} \frac{dx^a(\tau)}{d\tau} \frac{d\tau}{d\lambda} \frac{dx^b(\tau)}{d\tau} \frac{d\tau}{d\lambda}} = \int d\tau \sqrt{g_{ab} \frac{dx^a(\tau)}{d\tau} \frac{dx^b(\tau)}{d\tau}} \end{aligned} \tag{32}$$

However, the length element,

$$\begin{aligned} \mathcal{L}_\lambda &\equiv g_{ab}(x) \frac{dx^a(\lambda)}{d\lambda} \frac{dx^b(\lambda)}{d\lambda} \\ &= g_{ab} \frac{dx^a(\tau)}{d\tau} \frac{dx^b(\tau)}{d\tau} \left(\frac{d\tau}{d\lambda} \right)^2 = \mathcal{L}_\tau \left(\frac{d\tau}{d\lambda} \right)^2 \end{aligned} \tag{33}$$

Thus we may always choose a parameterization of the path, by parameter λ, such that $\mathcal{L}_\lambda = 1$. Then the parameter λ measures proper distance along the path.

Now we write the length of a path between two points $x_0 = x(\lambda_0)$ and $x_1 = x(\lambda_1)$ as,

$$s = \int_{\lambda_0}^{\lambda_1} d\lambda \sqrt{\mathcal{L}_\lambda} \tag{34}$$

and we vary this path length, keeping the points x_0 and x_1 fixed, so that $\delta x(\lambda) = 0$ for $\lambda = \lambda_0$ or λ_1. Then we find (in the sense of Euler-Lagrange),

$$\delta s = \int d\lambda \delta \left(\sqrt{\mathcal{L}_\lambda} \right) = \int d\lambda \frac{\delta \mathcal{L}_\lambda}{2\sqrt{\mathcal{L}_\lambda}} \tag{35}$$

and now we make the choice that our original path, before the variation, was in a proper distance parameterization, so,

$$\delta s \;=\; \frac{1}{2}\int d\lambda \delta \mathcal{L}_{\lambda} \tag{36}$$

Now we may evaluate this variation using Euler-Lagrange. Recall for an action,

$$s = \int_{\lambda_0}^{\lambda_1} d\lambda L(x^a, \frac{dx^a}{d\lambda}) \tag{37}$$

then the variation $\delta s = 0$ for fixed end point data $\delta x^a(\lambda_0) = \delta x^a(\lambda_1) = 0$ implies;

$$\frac{d}{d\lambda}\left(\frac{\partial L}{\partial(\frac{dx^a}{d\lambda})}\right) = \frac{\partial L}{\partial x^a} \tag{38}$$

Proof:

$$\begin{aligned} \delta s &= \int_{\lambda_0}^{\lambda_1} d\lambda \delta L(x^i, \frac{dx^i}{d\lambda}) \\ &= \int_{\lambda_0}^{\lambda_1} d\lambda \left(\delta x^i \frac{\partial L}{\partial x^i} + \delta(\frac{dx^i}{d\lambda}) \frac{\partial L}{\partial(\frac{dx^i}{d\lambda})} \right) \end{aligned} \tag{39}$$

then recall $\delta(\frac{dx^i}{d\lambda}) = \frac{d}{d\lambda}(\delta x^i)$, and so,

$$\begin{aligned} \delta s &= \int_{\lambda_0}^{\lambda_1} d\lambda \left(\delta x^i \frac{\partial L}{\partial x^i} + \frac{\partial L}{\partial(\frac{dx^i}{d\lambda})} \frac{d}{d\lambda}(\delta x^i) \right) \\ &= \int_{\lambda_0}^{\lambda_1} d\lambda \left(\delta x^i \frac{\partial L}{\partial x^i} - \delta x^i \frac{d}{d\lambda}\left(\frac{\partial L}{\partial(\frac{dx^i}{d\lambda})} \right) \right) + \left[\delta x^i \frac{\partial L}{\partial(\frac{dx^i}{d\lambda})} \right]_{\lambda_0}^{\lambda_1} \\ &= \int_{\lambda_0}^{\lambda_1} d\lambda \delta x^i \left(\frac{\partial L}{\partial x^i} - \frac{d}{d\lambda}\left(\frac{\partial L}{\partial(\frac{dx^i}{d\lambda})} \right) \right) \end{aligned} \tag{40}$$

Requiring that $\delta s = 0$ for all possible variations of path δx^i then implies the 'E-L' equations. *End of proof.*

Now the equations we require come from varying the Lagrangian,

$$L = \frac{1}{2}\mathcal{L}_\lambda = \frac{1}{2}g_{ab}(x)\frac{dx^a(\lambda)}{d\lambda}\frac{dx^b(\lambda)}{d\lambda} \tag{41}$$

where we recall g_{ab} is a function of x^i (but not $\frac{dx^a(\lambda)}{d\lambda}$). Thus the E-L equations give,

$$\begin{aligned}
\frac{d}{d\lambda}\left(\frac{\partial L}{\partial(\frac{dx^c}{d\lambda})}\right) &= \frac{\partial L}{\partial x^c} \\
\frac{d}{d\lambda}\left(g_{cb}\frac{dx^b}{d\lambda}\right) &= \frac{1}{2}\frac{dx^a}{d\lambda}\frac{dx^b}{d\lambda}\frac{\partial g_{ab}(x)}{\partial x^c} \\
g_{cb}\frac{d^2x^b}{d\lambda^2} + \frac{dx^b}{d\lambda}\frac{d}{d\lambda}g_{cb}(x) &= \frac{1}{2}\frac{dx^a}{d\lambda}\frac{dx^b}{d\lambda}\frac{\partial g_{ab}(x)}{\partial x^c} \\
g_{cb}\frac{d^2x^b}{d\lambda^2} + \frac{dx^b}{d\lambda}\frac{\partial g_{cb}(x)}{\partial x^a}\frac{dx^a}{d\lambda} &= \frac{1}{2}\frac{dx^a}{d\lambda}\frac{dx^b}{d\lambda}\frac{\partial g_{ab}}{\partial x^c}
\end{aligned} \tag{42}$$

Then 'raising' the c index (ie. multiplying by the inverse matrix g^{dc} and using $g^{dc}g_{cb} = \delta^d_b$;

$$\begin{aligned}
0 &= g^{dc}g_{cb}\frac{d^2x^b}{d\lambda^2} + g^{dc}\frac{dx^b}{d\lambda}\frac{\partial g_{cb}(x)}{\partial x^a}\frac{dx^a}{d\lambda} - g^{dc}\frac{1}{2}\frac{dx^a}{d\lambda}\frac{dx^b}{d\lambda}\frac{\partial g_{ab}}{\partial x^c} \\
&= \delta^d_b\frac{d^2x^b}{d\lambda^2} + g^{dc}\frac{dx^a}{d\lambda}\frac{dx^b}{d\lambda}\left(\frac{\partial g_{cb}(x)}{\partial x^a} - \frac{1}{2}\frac{\partial g_{ab}}{\partial x^c}\right) \\
&= \frac{d^2x^d}{d\lambda^2} + g^{dc}\frac{dx^a}{d\lambda}\frac{dx^b}{d\lambda}\left(\frac{\partial g_{cb}(x)}{\partial x^a} - \frac{1}{2}\frac{\partial g_{ab}}{\partial x^c}\right)
\end{aligned} \tag{43}$$

and we finally conclude that a geodesic follows a path that obeys $\delta s = 0$ and hence the *geodesic equation*;

$$\frac{d^2x^d}{d\lambda^2} + \Gamma^d{}_{ab}\frac{dx^a}{d\lambda}\frac{dx^b}{d\lambda} = 0 \tag{44}$$

where we have defined,

$$\Gamma^d{}_{ab}(x(\lambda)) \equiv \frac{1}{2}\,g^{dc}\left(\frac{\partial g_{cb}}{\partial x^a} + \frac{\partial g_{ca}}{\partial x^b} - \frac{\partial g_{ab}}{\partial x^c}\right)\Bigg|_{x^a(\lambda)} \tag{45}$$

and recall that we picked a parameterization $\mathcal{L}(x(\lambda)) = g_{ab}\frac{dx^a}{d\lambda}\frac{dx^b}{d\lambda}\Big|_{x^a(\lambda)} = 1$.

This is a set of coupled second order o.d.e.s in the parameter λ which we may solve for $x^a(\lambda)$. In order to solve it we must therefore give both the initial position x^a and also the velocity $\frac{dx^a}{d\lambda}$ at some starting value of λ. Knowing these one may then integrate the equation to deduce the curve which is given parameterically by the solution $x^a(\lambda)$.

An important comment here is that we must remember the solution is constrained by the condition $g_{ab}\frac{dx^a}{d\lambda}\frac{dx^b}{d\lambda} = 1$. This is a constraint also on our initial data. It is a non-trivial fact (that we will check later) that the solution then preserved this condition, provided it is true for the initial data.

Comment: An important point should be noted here. Whilst we have derived the geodesic equation for a surface embedded in Euclidean space, we have made no explicit reference to this embedding as we have left the metric a general matrix (rather than giving its explicit form in terms of the embedding function f).

Comment: Suppose our surface is a plane and so we embed it as f =constant. Then we saw before that the induced metric is just the 2-d Euclidean metric and then $\Gamma^c{}_{ab} = 0$, and hence,

$$\frac{d^2x^c}{d\lambda^2} = 0 \tag{46}$$

which has general solution,

$$x^c = a^c + b^c\lambda \tag{47}$$

for constant vectors a^c and b^c which give the starting point and direction of the line respectively.

Comment: Suppose that our surface is not a plane. Let us choose to consider a geodesic starting at some point on the curve which we will take (w.l.o.g) to be the point $x = y = 0$. As above arrange the surface so that the embedding has $\partial f/\partial x = \partial f/\partial y = 0$ at this point. Recall that then at this point the metric is just that of 2-d Euclidean space.

Now a *general curve* going through $x = y = 0$ with parameter λ being zero there can be written near this point as,

$$x^c = b^c\lambda + c^c\lambda^2 + O(\lambda^3) \tag{48}$$

by simply Taylor expanding. The constants b^c determine the tangent to the curve. The constants c^c determine the curvature of the curve at the point, with $c^c \sim$ the 'inverse radii of curvatures'.

Now from above expression we see that near $x = y = 0$,

$$\Gamma^c{}_{ab} = O(x, y) \tag{49}$$

and so a *geodesic* passing through this point will have solution,

$$x^c = b^c \lambda + O(\lambda^3) \tag{50}$$

where b^c again gives the direction of the curve through the point.

Hence we see that 'locally' a surface is a plane and a geodesic in that surface is a 'straight line' having no curvature there. It does however deviate from a straight-line as $O(\lambda^3)$. Geodesics generalize the notion of a straight line to a curved geometry, having the property that they are locally straight.

2 'Modern' (Riemannian) Geometry

In classical geometry we have examined a 2-d curved geometry by embedding it into 3-d Euclidean space. However, in all the above we have seen that the notion of vectors, distances, geodesics and local flatness actually reduce to concepts that can be expressed in an 'intrinsically' 2-d manner. The essence of modern differential geometry is to describe geometry from a purely intrinsic perspective without every needing to resort to a higher dimensional embedding space.

Our discussion here will try to avoid giving a formal development of differential geometry. Many of the subtleties are global ones and are not strictly relevant for a first course in GR.

The starting point of geometry is that any geometry locally looks like flat space (as we have seen above for embedded curves). To that end the basic object we employ is the space $\mathbb{R}^d$ (or at least subsets of it).

When we have an embedded surface, the notion of tangent vectors, metrics that describe distance, and geodesics are clear from our understanding of Euclidean geometry. The first task is to understand what these objects are from an intrinsic perspective. A pragmatic way to define tensors (objects that include vectors and the metric) is via their transformation under a coordinate change.

2.1 Coordinate transformations and tensors fields on $\mathbb{R}^d$

The space $\mathbb{R}^d$ is topologically the same as d-dimensional Euclidean space. A natural coordinate system on $\mathbb{R}$ is given by the coordinates $x^i = \{x^1, x^2, \ldots, x^d\}$. There exist (infinitely many) other coordinates; for example the generalization of spherical coordinates.

The reason we are now not considering Euclidean space $\mathbb{E}^d$, but rather $\mathbb{R}^d$, is that we will not consider the geometry of the space to be the same (only the topology – and strictly speaking the differential structure).

We will now consider a change of coordinates $x^i \to x'^{i'}$ so that $x' = x'(x)$ – for example one might have in mind the Cartesian and Spherical coordinates.

We shall restrict ourselves to 'well behaved' coordinate transformations - namely smooth and invertible (so that $x = x(x')$ exists).

Comment: We are not really allowed to transform from cartesians to spherical coordinates as the assumptions above break down on the axis of symmetry.

Under a change of coordinates; $x^i \to x'^{i'} = x'^{i'}(x)$, we define a transformation matrix, $\mathbf{M}$, with components,

$$M^{i'}{}_j \equiv \frac{\partial x'^{i'}(x)}{\partial x^j} \tag{51}$$

This is the Jacobian matrix for the transformations, with the Jacobian given as $J = \det \mathbf{M}$.

As we know, a function transforms trivially under a coordinate transform, namely;

$$f(x) = f(x(x')) \qquad \text{eg. ;} \qquad f = x^2 = (x' - 1)^2 \tag{52}$$

and so we may think of the function as a function of the new coordinates x' instead of x. We say a function is a *scalar* under coordinate transformations. However there are other objects in geometry which do transform in a particular way - the *tensors*.

A **vector field** $v^i(x)$ has the property that it transforms as,

$$v^{i'}(x(x')) = M^{i'}{}_j(x) v^j(x) \tag{53}$$

Thus we see that even sitting at the same actual point the value of the vector changes.

At some point with coordinates x, the **value of a vector field** $v^i(x)$ is a **vector** v^i at that point. This is also called a **tangent vector**. (Sometimes also referred to as a 'contra-variant vector').

The set of all vectors at a point are the **tangent vector space** at the point.

Example: $\{v^r, v^\theta\} = \{0, 1\}$, and a transform $x = r\cos\theta$, $y = r\sin\theta$. Then,

$$\begin{aligned} \begin{pmatrix} v^x \\ v^y \end{pmatrix} &= \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -r\sin\theta \\ r\cos\theta \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix} \end{aligned} \tag{54}$$

Comment: A curve in the geometry is described parametrically as $x^i = x^i(\lambda)$. The tangent vector to a curve is defined as its velocity, $v^i = dx^i(\lambda)/d\lambda$, and transforms correctly as a vector due to the chain rule:

$$v^{i'} = \frac{dx'^{i'}(\lambda)}{d\lambda} = \frac{dx'^{i'}(x)}{\partial x^j}\frac{dx^j(\lambda)}{d\lambda} = M^{i'}{}_j \frac{dx^j(\lambda)}{d\lambda} \tag{55}$$

Fundamentally a vector at a point may be thought of as a tangent to curve through that point.

A **covector field** w_i has its index down and has the property that it transforms 'oppositely' as,

$$w_i(x) = w'_{j'}(x(x'))M^{j'}{}_i(x) \tag{56}$$

or inverting this,

$$w'_{i'} = w_j M^j{}_{i'} \tag{57}$$

where we have defined the components $M_{i'}{}^j$ to be those of the the inverse matrix to $\mathbf{M}$, so that,

$$M_{i'}{}^j = \left(\mathbf{M}^{-1}\right)_{i'}{}^j \tag{58}$$

and then,

$$M^{i'}{}_j M^j{}_{k'} = \delta^{i'}_{k'} \quad \text{and} \quad M^i{}_{j'} M^{j'}{}_k = \delta^i_k \tag{59}$$

where δ^i_j is the Kronecker delta, so that,

$$\delta^i_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{60}$$

(ie. $\mathbf{M} \cdot \mathbf{M}^{-1} = \mathbf{M}^{-1} \cdot \mathbf{M} = \mathbf{1}$)

Due to our assumptions on the allowed coordinate transformations, the inverse matrix $M^j{}_{i'}$ always exists.

At a point x the value of a covector field $w_i(x)$ is a **covector** or **cotangent vector** w_i at the point. (Also sometimes known as a covariant vector).

The set of all covectors at a point form the **cotangent space**.

Due to the property of partial derivatives,

$$M^j{}_{i'} \equiv \frac{\partial x^j(x)}{\partial x'^{i'}} \tag{61}$$

so that we may check;

$$M^{i'}{}_j M^j{}_{k'} = \frac{\partial x'^{i'}(x)}{\partial x^j} \frac{\partial x^j(x)}{\partial x'^{k'}} = \frac{\partial x'^{i'}}{\partial x'^{k'}} = \delta^{i'}_{k'} \tag{62}$$

Comment: given a function $f(x)$ there is a natural covector, with components in the coordinates x^i defined as,

$$w_i \equiv \frac{\partial f(x)}{\partial x^i} \tag{63}$$

Again, due to the chain rule this transforms correctly.

$$w_{i'} = \frac{\partial f(x')}{\partial x'^{i'}} = \frac{\partial f(x)}{\partial x^j} \frac{\partial x^j(x')}{\partial x'^{i'}} = \frac{\partial f(x)}{\partial x^j} M^j{}_{i'} \tag{64}$$

Fundamentally covector at a point may be thought of as a differentials of function at that point.

Giving a vector or covector in one coordinate system this then defines it in any coordinate system.

A (q,r) **tensor field** has q 'up' indices and r 'down' indices, and in general is defined by giving its components $T^{i_1 i_2 \dots i_q}_{j_1 j_2 \dots j_r}$ in some coordinate system x^i. Under a transform to new coordinates $x^{i'}$ then the new components are,

$$T'^{i'_1 i'_2 \dots i'_q}_{j'_1 j'_2 \dots j'_r}(x(x')) = M^{i'_1}{}_{i_1} \dots M^{i'_q}{}_{i_q} \left(T^{i_1 i_2 \dots i_q}_{j_1 j_2 \dots j_r}(x) \right) M^{j_1}{}_{j'_1} \dots M^{j_r}{}_{j'_r} \tag{65}$$

At a point, the value of a tensor field gives a (q,r) **tensor** at the point.

The set of all (q,r) tensors at a point is the (q,r) **tensor space** at the point.

Comment: The metric g_{ij} is a $(0,2)$ tensor. A vector is a $(1,0)$ tensor.

Comment: We may view a function as a $(0,0)$ tensor. We call this a scalar quantity as It transforms trivially; $f(x) \to f(x')$.

Comment: Note that the Kronecker delta δ^i_j is a $(1,1)$ tensor. It has the special property that it is invariant under a coordinate transform;

$$M^{i'}{}_i \delta^i_j M^j{}_{j'} = M^{i'}{}_i M^i{}_{j'} = \delta^{i'}_{j'} \tag{66}$$

We say it is an invariant tensor.

2.2 Tensor operations

We may linearly combine two (q,r) tensors, **A** and **B**, to give a third (q,r) tensor **C**, defined in components in a coordinate system x^i as,

$$C^{i_1 i_2 \dots i_q}_{j_1 j_2 \dots j_r} \equiv f(x) A^{i_1 i_2 \dots i_q}_{j_1 j_2 \dots j_r} + g(x) B^{i_1 i_2 \dots i_q}_{j_1 j_2 \dots j_r} \tag{67}$$

for functions f, g. Note that the transformation of the components of **A** and **B** imply that the components of **C** correctly transform as a tensor.

Thus we see the tensor space at a point is a vector space with dimension d^{q+r}.

We may take an outer product of a (q_1, r_1) tensor **A** and a (q_2, r_2) tensor **B** to give a $(q_1 + q_2, r_1 + r_2)$ tensor **C** defined by its components as,

$$C^{i_1 \dots i_{q_1} j_1 \dots j_{q_2}}_{a_1 \dots a_{r_1} b_1 \dots b_{r_2}} \equiv A^{i_1 \dots i_{q_1}}_{a_1 \dots a_{r_1}} \cdot B^{j_1 \dots j_{q_2}}_{b_1 \dots b_{r_2}} \tag{68}$$

Again we note that the components of $\mathbf{C}$ transform correctly as a tensor.

Finally we may contract the a^{th} 'up' and b^{th} 'down' index of a (q,r) tensor $\mathbf{A}$ to give a $(q-1,r-1)$ tensor $\mathbf{B}$, defined as,

$$B^{i_1\ldots i_{a-1}i_{a+1}\ldots i_q}_{j_1\ldots j_{b-1}j_{b+1}\ldots j_r} \equiv A^{i_1\ldots i_{a-1}k\, i_{a+1}\ldots i_q}_{j_1\ldots j_{b-1}k\, j_{b+1}\ldots j_r} \tag{69}$$

Just as we saw Kronecker delta δ^i_j is a tensor, a similar calculation confirms this contraction yields a well defined tensor $\mathbf{B}$.

Comment: Suppose we have vector fields v^i and w^i. We may view their dot product $g_{ij}v^iw^j$ as a suitable outer product of the metric g, and these vectors, and then the contraction over all the indices. We are reduced to a $(0,0)$ tensor ie. a function. Thus this dot product is a scalar quantity.

Given a tensor we may symmetrize or antisymmetrise its upper or lower indices. We use the notation;

$$\begin{aligned} T_{(ij)} &= \frac{1}{2}\left(T_{ij} + T_{ji}\right) \\ T_{[ij]} &= \frac{1}{2}\left(T_{ij} - T_{ji}\right) \end{aligned} \tag{70}$$

With more indices we have,

$$T_{(i_1 i_2 \ldots i_n)} = \frac{1}{n!}\left(T_{i_1 i_2 \ldots i_n} + \text{sym}\right) \tag{71}$$

and likewise for antisymmetry.

Comment: Since these are linear operations, the resulting symmetrized or antisymmetrized objects are also tensors.

Important comment: We may *not* add tensors at different points. For example given vector fields $A^i(x)$ and $B^i(x)$ it is straightforward to check that the object,

$$C^i(x) = A^i(x) + B^i(x+v) \tag{72}$$

for some constant v^i does *not* transform as a vector.

Is this surprising? We are used to adding vectors in Euclidean space that live at different points. However, in our curved surface example above it would be less ‘natural’ to compare vectors at different points on the surface. In order to compare vectors at different points we need a notion of parallel transport which allows us to take a vector at one position and move it somewhere else preserving its ‘direction’. For that we need a notion of geometry.

2.3 The metric tensor

The metric is defined to be a $(0, 2)$ tensor g_{ij} that is *symmetric* so that $g_{ij} = g_{ji}$. At any point this is a symmetric matrix and hence has real eigenvalues. The metric is required to have all its eigenvalues being strictly positive at all points. A corollary of this is that $\det g_{ij} > 0$ everywhere, and hence g_{ij} is invertable.

The inverse of the metric, g^{ij}, is a symmetric $(2, 0)$ tensor, defined as the inverse of the metric at every point. This implies that *in any coordinate system* we must have,

$$g^{ij} g_{jk} = \delta^i_k \tag{73}$$

Recall this makes sense as the $(1, 1)$ tensor δ^i_j is invariant - its components are indeed the same in any coordinate system.

The metric provides a notion of the norm of a vector v^i as $|v|^2 = g_{ij} v^i v^j$.

We measure the distance s along a curve by integrating the norm of the tangent vector as,

$$s = \int d\lambda \sqrt{g_{ij} v^i v^j} \tag{74}$$

where the tangent vector $v^i = dx^i(\lambda)/d\lambda$.

The metric also provides a notion of angle, θ, between two vectors v, w at a point,

$$|v||w| \cos\theta = v^i w^j g_{ij} \tag{75}$$

Comment: The distance and angle measures above are independent of the choice of coordinates. It is a coordinate invariant quantity. All measurable quantities in geometry must be coordinate invariant.

Comment: (At least In GR) The metric tensor completely specifies the geometry. Two metrics that differ by a coordinate transformation describe the *same* geometry.

Comment: In order to specify the geometry of our 2-surface embedded in $\mathbb{E}^3$ that we started with we see we should take $\mathbb{R}^2$ and endow it with the metric tensor,

$$g_{ij} = \begin{pmatrix} 1+\left(\frac{\partial f}{\partial x}\right)^2 & \frac{\partial f}{\partial x}\frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial x}\frac{\partial f}{\partial y} & 1+\left(\frac{\partial f}{\partial y}\right)^2 \end{pmatrix} \tag{76}$$

or any coordinate transformation of this. However in general you *cannot* take an n-dimensional geometry and embed it into $\mathbb{E}^{n+1}$.

2.4 Tensors and the metric

The metric gives an important new structure for tensors, namely the ability to 'raise' and 'lower' indices. If we have a (q, r) tensor $T^{i_1 \ldots i_q}_{j_1 \ldots j_r}$, then we may define a $(q-1, r+1)$ tensor, A, by *lowering* the n^{th} index with the metric as,

$$A^{i_1 \ldots i_{n-1} i_{n+1} \ldots i_{q-1}}_{a j_1 \ldots j_r} = g_{a\, i_n} T^{i_1 \ldots i_q}_{j_1 \ldots j_r} \tag{77}$$

Likewise we can define a new $(q+1, r-1)$ tensor B, by *raising* the m^{th} index with the metric as,

$$B^{a i_1 \ldots i_{q-1}}_{j_1 \ldots j_{m-1} j_{m+1} \ldots j_r} = g^{a\, j_m} T^{i_1 \ldots i_q}_{j_1 \ldots j_r} \tag{78}$$

New notation: We introduce the notation that we order both the upper and lower indices together. Then we can simply raise or lower an index. Writing the tensor above as $T^{i_1 \ldots i_q}{}_{j_1 \ldots j_r}$ then we may understand,

$$A^{i_1 \ldots \hat{i}_n \ldots i_{q-1}}_{a j_1 \ldots j_r} = T^{i_1 \ldots i_{n-1}}{}_{a}{}^{i_{n+1} \ldots i_q}{}_{j_1 \ldots j_r} \tag{79}$$

and

$$B^{ai_1\cdots i_{q-1}}_{j_1\cdots j_{m-1}j_{m+1}\cdots j_r} = T^{i_1\cdots i_q}{}_{j_1\cdots j_{m-1}}{}^{a}{}_{j_{m+1}\cdots j_r} \tag{80}$$

2.5 Geodesics and the Christoffel symbol

Our discussion of geodesic curves follows exactly the same lines as that above. Such a curve between two points extremizes (minimises) its length. The geodesic equation is derived in the same manner as before – we note that derivation was essentially intrinsic. Using the **affine parameterization**,

$$g_{ab}\frac{dx^a}{d\lambda}\frac{dx^b}{d\lambda} = 1 \tag{81}$$

one finds the **geodesic equation**,

$$\frac{d^2x^c}{d\lambda^2} + \Gamma^c{}_{ab}\frac{dx^a}{d\lambda}\frac{dx^b}{d\lambda} \quad = \quad 0 \tag{82}$$

where as before we have defined,

$$\Gamma^c{}_{ab}(x) \equiv \frac{1}{2}g^{cd}\left(\frac{\partial g_{db}}{\partial x^a} + \frac{\partial g_{da}}{\partial x^b} - \frac{\partial g_{ab}}{\partial x^d}\right) \tag{83}$$

Note that a solution of this geodesic equation is always compatible with having $g_{ab}\frac{dx^a}{d\lambda}\frac{dx^b}{d\lambda} = 1$.

The object $\Gamma^c{}_{ab}$ is called the *Christoffel symbol* and will play an important role later. However, for now we simply note that it is symmetric in its lower indices but is *not* a tensor as it does not transform correctly. Under a transformation we find;

$$\begin{aligned}
\Gamma^{c'}{}_{a'b'}(x'(x)) &= M^{c'}{}_c\Gamma^c{}_{ab}M^a{}_{a'}M^b{}_{b'} - \frac{\partial^2 x^{c'}(x)}{\partial x^a \partial x^b}M^a{}_{a'}M^b{}_{b'} \\
&= M^{c'}{}_c\Gamma^c{}_{ab}M^a{}_{a'}M^b{}_{b'} + M^{c'}{}_c\frac{\partial^2 x^c(x')}{\partial x'^{a'}\partial x'^{b'}}
\end{aligned} \tag{84}$$

where we explicitly see the trailing term violates the tensor transformation law. (See example sheet 2)

In fact rather than being a tensor, it is a *connection*, which is something that transforms as the above. The Christoffel symbol is known as the *Levi-Civita connection.*

The Christoffel symbol has a symmetry in its lower 2 indices,

$$\Gamma^c{}_{ab} = \Gamma^c{}_{ba} \tag{85}$$

One can also check that (see exercise sheet 3),

$$\frac{\partial g_{ab}}{\partial x^c} = g_{ma}\Gamma^m{}_{bc} + g_{mb}\Gamma^m{}_{ac} \tag{86}$$

2.6 Riemann Normal coordinates

Consider a point at the origin $x^i = 0$. (Note that by taking a coordinate transform of the form $x'^{i'} = x^i + a^i$ for some a^i we can always map any particular point to the origin of coordinates).

Consider a coordinate transformation that locally is written near the origin $x^i = 0$ as a Taylor expansion,

$$x^{i'} = b^{i'}{}_j x^j + \frac{1}{2} b^{i'}{}_j c^j{}_{mn} x^m x^n + \dots \tag{87}$$

where the $b^{i'}{}_j$ and $c^{i'}{}_{jk}$ are constant matrices (in the x coordinates). We take the matrix $b^{i'}{}_j$ to be invertible with inverse $b^i{}_{j'}$ so that,

$$b^{i'}{}_j b^j{}_{k'} = \delta^{i'}_{k'} \ , \quad b^i{}_{j'} b^{j'}{}_k = \delta^i_k \tag{88}$$

Now it is required to be invertible so that the inverse coordinate transform exists. We take c to be symmetric in its lower indices, $c^i{}_{jk} = c^i{}_{kj} = c^i{}_{(jk)}$ as only this symmetric part contributes above.

Then the matrix $M^{i'}{}_i$ takes the form,

$$M^{i'}{}_i \equiv \frac{\partial x'^{i'}(x)}{\partial x^j} = b^{i'}{}_i + b^{i'}{}_m c^m{}_{ij} x^j + O(x^2) \tag{89}$$

The inverse matrix $M^i{}_{i'}$ takes the form,

$$M^i{}_{i'} = b^i{}_{i'} - c^i{}_{mj} b^m{}_{i'} x^j + O(x^2) \tag{90}$$

to ensure that $M^i{}_{i'} M^{i'}{}_j = \delta^i_j + O(x^2)$, $M^{i'}{}_i M^i{}_{j'} = \delta^{i'}_{j'} + O(x^2)$.

Consider Taylor expanding the metric at the origin as,

$$g_{ij}(x) = g_{ij}|_0 + \frac{\partial}{\partial x^k} g_{ij}\bigg|_0 x^k + \frac{1}{2}\frac{\partial^2}{\partial x^k \partial x^l} g_{ij}\bigg|_0 x^k x^l + \dots \tag{91}$$

Then the metric transforms in the new coordinates to,

$$g'_{i'j'}(x'(x)) = b^i{}_{i'} b^j{}_{j'}\, g_{ij}|_0 + b^i{}_{i'} b^j{}_{j'} \left(\frac{\partial}{\partial x^m} g_{ij} - g_{in} c^n{}_{mj} - g_{nj} c^n{}_{mi} \right)\bigg|_0 x^m + O(x^2)$$

Now there always exists an appropriate choice of matrix $b^i{}_{i'}$ such that,

$$b^i{}_{i'} b^j{}_{j'}\, g_{ij}|_0 = \delta_{i'j'} \tag{92}$$

Furthermore we have previous determined that,

$$\frac{\partial g_{ab}}{\partial x^c} = g_{ma} \Gamma^m{}_{bc} + g_{mb} \Gamma^m{}_{ac} \tag{93}$$

and we see we can set the $O(x)$ term to zero by taking;

$$c^m{}_{ab} = \Gamma^m{}_{ab}|_0 \tag{94}$$

(which is consistent with the symmetry properties of c) so then,

$$g'_{i'j'}(x'(x)) = \delta_{i'j'} + O(x^2) \tag{95}$$

As discussed in example sheet 3, this choice of $c^m{}_{ab}$ is unique.

These coordinates are called *Riemann normal coordinates.* As earlier for the surface we see that for a geometry we may pick any point and use a coordinate transform to render the metric the Euclidean metric up to quadratic order. The quadratic deviations indicate the geometry is in fact curved, not the Euclidean one. Thus we learn the important fact that all geometries are 'locally' Euclidean. This will be crucial later.

[**Aside:**

Now there always exists an appropriate choice of matrix $b^i{}_{i'}$ such that,

$$b^i{}_{i'} b^j{}_{j'} \, g_{ij}\big|_0 = \delta_{i'j'} \tag{96}$$

The matrix $b^i{}_{i'}$ is equal to,

$$\mathbf{b} = \mathbf{O} \cdot \mathbf{D} \cdot \mathbf{O}' \tag{97}$$

using matrix notation, where $\mathbf{O}$ is a rotation that diagonalizes the symmetric matrix $\mathbf{g}$ at $x = 0$, so that,

$$\mathbf{O}^T \cdot (\mathbf{g}|_0) \cdot \mathbf{O} = \mathbf{\Lambda} \; , \quad \mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ & & & \\ \vdots & & \lambda_{d-1} & 0 \\ 0 & \dots & 0 & \lambda_d \end{pmatrix} \tag{98}$$

and the matrix $\mathbf{D}$ is diagonal with,

$$\mathbf{D} = \mathbf{\Lambda}^{-1/2} = \begin{pmatrix} \lambda_1^{-1/2} & 0 & \dots & 0 \\ 0 & \lambda_2^{-1/2} & & \vdots \\ & & & \\ \vdots & & \lambda_{d-1}^{-1/2} & 0 \\ 0 & \dots & 0 & \lambda_d^{-1/2} \end{pmatrix} \tag{99}$$

and $\mathbf{O}'$ is then an arbitrary rotation - ie. the matrices that leave δ_{ij} invariant - which represents the freedom in choosing the $b^i{}_{i'}$.

Note that the λ_i are the eigenvalues of the metric matrix $\mathbf{g}$ and hence must all be strictly positive, so there is not concern over taking the negative fractional power. (See example sheet 3 for more details). **End of aside.**]

Comment: Consider taking the above Riemann normal coordinates so that the metric mapped from the origin point is locally Euclidean. Then the Christoffel connection also vanishes at that point and behaves linearly about it,

$$\Gamma^c{}_{ab} = O(x) \tag{100}$$

Hence a general geodesic passing through this point will have solution,

$$x^c = b^c \lambda + O(\lambda)^3 \tag{101}$$

and thus appears locally as a straight line - ie. no quadratic terms about $x = 0$.

2.7 Lie derivatives and symmetries

Given a vector field $v^i(x)$ we can define a flow, which is a set of curves parameterized by a parameter λ and their starting point $x_{(0)}$, which we write $x^i_{(v)}(\lambda, x_{(0)})$, with the properties that,

$$v^i(x_{(v)}(\lambda, x_{(0)})) = \frac{dx^i_{(v)}(\lambda, x_{(0)})}{d\lambda} \tag{102}$$

for any starting point $x_{(0)}$, and also,

$$x^i_{(0)} = x^i_{(v)}(0, x_{(0)}) \tag{103}$$

Note that a flow obeys;

$$x^i_{(1)} = x^i_{(v)}(\lambda_1, x_{(0)}) \ , \quad x^i_{(v)}(\lambda_2, x_{(1)}) = x^i_{(v)}(\lambda_1 + \lambda_2, x_{(0)}) \tag{104}$$

The flow is a (dense) set of curves in space. For every point there is a unique curve in the flow going through it. We say that the vector field v^i **generates the flow** $x^i_{(v)}(\lambda, x_0)$.

Suppose we have a function $f(x)$. How does it vary as we move along a flow line? The answer is given by the **Lie Derivative**, defined on a function as,

$$Lie(v, f) \equiv v^i \frac{\partial f}{\partial x^i} \tag{105}$$

Note that using the fact, $v^i = dx^i_{(v)}/d\lambda$, then,

$$Lie(v, f) = \frac{\partial f}{\partial x^i} \frac{dx^i_{(v)}}{d\lambda} = \frac{df(x_{(v)}(\lambda))}{d\lambda} \tag{106}$$

so the Lie derivative at a point p gives the rate of change of the function f at p along the flow curve going through p.

Suppose we have a function $f(x)$ and a flow generated by $v^i(x)$ such that $Lie(v, f) = df/d\lambda = 0$. We say that the flow is a **symmetry** of f, and that $v^i(x)$ generates this symmetry.

Example: Consider $\mathbb{E}^3$ in spherical polar coordinates (r, θ, ϕ). Then the flow generated by $v^i = (0, 0, 1)$ is,

$$x^i_{(v)} = (r_0, \theta_0, \phi_0 + \lambda) \tag{107}$$

so it is a flow in the aximuthal direction. Suppose we have a function $f = f(r, \theta)$ - so it is axisymmetric. Then,

$$Lie(v, f) = \frac{\partial f}{\partial x^i} \frac{dx^i_{(v)}}{d\lambda} = \frac{\partial f}{\partial \phi} = 0 \tag{108}$$

Thus $v^i = (0, 0, 1)$ generates the symmetry of this axisymmetric function.

Now suppose we want to do the same with a vector field $w^i(x)$, namely to compute its rate of change along the flow generated by $v^i(x)$. We can use coordinates **adapted to the flow**, using the flow curves to define our coordinates. Then we may (at least in some region) choose $v^i(x) = (1, 0, \ldots, 0)$. In these coordinates we want,

$$Lie(v, w)^i = \frac{\partial}{\partial x^1} w^i \tag{109}$$

so the rate of change of w^i in the direction of the flow (ie. in the x^1 direction).

Let's try to write down the tensor expression, valid in any coordinate system. We might try,

$$Lie(v, w)^i = v^j \frac{\partial}{\partial x^j} w^i \ ? \tag{110}$$

Whilst this obviously agrees for the particular coordinate system above, it does **not** transform correctly as a $(1, 0)$ tensor. So this is wrong.

In fact the correct tensor expression is,

$$Lie(v,w)^i \equiv v^j \frac{\partial}{\partial x^j} w^i - w^j \frac{\partial}{\partial x^j} v^i \tag{111}$$

This fixes the problem with transformation (Ex to check!), and agrees with our expectations in the special adapted coordinate system where the second term vanishes.

Comments: It is linear in each argument and it obeys a 'product' rule, for a function f,

$$Lie(v, fw)^i = f Lie(v,w)^i + w^i Lie(v,f) \tag{112}$$

BUT it isn't quite like a derivative we are used to. Unlike the Lie derivative of a function which at a point depends only on the direction of the flow v^i there, now for a vector field the Lie derivative depends on the local properties of the flow, so both v^i **and** $\partial v^i / \partial dx^j$ at the point.

In order to get a more usual derivative, we will need to introduce more structure as we shall see later when we discuss the covariant derivative.

We may define a Lie derivative of a covector field $\omega_i(x)$ along a flow $v^i(x)$. One finds that,

$$Lie(v,\omega)_i \equiv v^j \frac{\partial}{\partial x^j} \omega_i + \omega_j \frac{\partial}{\partial x^i} v^j \tag{113}$$

For a general (q,r) tensor the expression is;

$$\begin{aligned} \mathrm{Lie}(v,T)^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q} \equiv\ & v^k \partial_k T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q} \\ & - T^{k i_2 \ldots i_p}_{j_1 j_2 \ldots j_q} \partial_k v^{i_1} - T^{i_1 k \ldots i_p}_{j_1 j_2 \ldots j_q} \partial_k v^{i_2} - \ldots - T^{i_1 i_2 \ldots k}_{j_1 j_2 \ldots j_q} \partial_k v^{i_p} \\ & + T^{i_1 i_2 \ldots i_p}_{k j_2 j_q} \partial_{j_1} v^k + T^{i_1 i_2 \ldots i_p}_{j_1 k \ldots j_q} \partial_{j_2} v^k + \ldots + T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots k} \partial_{j_q} v^k \end{aligned}$$

Note the new notation: $\partial_i = \partial / \partial x^i$.

These have the property we want, namely they are tensor expressions, and in a coordinate system adapted to the flow, so that $v^i = (1, 0, \ldots, 0)$ then,

$$\text{Lie}(v,T)^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q} = \frac{\partial}{\partial x^1} T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q} \tag{114}$$

We say that v^i generates a **symmetry of a tensor** $T^{i_1 i_2 \ldots i_p}_{j_1 j_2 \ldots j_q}$ if $Lie(v,T) = 0$.

In particular the Lie derivative of the metric g_{ij} is;

$$\text{Lie}(v,g)_{ij} = v^k \partial_k g_{ij} + g_{ik} \partial_j v^k + g_{jk} \partial_i v^k \tag{115}$$

A *symmetry* of the metric is called an **isometry**. If for some vector v^i we have $\text{Lie}(v,g) = 0$ we say v^i generates an isometry of the metric. In fact then v^i is called a 'Killing' vector.

Ex. For a 2-sphere we have the metric $ds^2 = d\theta^2 + \sin^2\theta d\phi^2$, and $v^i = (v^r, v^\theta) = (0, 1)$ is an isometry of the sphere. There are others - see the example sheet!

It means that if you 'flow along in the ϕ direction' then the geometry remains invariant.

2.8 Christoffel symbol and covariant derivatives

For convenience we now use the standard notation: $\partial_a \equiv \frac{\partial}{\partial x^a}$

Consider a function $\phi(x)$ on our geometry. Suppose we are interested in computing its gradient. The natural quantity to compute is, $\partial_a \phi$. We have earlier mentioned that this quantity indeed transforms as a $(0,1)$ tensor (see Ex Sheet 2).

However, suppose we try to do the same thing for a vector. We might try to compute its gradient as, ${dv_a}^b \equiv \partial_a v^b$. Let us see how this transforms,

$$\begin{aligned}
{dv_{a'}}^{b'} &= \frac{\partial v^{b'}}{\partial x^{a'}} = \frac{\partial x^a}{\partial x^{a'}}\frac{\partial}{\partial x^a}\left(v^b \frac{\partial x^{b'}}{\partial x^b}\right) \\
&= \frac{\partial x^a}{\partial x^{a'}}\frac{\partial x^{b'}}{\partial x^b}\frac{\partial v^b}{\partial x^a} + v^c \frac{\partial x^a}{\partial x^{a'}}\frac{\partial x^{b'}}{\partial x^a \partial x^c} \\
&= {M^a}_{a'} {dv_a}^b {M^{b'}}_b + v'^{c'} {M^a}_{a'} {M^c}_{c'} \frac{\partial x^{b'}}{\partial x^a \partial x^c} \qquad (116)
\end{aligned}$$

Now the first term is precisely the transformation of a $(1,1)$ tensor, *but* the second term doesn't behave nicely at all.

So our usual notion of derivative, simply using ∂_μ, doesn't work as it doesn't give tensor $(q, r+1)$ quantities when you act on general (q, r) tensors. In order to proceed recall,

$$\Gamma^{b'}{}_{a'c'} = M^{b'}{}_b \Gamma^b{}_{ac} M^a{}_{a'} M^c{}_{c'} - M^a{}_{a'} M^c{}_{c'} \frac{\partial^2 x^{b'}(x)}{\partial x^a \partial x^c} \qquad (117)$$

so that,

$$v'^{c'} \Gamma^{b'}{}_{a'c'} = M^{b'}{}_b (v^c \Gamma^b{}_{ac}) M^a{}_{a'} M^c{}_{c'} - v'^{c'} M^a{}_{a'} M^c{}_{c'} \frac{\partial^2 x^{b'}(x)}{\partial x^a \partial x^c} \qquad (118)$$

Hence we may define a new derivative, the *covariant* derivative ∇_a as,

$$\nabla_a v^b \equiv \partial_a v^b + \Gamma^b{}_{ac} v^c \qquad (119)$$

and this *does* transform correctly as a tensor. The naughty second terms in the transformation of ${dv_{a'}}^{b'}$ above and $\Gamma^{c'}{}_{a'b'}$ cancel. The key point is that neither $\partial_a v^b$ nor this connection transform properly as tensors. However their combination above exactly does (see example sheet 3).

What is this object ∇_i? It is the natural generalisation of the partial derivative ∂_i we use in Euclidean space to a curved geometry. This is best seen by recalling that at any point p (take it to be at $x = 0$) we may choose Riemann Normal coordinates so that $g_{ij} = \delta_{ij} + O(x^2)$ and so $\Gamma^k{}_{ij} = O(x)$. Then locally at this point,

$$\nabla_a v^b|_0 = \partial_a v^b|_0 + \Gamma^b{}_{ac} v^c|_0 = \partial_a v^b|_0 \qquad (120)$$

so locally the covariant derivative *is* just the partial derivative.

A very useful fact is that in Riemann Normal coordinates if you see a ∂_i then to get a general tensor expression in general coordinates you just have to replace $\partial_i \to \nabla_i$.

This construction generalizes for a (q,r) tensor $T^{i_1 \dots i_q}{}_{j_1 \dots j_r}$;

$$\begin{aligned} \nabla_a T^{i_1 \dots i_q}{}_{j_1 \dots j_r} \equiv & \; \partial_a T^{i_1 \dots i_q}{}_{j_1 \dots j_r} \\ & + \Gamma^{i_1}{}_{ab} T^{b i_2 \dots i_q}{}_{j_1 \dots j_r} + \Gamma^{i_2}{}_{ab} T^{i_1 b i_3 \dots i_q}{}_{j_1 \dots j_r} + \dots \\ & - \Gamma^{b}{}_{a j_1} T^{i_1 \dots i_q}{}_{b j_2 \dots j_r} - \Gamma^{b}{}_{a j_2} T^{i_1 \dots i_q}{}_{j_1 b j_3 \dots j_r} + \dots \end{aligned} \quad (121)$$

For convenience we define the covariant derivative on a function (a $(0,0)$ tensor) simply to be the partial derivative;

$$\nabla_i f \equiv \partial_i f \quad (122)$$

In particular we find for a covector,

$$\nabla_a w_b \equiv \partial_a w_b - \Gamma^{c}{}_{ab} w_c \quad (123)$$

The covariant derivative has the usual 'derivative' properties. It obeys a product (Liebnitz) rule,

$$\nabla_a \left(A^{i_1 \dots}{}_{j_1 \dots} B^{m_1 \dots}{}_{n_1 \dots} \right) = A^{i_1 \dots}{}_{j_1 \dots} \nabla_a B^{m_1 \dots}{}_{n_1 \dots} + A^{i_1 \dots}{}_{j_1 \dots} \nabla_a B^{m_1 \dots}{}_{n_1 \dots} \quad (124)$$

You can prove this either by direct calculation or by observing that in Riemann Normal coordinates at a point $x = 0$, then,

$$\nabla_a \left(A^{i_1 \dots}{}_{j_1 \dots} B^{m_1 \dots}{}_{n_1 \dots} \right) |_0 = \partial_0 \left(A^{i_1 \dots}{}_{j_1 \dots} B^{m_1 \dots}{}_{n_1 \dots} \right) |_0 = A^{i_1 \dots}{}_{j_1 \dots} \partial_a B^{m_1 \dots}{}_{n_1 \dots} |_0 + A^{i_1 \dots}{}_{j_1 \dots} \partial_a B^{m_1 \dots}{}_{n_1 \dots} |_0$$

using the product/Leibnitz rule of partial derivatives. Now using the *very useful fact* above we know that going back to a general coordinate system we simply observe the answer is equation (124) when we take $\partial_i \to \nabla_i$.

Properties of ∇_a:

Most importantly it vanishes on the metric;

$$\nabla_a g_{bc} = \partial_a g_{bc} - \Gamma^{d}{}_{ab} g_{dc} - \Gamma^{d}{}_{ac} g_{bd} = 0 \quad (125)$$

by the property of the Christoffel symbol.

We say it is *metric compatible*.

Note: this is simple to see in RN coordinates (say about $x = 0$); then,

$$\nabla_a g_{bc}|_0 = \partial_a g_{bc}|_0 = \partial_a \delta_{bc}|_0 = 0 \tag{126}$$

It also vanishes on the invariant tensor δ^i_j;

$$\begin{aligned} \nabla_a \delta^b_c &= \partial_a \delta^b_c + \Gamma^b{}_{ad} \delta^d_c - \Gamma^d{}_{ac} \delta^b_d \\ &= \Gamma^b{}_{ad} \delta^d_c - \Gamma^d{}_{ac} \delta^b_d = \Gamma^b{}_{ac} - \Gamma^b{}_{ac} = 0 \end{aligned} \tag{127}$$

by the property of the Christoffel symbol.

Again this is straightforward to derive in RN coordinates.

Likewise $\nabla_a g^{bc} = 0$ as we shall show. Recall $g^{bc} g_{cd} = \delta^b_d$ so,

$$0 = \nabla_a \left(g^{bc} g_{cd} \right) = g_{cd} \nabla_a g^{bc} + g^{bc} \nabla_a g_{cd} = g_{cd} \nabla_a g^{bc} \tag{128}$$

and hence $\nabla_a g^{bc} = 0$.

Hence we may write $\nabla^i = g^{ij} \nabla_i = \nabla_i g^{ij}$.

Comment: This covariant derivative is the *unique* derivative that is metric compatible and built solely from the metric (without any other data such as torsion).

Laplacian: Consider the Laplacian on a function;

$$\nabla^2 \phi = g^{ij} \nabla_i \nabla_j \phi = \nabla^i \nabla_i \phi = \nabla_i \nabla^i \phi \tag{129}$$

We may write this out as;

$$\begin{aligned} \nabla^2 \phi &= g^{ij} \nabla_i \nabla_j \phi = g^{ij} \nabla_i \left(\partial_j \phi \right) \\ &= g^{ij} \left(\partial_i \partial_j - \Gamma^k{}_{ij} \partial_k \right) \phi \end{aligned} \tag{130}$$

In fact one may also write this as,

$$\nabla^2 \phi = \frac{1}{\sqrt{g}} \partial_i \left((\sqrt{g}) g^{ij} \partial_j \phi \right) \tag{131}$$

where $g = \det g_{ij}$. (See example sheet 4).

2.9 Parallel transport

The covariant derivative allows us to determine the notion of vectors being parallel at *different* points.

Suppose we have a curve that joins two points $x(\lambda_0)$ and $x(\lambda_1)$, and the tangent to that curve is $v^i = dx^i/d\lambda$. Then we say a vector field w^i is *parallel* transported along that curve if;

$$v^i \nabla_i w^j \big|_{x(\lambda)} = 0 \tag{132}$$

We say then that $w^i(\lambda_0)$ is *parallel* to $w^i(\lambda_1)$ with respect to the curve.

Note however this notion of parallel depends on the curve joining the two points. In general two curves going from $x(\lambda_0)$ to $x(\lambda_1)$ will not agree on what is parallel.

It is a special property of Euclidean space that the notion of parallel is independent of the curve. In fact as we see later, it is precisely the presence of *curvature* that means different curves give rise to different parallel transport for different curves.

We may also think of the equation above as a way to evolve a vector from one point, $x(\lambda_0)$, to another at $x(\lambda_1)$.

$$\begin{aligned} 0 = v^i \nabla_i w^j &= \frac{dx^i}{d\lambda}\frac{\partial w^j}{\partial x^i} - \Gamma^j_{ik} v^i w^k \\ &= \frac{dw^j(\lambda)}{d\lambda} - \Gamma^j_{ik} v^i w^k \big|_{x=x(\lambda)} \end{aligned} \tag{133}$$

which we view as a first order o.d.e. in λ which is integrated along the curve in λ, using an initial condition, $w^i(\lambda_0)$, to obtain the value of $w^i(\lambda_1)$.

A more geometric way to understand a geodesic is a curve whose tangent vector is parallel along the curve - ie. as ‘straight as possible. For a geodesic,

$$v^i \nabla_i v^j \big|_{x(\lambda)} = 0 \tag{134}$$

(see Ex Sheet 4). This is simply shown as;

$$
\begin{aligned}
0 = v^i \nabla_i v^j &= v^i \partial_i v^j + v^i \Gamma^j{}_{ik} v^k = \frac{dx^i}{d\lambda}\frac{\partial v^j}{\partial x^i} + \Gamma^j{}_{ik}\frac{dx^i}{d\lambda}\frac{dx^k}{d\lambda} \\
&= \frac{dv^j}{d\lambda} + \Gamma^j{}_{ik}\frac{dx^i}{d\lambda}\frac{dx^k}{d\lambda} \\
&= \frac{d^2 x^j}{d\lambda^2} + \Gamma^j{}_{ik}\frac{dx^i}{d\lambda}\frac{dx^k}{d\lambda} \qquad (135)
\end{aligned}
$$

3 Special Relativity

3.1 Brief review of Special Relativity

The basis of Special Relativity (SR) is the **Lorentz** transformation. Consider two inertial observers with frames F, F' with spacetime coordinates t, x, y, z and t', x', y', z' respectively. Suppose frame F' is moving at a constant velocity v in the x-axis direction of frame F so that the origin spacetime point of F and F' coincide ie. the point $t' = x' = y' = z' = 0$ is the same as $t = x = y = z = 0$. Then SR states that the spacetime coordinates of the two frames are related by a **Lorentz** transformation,

$$t' = \gamma \left(t - \frac{v}{c^2} x \right) \ , \quad x' = \gamma \left(x - vt \right) \ , \quad y' = y \ , \quad z' = z \tag{136}$$

where the gamma factor,

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{137}$$

and c is the speed of light, $c = 10^8 ms^{-1}$.

Special Relativity (SR) is based on the Minkowski space-time geometry. It is essentially the natural extension of Euclidean space to include time. We begin with $\mathbb{R}^4$ with coordinates x^μ with $\mu = 0, 1, 2, 3$ and $x^\mu = \{t, x, y, z\}$. Note the use of Greek indices to denote space-time coordinates. We continue to use $x^i = \{x, y, z\}$ with Roman indices $i = 1, 2, 3$ to represent only the spatial coordinates.

We now may immediately identify this Lorentz transformation as a *special* type of coordinate transform $x^\mu \to x'^{\mu'} = x'^{\mu'}(x)$ which can be written as;

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma \frac{v}{c^2} & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \tag{138}$$

and similarly for boosts in the y and z directions.

The key point is that the Lorentz transformation preserves the length of a space-time interval, so;

$$
\begin{aligned}
(\Delta s)^2 &= -c^2(\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 \\
&= -c^2(\Delta t')^2 + (\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2
\end{aligned} \tag{139}
$$

In more 'grown-up' language the Lorentz transformation leaves *invariant* the line element or *spacetime interval*,

$$
ds^2 = -c^2 dt^2 + \delta_{ij} dx^i dx^j = \eta_{\mu\nu} dx^\mu dx^\nu \tag{140}
$$

which defines a (Lorentzian) metric,

$$
\eta_{\mu\nu} = \begin{pmatrix} -c^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{141}
$$

Note this is *not* a Riemannian metric - it has a negative eigenvalue.

Relativistic units:

Recall from astrophysics/cosmology the familiar choice of measuring time in units of years, and measuring distance in *light years.* In these (natural) units the speed of light is simply $c = 1$ and becomes dimensionless.

Example: Let us choose units of minutes for time. It takes 8 minutes for light to reach us from the sun so time $T = 8$ in these units. The sun is 8 light minutes away so is a distance $D = 8$ in these units. Hence the speed of light, $c = D/T = 1$.

From this point on we will work in analogous *natural units* where $c = 1$ and is dimensionless.

More conventionally one derives units of time from distance. Ex. 1 light meter $\simeq 0.3 \times 10^{-8} sec$

In these natural units

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{142}$$

so that, so that $\eta_{00} = -1$, $\eta_{0i} = \eta_{i0} = 0$ and $\eta_{ij} = \delta_{ij}$. This is the **Minkowski metric** and gives the geometry the **Minkowski spacetime**.

The spacetime interval

Then we have that a space-time interval dx^μ is;

- *spacelike* if $ds^2 > 0$; ie. an infinitessimal displacement dx^μ corresponds to a *distance* separation $\sqrt{+ds^2}$.
- *timelike* if $ds^2 < 0$; ie. an infinitessimal displacement dx^μ corresponds to a *time* separation $\sqrt{-ds^2}$.
- *null* or *light like* if $ds^2 = 0$.

Note that this space-time interval is invariant under Lorentz transformation so all observers agree on which class an interval is.

Inertial observers: Recall that inertial observers just follow straight lines;

$$x^\mu(\tau) = a^\mu + v^\mu \tau \tag{143}$$

for constants a^μ and v^μ where τ is the proper time measured by that observer, and v^μ is their 4-velocity.

Since τ is proper time, we must have $ds^2 = -d\tau^2$. Hence,

$$\eta_{\mu\nu} v^\mu v^\nu = \eta_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = \left(\frac{ds}{d\tau} \right)^2 = -\left(\frac{d\tau}{d\tau} \right)^2 = -1 \tag{144}$$

Hence the 4-velocity is timelike with unit norm, so we may write it as,

$$v^\mu = \left(\gamma, \gamma v^i \right) \tag{145}$$

where v^i is the 3-velocity of the particle in the inertial frame of the coordinates x^μ.

Accelerated observers follow a general curve $x^\mu(\tau)$.

Then instantaneous 4-velocity is the derivative of space-time position of the observer with respect to *their* proper time;

$$v^\mu = \frac{dx^\mu(\tau)}{d\tau} \tag{146}$$

and as above is unit norm and timelike.

Rest frame: The rest frame of an internal observe is one where $v^\mu = (1, 0, 0, 0)$ - remember for an inertial observer v^μ is constant, independent of τ. For an accelerated observer at some time τ one may Lorentz transform to their *instantaneous rest frame* where $v^\mu(\tau) = (1, 0, 0, 0)$ for a particular values of τ.

Comment: The component v^0 gives the time dilation of a clock carried by the moving observer as measured by a static observer in the reference frame x^μ whose passes through the same space-time point.

4-momentum of the particle is $p^\mu = mv^\mu$ with m the *rest mass* of the particle. Recall that p^0 is interpreted as the energy of the particle, E, measured by a static observer in the reference frame x^μ who meets the particle; p^i are the usual 3-momentum.

Comment: In the rest frame of the particle we have the famous $p^\mu = (m, 0, 0, 0)$, so $E = m$ (recall we have units so $c = 1$ - really $E = mc^2$). As measured in the reference frame x^μ, a moving particle has energy $E = \gamma\, m$, and, (putting back factors of 'c');

$$\gamma mc^2 = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} \simeq mc^2 \left(1 + \frac{1}{2}\frac{v^2}{c^2} + \ldots \right) = mc^2 + \frac{1}{2}mv^2 + \ldots \tag{147}$$

so the energy as low speeds has a rest mass component and then the usual Newtonian kinetic energy.

Generally an observer at point p with 4-velocity $\bar{v}^\mu$ measures a particle with 4-momentum p^μ at point p to be;

$$E = -\eta_{\mu\nu}\bar{v}^\mu p^\nu = -\bar{v}_\mu p^\mu \tag{148}$$

In relativistic mechanics it is the 4-momentum that is conserved - hence both energy and momentum as usual.

The generalisation of Newton's law is;

$$f^\mu = \frac{dp^\mu}{d\tau} \tag{149}$$

for a 4-vector force f^μ. Correspondingly the 4-vector acceleration is $f^\mu = ma^\mu$. We see for an inertial observer there is no force as p^μ is constant in τ.

Comment: In the instantaneous rest frame of an accelerates particle, so that $v^\mu = (1, 0, 0, 0)$ the force 4-vector should take the form $f^\mu = (0, f^i)$, with f^i the usual Newton force 3-vector.

Light rays and photons

Recall that the light rays follow null straight lines;

$$x^\mu(\lambda) = a^\mu + \lambda b^\mu \tag{150}$$

for b^μ a null vector.

Photons have null 4-momentum p^μ, and their energy is again given by $E = -\bar{v}_\mu p^\mu$, or in the reference frame of the coordinates $E = p^0$.

4 The geometry of Special Relativity

Let us for the moment ignore the fact that the Minkowski metric doesn't have all positive eigenvalues, and try to treat it as the metric and see what we find. In fact we will see that many of the physical features of SR have a simply geometric description.

4.1 Lorentz transformations

The set of coordinate transformations that leave the Minkowski metric invariant - ie. the *symmetries of the Minkowski geometry* - are the translations together with Lorentz transformations,

$$x'^{\mu'} = a^{\mu'} + b^{\mu'}{}_{\mu} x^{\mu} \tag{151}$$

with $a^{\mu'}, b^{\mu'}{}_{\mu}$ constants and $b^{\mu'}{}_{\mu} \in O(1,3)$ where $O(1,3)$ are the Lorentz transformations. These are a *global* transformation - all points are transformed in the same way.

The definition of an $O(1,3)$ matrix $\mathbf{A}$ is one that leaves invariant $\eta_{\mu\nu}$ so,

$$\eta = \mathbf{A}^T \cdot \eta \cdot \mathbf{A} \tag{152}$$

Again the determinant is ± 1.

In fact we will always concern ourselves only with the transformations preserving *both* spatial orientation and the direction of time - the *proper orthochronous* Lorentz transformations. This is the group $SO^+(1,3)$ and forms a subgroup of the Lorentz group $O(1,3)$. Note that $\det \mathbf{A} = +1$ for $\mathbf{A} \in SO^+(1,3)$.

[We may write; $O(1,3) = SO^+(1,3) \times \{1, T, P, T \cdot P\}$, where T it time reversal $(t, x^i) \to (-t, x^i)$, and P is parity reveral $(t, x^i) \to (t, -x^i)$.]

Now we see that $M^{\mu'}{}_{\mu} = b^{\mu'}{}_{\mu} \in SO^+(1,3)$, and again if $M^{\mu'}{}_{\mu} \in SO^+(1,3)$ then its inverse is also, $M^{\mu}{}_{\mu'} \in SO^+(1,3)$.

The metric $\eta_{\mu\nu}$ transforms as,

$$\eta'_{\mu'\nu'} = \eta_{\mu\nu} M^{\mu}{}_{\mu'} M^{\nu}{}_{\nu'} \tag{153}$$

and hence if $M^{\mu}{}_{\mu'} \in SO^{+}(1,3)$ then we have $\eta_{\mu\nu}$ is invariant;

$$\eta_{\mu'\nu'} = \eta_{\mu\nu} M^{\mu}{}_{\mu'} M^{\nu}{}_{\nu'} \tag{154}$$

Since the above is symmetric in μ, ν this constitutes 10 equations. The $SO^{+}(1,3)$ matrix has 16 components. The $SO^{+}(1,3)$ Lorentz transformation has $16 - 10 = 6$ parameters. There are 3 rotations and 3 boosts, as we discuss later in more detail.

The Lorentz transformations together with translations in time and space give the *Poincare* transformations.

Comment: This is analogous to the Galilean group (the translations + rotations) that leave the Euclidean metric invariant (see Ex Sheet 4).

4.2 The geometry of Minkowski space-time

Consider the Minkowski geometry. In Minkowski coordinates this defines a metric,

$$g_{\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{155}$$

However, in general the metric in other coordinates will look like;

$$g_{\mu'\nu'} = M^{\mu}{}_{\mu'} M^{\nu}{}_{\nu'} \eta_{\mu\nu} \tag{156}$$

For example, in spherical coordinates (t, r, θ, ϕ) we have,

$$g_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2 \sin^2\theta \end{pmatrix} \tag{157}$$

An infinitesimal displacement dx^μ in the coordinates is timelike/spacelike/null is $ds^2 = g_{\mu\nu}dx^\mu dx^\nu < 0, > 0 = 0$ respectively. If timelike then the infinitesimal proper time is $d\tau = \sqrt{-ds^2}$, and it space like the infinitesimal proper length is ds.

4.2.1 Vectors in Minkowski geometry

A vector v^μ at a point has norm $|v|^2 = g_{\mu\nu}v^\mu v^\nu$ and is *timelike* if $|v|^2 < 0$, *spacelike* if $|v|^2 > 0$ and *null* if $|v|^2 = 0$.

Key point: Since the norm is coordinate invariant - ie. is a scalar - then all observers using any coordinates must agree on whether a vector is timelike, spacelike or null.

In Minkowski coordinates the norm is just $|v|^2 = \eta_{\mu\nu}v^\mu v^\nu = -(v^0)^2 + (v^i)^2$.

In a coordinate system where the coordinate $t = x^0$ is timelike, meaning, the interval $dx^\mu = (dt, 0, 0, 0)$ is timelike, then we say a timelike or null vector is *future directed* if $v^0 > 0$. Otherwise it is *past directed*.

For coordinate transforms that preserve the direction of time (such as proper Lorentz) they also agree of future or past direction. *We shall restrict to such coordinate transforms from now on.*

As usual we define the product between vectors v, u as $u \cdot v = g_{\mu\nu}u^\mu v^\nu$. If this vanishes we say the vectors are *orthogonal*. Note a null vector is orthogonal to itself!

4.2.2 Example: proper Lorentz transforms of vectors

Consider Minkowski coordinates, so $g_{\mu\nu} = \eta_{\mu\nu}$. Then by a (proper) Lorentz transformation we may always find a frame so that a vector v^μ that is...

- ... future timelike with norm $|v|^2 = -a^2$ has components $v^\mu = (|a|, 0, 0, 0)$. This is its *rest frame*.
- ... spacelike with norm $|v|^2 = +b^2$ has components, $v^\mu = (0, b, 0, 0)$.

- ... null has components $v^\mu = (1, 1, 0, 0)$.

[From a group theory point of view; the proper Lorentz transform preserves a vector's norm, and for timelike/null vectors preserves future/past directedness, but otherwise has a *transitive* action.]

4.2.3 Curves in Minkowski geometry

Consider a curve $x^\mu(\lambda)$. A curve is defined to be a *timelike curve* if the tangent vector $v^\mu = dx^\mu/d\lambda$ is everywhere timelike. Likewise we have *space like* and *null* curves.

Comment: Note generally a curve will have indefinite character, where the tangent varies at different points between timelike, space like and null, so these timelike, space like and null curves are rather special ones.

Spacelike curve

The proper length s along a curve is measured as,

$$s = \int d\lambda \sqrt{\left(g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \right)} \tag{158}$$

and is reparameterization invariant.

We may use the proper length s as a parameter, so $x^\mu(s)$. This gives an affine parameterization as then,

$$+1 = g_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} \tag{159}$$

Timelike curve

The proper time τ along a curve is measured as,

$$\tau = \int d\lambda \sqrt{\left(-g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \right)} \tag{160}$$

and is reparameterization invariant.

We may use the proper time τ as a parameter for the trajectory, so $x^\mu(\tau)$. This gives an affine parameterization as then,

$$-1 = g_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \tag{161}$$

Null curve

A null curve $x^\mu(\lambda)$ has $g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} = 0$. Note that a massless curve is *always* in an affine parameterization. Any reparameterization is also affine.

4.2.4 Geodesics

Recall geodesics are curves with tangents $v^\mu = dx^\mu / d\lambda$ such that in an affine parameterization so, $g_{\mu\nu} v^\mu v^\nu =$ constant then,

$$v^\mu \nabla_\mu v^\nu = 0 \qquad \Longrightarrow \qquad \frac{d^2 x^\mu}{d\lambda^2} + \Gamma^\mu{}_{\alpha\beta} \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} = 0 \tag{162}$$

Since $|v|^2$ is constant we see a geodesic will either be timelike, space like or null.

For a spacelike or timelike geodesic we may always choose the proper length s or time τ respectively as the affine parameter.

In Minkowski coordinates then $g_{\mu\nu} = \eta_{\mu\nu}$ so that $\Gamma^\mu{}_{\alpha\beta} = 0$. Then the geodesics simply obey for constant a^μ, b^μ,

$$x^\mu = a^\mu + b^\mu \lambda \;, \quad \text{with} \quad \eta_{\mu\nu} b^\mu b^\nu = \begin{cases} -1 & \text{timelike} \;, \quad \lambda = \tau \\ 0 & \text{null} \\ +1 & \text{spacelike} \;, \quad \lambda = s \end{cases} \tag{163}$$

Thus we see that a *key* physical point of SR is geometric;

Non-accelerated massive particles/observers follow timelike geodesics.

Non-accelerated massless particles follow null geodesics.

4.2.5 'Newton's laws'

In general the relativistic 'Newton' laws for a massive particle with trajectory $x^\mu(\tau)$, for affine proper time τ so $v^\mu = \frac{dx^\mu}{d\tau}$ and $|v|^2 = -1$, takes the form;

$$a^\mu = v^\alpha \nabla_\alpha v^\mu \ , \quad f^\mu = m a^\mu \tag{164}$$

where m is the rest mass and the vectors a^μ and f^μ are the 4-acceleration and 4-force.

Note: If $a^\mu = 0$, then the particle follows a geodesic.

Check: In Minkowski coordinates;

$$a^\mu = \frac{dx^\alpha}{d\tau}\nabla_\alpha v^\mu = \frac{dx^\alpha}{d\tau}\frac{\partial}{\partial x^\alpha} v^\mu = \frac{d^2 v^\mu}{d\tau^2} \tag{165}$$

as we saw earlier.

4.2.6 Isometries of Minkowski

We said earlier that the Poincare transformations are symmetries of the Minkowski metric. More precisely the Poincare transformations generate the *isometries* of the geometry. In Minkowski coordinates they look like;

$$x'^{\mu'} = a^{\mu'} + b^{\mu'}{}_{\mu} x^\mu \tag{166}$$

with $a^{\mu'}, b^{\mu'}{}_{\mu}$ constant and $b^{\mu'}{}_{\mu} \in SO^+(1,3)$. In other coordinate systems the Poincare transformations may look very complicated - for example in spherical coordinates.

We may write the three boosts and three rotations as follows;

Rotations: There are 3 rotations; about x, y and z;

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta_1 & \sin\theta_1 \\ 0 & 0 & -\sin\theta_1 & \cos\theta_2 \end{pmatrix} , \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_2 & \sin\theta_2 & 0 \\ 0 & -\sin\theta_2 & \cos\theta_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_3 & 0 & \sin\theta_3 \\ 0 & 0 & 1 & 0 \\ 0 & -\sin\theta_3 & 0 & \cos\theta_3 \end{pmatrix}$$

and two more about x and y.

Boosts: There are 3 boosts; along x, y and z,

$$\mathbf{M} \;=\; \begin{pmatrix} \cosh\alpha_1 & -\sinh\alpha_1 & 0 & 0 \\ -\sinh\alpha_1 & \cosh\alpha_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} \cosh\alpha_2 & 0 & -\sinh\alpha_2 & 0 \\ 0 & 1 & 0 & 0 \\ -\sinh\alpha_2 & 0 & \cosh\alpha_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$
$$\begin{pmatrix} \cosh\alpha_3 & 0 & 0 & -\sinh\alpha_3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\alpha_3 & 0 & 0 & \cosh\alpha_3 \end{pmatrix}$$

where $\cosh\theta = \gamma$ and $\sinh\theta = \gamma\, v$.

[Hence this looks very close to a rotation - in fact it is a rotation if you analytically continue time t to Euclidean time τ as $t = i\tau$ and take $\alpha = i\theta$ and θ is a usual rotation.]

Any element of the Lorentz group can be written in terms of *generators*, matrices $\mathbf{T}^{(A)}$ with $A = 1, 2, \ldots, 6$, where,

$$\mathbf{M} = e^{\mathbf{T}^{(A)}\theta^{(A)}} = \mathbf{1} + \left(\mathbf{T}^{(A)}\theta^{(A)}\right) + \frac{1}{2!}\left(\mathbf{T}^{(A)}\theta^{(A)}\right)^2 + \ldots \tag{167}$$

and the six quantities $\theta^{(A)}$ parameterize the transformation, just like 3 angles parameterize a rotation in 3-d Euclidean space.

For the Lorentz group we may write these generators conveniently as,

$$\theta^{(A)} T^{(A)\mu}{}_{\nu} = c^{\alpha\beta}\left(\delta^{\mu}_{\alpha}\eta_{\nu\beta} - \delta^{\mu}_{\beta}\eta_{\nu\alpha}\right) \tag{168}$$

for constants $c^{\alpha\beta}$.
The quantity $c^{\alpha\beta}$ is antisymmetric, so $c^{\alpha\beta} = c^{[\alpha\beta]}$, and an antisymmetric 4×4 matrix has 6 components. To agree with the previous matrices $\mathbf{M}$ above we have;

$$c^{[\alpha\beta]} = \begin{pmatrix} 0 & -\alpha_1 & -\alpha_2 & -\alpha_3 \\ & 0 & \theta_2 & \theta_3 \\ & & 0 & \theta_1 \\ & & & 0 \end{pmatrix} \tag{169}$$

Using these generators we may simply write down the vector fields that generate the isometries of Minkowski spacetime. In Minkowski coordinates these are,

$$v^\mu = a^\mu + \theta^{(A)} T^{(A)\mu}{}_{\gamma} x^\gamma = a^\mu + c^{\alpha\beta}\left(\delta^\mu_\alpha \eta_{\gamma\beta} - \delta^\mu_\beta \eta_{\gamma\alpha}\right) x^\gamma \tag{170}$$

for the 4 constants a^μ parameterizing translations, and the 6 constants $c^{[\alpha\beta]}$ parameterizing Lorentz transforms.

Recall that isometries obey,

$$Lie(v,g)_{\mu\nu} = v^\sigma \partial_\sigma g_{\mu\nu} + g_{\mu\sigma}\partial_\nu v^\sigma + g_{\nu\sigma}\partial_\mu v^\sigma \tag{171}$$

Then in Minkowski coordinates $g_{\mu\nu} = \eta_{\mu\nu}$ is constant and a^μ and $c^{\alpha\beta}$ are constant, so,

$$\begin{aligned}
Lie(v,g)_{\mu\nu} &= \eta_{\mu\sigma}\partial_\nu\left(c^{\alpha\beta}\left(\delta^\sigma_\alpha\eta_{\gamma\beta} - \delta^\sigma_\beta\eta_{\gamma\alpha}\right)x^\gamma\right) + \eta_{\nu\sigma}\partial_\mu\left(c^{\alpha\beta}\left(\delta^\sigma_\alpha\eta_{\gamma\beta} - \delta^\sigma_\beta\eta_{\gamma\alpha}\right)x^\gamma\right) \\
&= \eta_{\mu\sigma}c^{\alpha\beta}\left(\delta^\sigma_\alpha\eta_{\gamma\beta} - \delta^\sigma_\beta\eta_{\gamma\alpha}\right)\partial_\nu\left(x^\gamma\right) + \eta_{\nu\sigma}c^{\alpha\beta}\left(\delta^\sigma_\alpha\eta_{\gamma\beta} - \delta^\sigma_\beta\eta_{\gamma\alpha}\right)\partial_\mu\left(x^\gamma\right) \\
&= c^{\alpha\beta}\left(\eta_{\mu\alpha}\eta_{\gamma\beta} - \eta_{\mu\beta}\eta_{\gamma\alpha}\right)\delta^\gamma_\nu + c^{\alpha\beta}\left(\eta_{\nu\alpha}\eta_{\gamma\beta} - \eta_{\nu\beta}\eta_{\gamma\alpha}\right)\delta^\gamma_\mu \\
&= c^{\alpha\beta}\left(\eta_{\mu\alpha}\eta_{\nu\beta} - \eta_{\mu\beta}\eta_{\nu\alpha}\right) + c^{\alpha\beta}\left(\eta_{\nu\alpha}\eta_{\mu\beta} - \eta_{\nu\beta}\eta_{\mu\alpha}\right) \\
&= c^{\alpha\beta}\left(\eta_{\mu\alpha}\eta_{\nu\beta} - \eta_{\mu\beta}\eta_{\nu\alpha}\right) - c^{\beta\alpha}\left(\eta_{\nu\alpha}\eta_{\mu\beta} - \eta_{\nu\beta}\eta_{\mu\alpha}\right) \\
&= 0
\end{aligned} \tag{172}$$

which vanishes by symmetry.

5 Continuous matter in Minkowski spacetime

We have now understood how to think about point particle matter in Special relativity from a geometric point of view. Recall massive/massless

We will now think about writing down the physical laws that govern matter. We will start with electromagnetism, and then progress to fluids. An important concept we will need later is the *stress tensor*.

5.1 Physical laws as geometry

Since for $g_{\mu\nu} = \eta_{\mu\nu}$ then $\Gamma^{\mu}{}_{\alpha\beta} = 0$ then in Minkowski coordinates a covariant derivative $\nabla_\mu = \partial_\mu$.

Note that physical laws *must* be tensor expressions or else they would depend on our choice of coordinates!

Given physical laws in Minkowski coordinates, we get their general form by being careful to write down tensor expressions and noting that generally $\partial_\mu \to \nabla_\mu$.

5.2 Electromagnetism

In *Minkowski coordinates* we write electromagnetism in terms of $F_{\mu\nu}$, the antisymmetric field strength where given the electric field E^i and magnetic field B^i then,

$$F^{0i} = E^i \ , \quad F^{ij} = \epsilon^{ijk} B_k \tag{173}$$

The Maxwell equations can be written as,

$$\partial_\mu F^{\mu\nu} = j^\nu \ , \quad \partial_{[\nu} F_{\alpha\beta]} = 0 \tag{174}$$

for a 4-charge current j^μ where j^0 is the charge density ρ in the frame x^μ and j^i is the charge current.

Note that symmetric implies $0 = \partial_\mu \partial_\nu F^{\mu\nu} = \partial_\mu j^\mu$. This is the local conservation of charge equation;

$$\partial_\mu j^\mu = 0 \tag{175}$$

In these Minkowski coordinates we have,

$$\partial_t \rho = -\nabla \cdot \mathbf{j} \tag{176}$$

Hence, the total charge in a volume V, $Q = \int dV \rho$ is constant;

$$\frac{d}{dt} Q = -\int dV \partial_t \rho = -\int dV \nabla \cdot \mathbf{j} = -\int dS \hat{\mathbf{n}} \cdot \mathbf{j} \tag{177}$$

with $\hat{\mathbf{n}}$ the outer unit normal to the surface element, so $\hat{\mathbf{n}} \cdot \hat{\mathbf{n}} = 1$. We see that if the current through the surface vanishes, the total charge is conserved. Hence the equation $\partial_\mu j^\mu = 0$ governs charge conservation. In general we see from the above equation that the rate of change of charge in a volume is given by the total current entering the volume.

Now in a general coordinate system for the Minkowski geometry, so that $g_{\mu\nu} \neq \eta_{\mu\nu}$ and $\nabla_\mu \neq \partial_\mu$ as the tensor equations;

$$\nabla_\mu F^{\mu\nu} \ = \ j^\nu \quad , \quad \nabla_{[\nu} F_{\alpha\beta]} = 0 \tag{178}$$

where $F_{\mu\nu}$ is an antisymmetric tensor field, and j^μ is a vector field which obeys, $\nabla_\mu j^\mu = 0$.

5.3 Newtonian equilibrium continuum mechanics

Let us first review the stress tensor in usual Newtonian mechanics.

In order to describe continuous matter, possibly deformed by some external surface forces, we must introduce the *stress tensor*. This is the symmetric $(2, 0)$ tensor σ^{ij} that measures stresses within a material.

Consider slicing open a material at equilibrium along some surface. In order to keep it unchanged we have to apply forces to the surface that were provided by the *internal forces* we have removed.

Consider a surface element area ΔS with unit normal $\hat{\mathbf{n}}$. The force we must provide to match the internal force is $\mathbf{\Delta f}$. Then;

- $\mathbf{\Delta f}$ is a *normal stress* if $\mathbf{\Delta f}$ is parallel to $\hat{\mathbf{n}}$. The normal stress is defined as $\sigma_{normal} \equiv \frac{|\mathbf{\Delta f}|}{\Delta S}$, and can be thought of as a 'pressure'.
- $\Delta\mathbf{f}$ is a *shear stress* if $\Delta\mathbf{f}$ is normal to $\hat{\mathbf{n}}$ ie. tangent to the surface. The shear stress is defined as $\sigma_{shear} \equiv \frac{|\Delta\mathbf{f}|}{\Delta S}$.

Cauchy: In continuum mechanics the fundamental assumption is that at any point in a material we may write the infinitessimal internal force vector $\mathbf{df}$ acting on an infinitesimal surface with normal $\mathbf{dS} = \hat{\mathbf{n}}\, dS$ in terms of a $(0, 2)$ stress tensor $\sigma_{ij}(x)$ in Euclidean space;

$$df^i = \sigma^{ij}(x)\, dS_j \tag{179}$$

Conservation: Consider the internal forces acting over the surface of some closed volume of material V. Assuming the total external force on the body vanishes and it is in equilibrium, then the sum of the internal forces acting on a closed subset must also vanish. Taking Cartesian coordinates, so $g_{ij} = \delta_{ij}$ and $\nabla_i = \partial_i$ then,

$$0 = \int_S df^i = \int_S \sigma^{ij}(x)\, dS_j = \int_V \partial_j \sigma^{ij}(x)\, dV \tag{180}$$

Now this must be true for *any* surface S, which implies,

$$\partial_j \sigma^{ij}(x) = 0 \tag{181}$$

everywhere within the material.

Since σ^{ij} is a tensor, then the real expression in general coordinates (eg. spherical polar) in Euclidean space must be;

$$\nabla_j \sigma^{ij}(x) = 0 \tag{182}$$

Symmetry: Likewise the moments about any point, say x_0^j, from the surface forces df^i must also vanish for a closed surface. Hence,

$$\begin{aligned} 0 = \int_S \epsilon_{ijk}(x^j - x_0^j) df^k &= \int_S \epsilon_{ijk}(x^j - x_0^j)\sigma^{km}(x)\, dS_m = \int_V \partial_m \left(\epsilon_{ijk}(x^j - x_0^j)\sigma^{km}(x)\right) dV \\ &= \int_V \left(\epsilon_{ijk}(\partial_m x^j)\sigma^{km}(x) + (x^j - x_0^j)\partial_m \sigma^{km}(x)\right) dV \\ &= \int_V \epsilon_{ijk}\delta^j_m \sigma^{km}(x)\, dV = \int_V \epsilon_{ijk}\sigma^{kj}(x)\, dV \end{aligned} \tag{183}$$

where we have used the fact that $\partial_j \sigma^{ij}(x) = 0$. As this must be true for any closed surface,

$$\epsilon_{ijk}\sigma^{kj}(x) = 0 \quad \Longrightarrow \quad \sigma^{ij}(x) = \sigma^{ji}(x) \tag{184}$$

and hence the stress tensor is symmetric.

Note this symmetry means that σ^{ij} can be diagonalized by a coordinate transformation, leading to the existence of the *principle normal stresses.*

Comment: The component σ^{ij} gives;

σ^{ij} = (force in i direction) / (area normal to j)

but using force = d (momentum) / d(time) then;

σ^{ij} = (momentum in i direction) / (time * area normal to j)
= momentum flux in i direction through surface normal to j.

5.4 Relativistic continuum mechanics

Just as momentum and energy get combined into a 4-vector relativistic momentum, this momentum flux also is combined with energy flux and energy density to give a relativistic *symmetric* $(2,0)$ stress tensor $T^{\mu\nu}$.

Definition of the stress tensor: At a point we choose the rest frame of the matter there (ie. its total 3-momentum is zero at that point) and then;

- T^{00} = *rest* mass density of the matter
- $T^{0i} = 0$
- T^{ij} = the usual stress tensor ie. *momentum flux* in direction of x^i through surface normal to x^j.

Note that since T^{ij} is symmetric, then $T^{\mu\nu}$ must be symmetric - if it is symmetric in one frame, it must be in all frames.

In a general frame the stress tensor is;

- T^{00} = *energy density* of the matter
- T^{0i} = *energy flux* through surface normal to x^i (or *momentum density* of the matter in the x^i direction)
- T^{ij} = *momentum flux* in direction of x^i through surface normal to x^j.

For an observer moving with 4-velocity $\bar{v}^\mu$ they observe the local 4-momentum of the matter, p^μ, to be;

$$p^\mu = -\bar{v}_\nu T^{\mu\nu} \tag{185}$$

In Minkowski space-time, taking the usual Minkowski coordinates x^μ so that $g_{\mu\nu} = \eta_{\mu\nu}$ we may define the total 4-momentum measured in the frame x^μ in a subset V of material;

$$P^\mu = \int dV T^{0\mu} \tag{186}$$

Hence,

$$\frac{d}{dt}P^\mu = \int dV \frac{d}{dt}T^{0\mu} \tag{187}$$

However, as the material moves, the rate of change of this momentum must be equal to the sum of the 4-forces acting on this volume - these are the internal forces that act at the surface. Hence,

$$\frac{d}{dt}P^\mu = \int dS \hat{n}^i T^{i\mu} \tag{188}$$

Equating these expressions and using usual Minkowski coordinates x^μ so the metric is just $\eta_{\mu\nu}$ and $\nabla_\mu = \partial_\mu$ gives;

$$\int dV \frac{d}{dt} T^{0\mu} = \int dS \hat{n}^i T^{i\mu} = \int dV \partial_i T^{i\mu} \tag{189}$$

using the divergence theorem, and hence,

$$0 = \int dV \left(-\frac{d}{dt} T^{0\mu} + \partial_i T^{i\mu} \right) = \int dV \partial_\mu T^{\mu\nu} \tag{190}$$

Now since this must be true for any volume we have,

$$\partial_\mu T^{\mu\nu} = 0 \tag{191}$$

everywhere in the material.

Comment: For a state/equilibrium situation this reproduces the previous equilibrium condition $\partial_i T^{ij} = 0$.

We have used Minkowski coordinates. In general coordinates in the Minkowski geometry the answer is;

$$\boxed{\nabla_\mu T^{\mu\nu} = 0} \tag{192}$$

5.5 Stress tensor for electromagnetism

The stress tensor for electromagnetism is,

$$T_{\mu\nu} = F_{\mu\alpha} F_{\nu}{}^{\alpha} - \frac{1}{4} g_{\mu\nu} F_{\alpha\beta} F^{\alpha\beta} \tag{193}$$

Note that indeed $\nabla_\mu T^{\mu\nu} = 0$ provided the EM equations hold (which requires current conservation $\nabla_\mu j^\mu = 0$).

In Minkowski coordinates you should find that,

$$T^{00} \sim \mathbf{E}^2 + \mathbf{B}^2 \,, \quad T^{0i} \sim \mathbf{E} \times \mathbf{B} \tag{194}$$

giving the energy density and momentum density (the Poynting vector!).

5.6 Perfect fluid matter

A perfect fluid is matter that is described simply by a local density, ρ, and 4-velocity, u^μ (so that $u^2 = -1$). The pressure in the fluid is isotropic - no preferred direction - and hence is described only by a scalar function P. This pressure is determined just in terms of the density ρ by the *equation of state*, so that $P = P(\rho)$.
Examples:

- Cold matter matter ('Dust') fluid has no pressure, so $P = 0$.
- Hot gaseous matter has pressure, so $P = \frac{1}{3}\rho$.
- Radiation fluid (eg. the gas of photons) has pressure $P = \frac{1}{4}\rho$.

The stress tensor is then determined as,

$$T^{\mu\nu} = (\rho + P)\, u^\mu u^\nu + P\eta^{\mu\nu} \tag{195}$$

In general, $u^\mu = (\gamma, \gamma v^i)$ where the 3-velocity v^i is a function of position in the fluid. However at a point in the fluid we may move to the instantaneous rest frame so $u^\mu = (1, 0, 0, 0)$ and then,

$$T^{tt} = \rho \ , \quad T^{ti} = T^{it} = 0 \ , \quad T^{ij} = P\delta^{ij} \tag{196}$$

so that T^{tt} gives the local rest mass density, and T^{ij} the stress (which is just normal for a perfect fluid).

The conservation equation then determines the evolution - or equation of motion - of the fluid when combined with the equation of state. (See example sheet 5 for the details!)

Let us work with Minkowski coordinates so $\nabla_\mu = \partial_\mu$.

Let us project our evolution equation into the direction u^μ, recalling that $u^2 = -1$. Then,

$$u_\mu \partial_\nu T^{\mu\nu} = 0 \quad \Longrightarrow \quad u^\nu \partial_\nu \rho + (\rho + P)\, \partial_\nu u^\nu = 0 \tag{197}$$

Let us now project onto an orthogonal direction n^μ to the motion u^μ, so that $u^\mu n_\mu = 0$. Note that n^μ must be space like. Then,

$$n_\mu \partial_\nu T^{\mu\nu} = 0 \quad \Longrightarrow \quad n_\mu \left((\rho + P)\, u^\nu \partial_\nu u^\mu + \eta^{\mu\nu} \partial_\mu P \right) = 0 \tag{198}$$

and you can show that this second equation is equivalent to;

$$(\rho + P)\, u^{\nu} \partial_{\nu} u^{\mu} + (\eta^{\mu\nu} + u^{\mu} u^{\nu})\, \partial_{\mu} P = 0 \tag{199}$$

These reduce to the usual non-relativistic Navier-Stokes equations if we can find a coordinate frame where the fluid everywhere obeys; $v^i \ll 1$ and $P \ll \rho$. Then one finds (See example sheet 5):

$$\partial_t \rho + \partial_i \left(\rho\, v^i \right) = 0 \,, \quad \rho \left(\partial_t v^i + v^j \partial_j v^i \right) + \partial^i P = 0 \tag{200}$$

where we recall $\partial_t \sim v^i \partial_i$ for a fluid.

Note that for general coordinates in the Minkowski geometry we would replace all $\partial_\mu \to \nabla_\mu$.

6 Curved spacetime and a geometric origin to gravity

A space-time or *Lorentzian* metric is a *symmetric* $(0, 2)$ tensor $g_{\mu\nu}$. At every point it is a symmetric matrix and hence has real eigenvalues. A Lorentzian metric must everywhere have non-vanishing eigenvalues (so it is invertible) with 1 being negative and 3 positive. We say the *signature* is $(1, 3)$.

Ex. An example is obviously the Minkowski metric.

Then this Lorentzian metric $g_{\mu\nu}$ defines the geometry of space-time in exactly the same way a Riemannian metric g_{ij} defines the geometry of space.

Our previous discussion of Riemannian geometry works in the same way for space-time/Lorentzian geometry.

Two differences with Riemannian geometry;

- Geodesics $x^\mu(\lambda)$ with affine λ no longer minimise length; instead they still extremize,

$$L = \int d\lambda g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \tag{201}$$

  [One can always make a timelike curve between two points with arbitrarily small proper time - make it move near the speed of light!]

- The geometry is locally Minkowski (rather than locally Euclidean as in the Riemannian case).

The second point is of key physical significance. We may choose the analog of Riemann normal coordinates - *local inertial coordinates* or the *local inertial frame (LIF)* - so that, at some point, say $x = 0$,

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + O(x^2) \ , \quad \Gamma^\alpha{}_{\mu\nu} = O(x) \tag{202}$$

Are these coordinates unique? Almost, but given LIF coordinates at a point, we may construct other LIF coordinates by a (global) Lorentz transformation,

$$x'^{\mu'} = \Lambda^{\mu'}{}_\mu x^\mu \tag{203}$$

for a constant matrix $\mathbf{\Lambda} \in SO^+(1,3)$. This is the only freedom in specifying the LIF coordinates.

[The proof of this follows exactly the same method as for the Riemann normal coordinates in Riemannian geometry, worked out in Qu. 3 of Ex Sheet 3.]

An important consequence of this is that the tangent space at any point again divides into timelike ($|v|^2 = g_{\mu\nu} v^\mu v^\nu < 0$), space like ($|v|^2 > 0$) and null vectors ($|v|^2 = 0$) . Again curves can be timelike, space like, null.

A timelike geodesic, parameterized by proper time τ, passing through the point $x = 0$ locally takes the form,

$$x^\mu(\tau) = v^\mu \tau + O(\tau^3) \tag{204}$$

where, $\eta_{\mu\nu} v^\mu v^\nu = -1$ so we may write,

$$v^\mu = (\gamma, \gamma v^i) \tag{205}$$

as in SR. Note the absence of the quadratic terms indicate that these curves are as 'straight' as possible locally. Since there is a residual Lorentz invariance in the coordinates, for any given timelike geodesic - and hence inertial observer - we may choose LIF coordinates such that at the point $x = 0$, we are in the instantaneous rest frame of the observer.

Thus we arrive at the key point in our discussion:

An inhabitant of space-time locally sees the geometry to be Minkowski space-time.

Hence we may say:

The physics of SR is built into the local geometry of any space-time.

Following from this *all the geometry/physics of SR naturally generalises to a general curved space-time.* We simply write Lorentz compatible laws in a geometric way, as tensor equations, taking,

$$\begin{aligned} \eta_{\mu\nu} &\rightarrow g_{\mu\nu} \\ \partial_\mu &\rightarrow \nabla_\mu \end{aligned} \tag{206}$$

and these are the laws for any space-time metric $g_{\mu\nu}$, not just coordinate transforms of Minkowski space-time.

Ex: An accelerated massive particle follows a timelike trajectory, $a^\mu = v^\nu \nabla_\nu v^\mu$ for 4-acceleration a^μ. Note that in the LIF at a point on the trajectory,

$$a^\mu = v^\nu \nabla_\nu v^\mu = v^\nu \partial_\nu v^\mu = \frac{dx^\nu}{d\tau} \partial_\nu v^\mu = \frac{dv^\mu}{d\tau} \tag{207}$$

Ex: Massless particles follow null geodesics.

Ex: The stress energy tensor is conserved as $\nabla_\mu T^{\mu\nu} = 0$.

The scale at which the quadratic corrections to $g_{\mu\nu} = \eta_{\mu\nu} + O(x^2)$ become important is the curvature scale as we see later. Below this curvature scale the geometry is Minkowski and the physics of SR is recovered.

6.1 Newtonian spacetime, gravity redshift and light bending

This section and the following ones refer to calculations performed in Ex Sheet 6. Please review that question and its solutions.

Consider Minkowski space-time, in Minkowski coordinates $x^\mu = (t, x, y, z)$, so $g_{\mu\nu} = \eta_{\mu\nu}$, and deform it a little in a specific way;

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \ , \quad \text{with} \quad g_{\mu\nu} = \eta_{\mu\nu} - 2\,\epsilon\,\Phi\,\delta_{\mu\nu} + O(\epsilon^{3/2}) \tag{208}$$

where $\epsilon \ll 1$ controls this small deformation.

The time t and spatial x^i define the *Newtonian frame.*

Taking $\epsilon \to 0$ will be the *Newtonian limit*, and we identify Φ with the *Newtonian gravitational potential* $\tilde{\Phi}$ as;

$$\tilde{\Phi} = \epsilon\,\Phi \tag{209}$$

so,

$$\delta^{ij}\partial_i\partial_j\tilde{\Phi} = \nabla^2\tilde{\Phi} = 4\pi G_N \rho \tag{210}$$

At the moment we just identify $\epsilon\Phi$ with the Newtonian potential. Later we will use the Einstein equations to show the above equation holds.

We will think of ϵ simply as an accounting tool. We compute quantities only to lowest (non-trivial) order in ϵ as $\epsilon \to 0$ and keep only that lowest term. We may make the separation of $\tilde{\Phi}$ into $\epsilon\Phi$ as,

$$\epsilon = \max|\tilde{\Phi}| \quad \Longrightarrow \quad -1 \le \Phi \le +1 \tag{211}$$

and requiring $\epsilon \ll 1$ implies that the matter with density ρ spread over a characteristic length scale R should obey,

$$\frac{G_N \rho R^2}{c^2} \ll 1 \tag{212}$$

(reinstating the factors of c). So this approximation is good for low densities spread over suitably large areas.

We will take the gravitating matter density ρ and hence Φ to be static so $\partial_t \Phi = 0$ - (or at least very slowly varying).

Non-accelerated (or inertial) massive particles follow timelike geodesics. In the Newtonian limit, $\epsilon \to 0$, the velocity of the particles must be small. Thus we write the 4-velocity;

$$\frac{dx^\mu}{d\tau} = \left(1 + \epsilon\, f + \ldots,\ \epsilon^{1/2}\, v^i + \ldots\right) \tag{213}$$

and likewise we take the acceleration to be small too,

$$\frac{d^2 x^\mu}{d\tau^2} = \left(\epsilon \frac{df}{d\tau} + \ldots,\ \epsilon\, a^i + \ldots\right) \tag{214}$$

Then in the Newtonian frame defined by x^i both the velocity and acceleration are small in the Newtonian limit, $\epsilon \to 0$, going as;

$$\begin{aligned} \frac{dx^i}{d\tau} &\simeq \epsilon^{1/2}\, v^i \\ \frac{d^2 x^i}{d\tau^2} &\simeq \epsilon\, a^i \end{aligned} \tag{215}$$

By calculating and working to lowest order in ϵ you find the inverse metric $g^{\mu\nu}$ and Christoffel symbol are;

$$\begin{aligned} g^{\mu\nu} &= \eta^{\mu\nu} + 2\epsilon\, \Phi\, \delta^{\mu\nu} + \ldots \\ \Gamma^i{}_{tt} &= \Gamma^t{}_{it} = \Gamma^t{}_{ti} = \epsilon\, (\partial_i \Phi) + \ldots \\ \Gamma^k{}_{ij} &= \epsilon\, \left((\partial_k \Phi)\, \delta_{ij} - (\partial_i \Phi)\, \delta_{jk} - (\partial_j \Phi)\, \delta_{ik} \right) + \ldots \end{aligned} \tag{216}$$

with other components, $\Gamma^t{}_{tt} = \Gamma^i{}_{jt} = \Gamma^t_{ij} = 0$.

The condition $|\frac{dx^\mu}{d\tau}|^2 = -1$ then yields,

$$f = \frac{1}{2}\delta_{ij} v^i v^j - \Phi \tag{217}$$

in the Newtonian limit, and thus the time dilation of the particle relative to the Newtonian frame is;

$$\frac{dt}{d\tau} = 1 + \epsilon\, \left(\frac{1}{2}\delta_{ij} v^i v^j - \Phi \right) + \ldots \tag{218}$$

Thus we see now only the usual SR time dilation of a moving particle due to its speed, where we recall,

$$\gamma \simeq 1 + \frac{1}{2}\delta_{ij}v^i v^j + \dots \tag{219}$$

but we see also that there is a *gravitational time dilation* due to the potential Φ.

If the particle is not accelerating then it follows a timelike geodesic;

$$\frac{d^2x^\alpha}{d\tau^2} + \Gamma^\alpha{}_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = 0 \tag{220}$$

The time component of this yields in the Newtonian limit;

$$0 = \frac{d}{d\tau}\left(\frac{1}{2}\delta_{ij}v^i v^j + \Phi\right) + \dots \tag{221}$$

ie. conservation of the usual Newtonian energy, the sum of the KE and potential energy Φ. The spatial components give,

$$a^i = -(\partial_i \Phi) + \dots \tag{222}$$

or equivalently,

$$\frac{d^2x^i}{d\tau^2} = -(\partial_i \tilde{\Phi}) + \dots \tag{223}$$

recovering the usual Newton force law for the gravitational potential $\tilde{\Phi}$ *to leading order*.

We conclude that for this space-time, slow moving (wrt the Newtonian frame) inertial observers 'feel' a force of Newtonian gravity, and obey the usual Newtonian dynamics, of Newton's theory with gravity potential Φ.

Hence: Curved space-time can account for the force of gravity!

[*Comment:* We still need a theory that tells us why to pick *this* space-time.]

6.2 Gravitational redshift

We have seen that a massive inertial particle incurs a time dilation from the gravitational potential Φ. Consider two particles sitting at fixed positions $x^i_{(1)}$ and $x^i_{(2)}$ with proper times $\tau_{(1)}$ and $\tau_{(2)}$ respectively - hence they must be accelerated, eg. by a rocket. Then their 4-velocities (from above) will be

$$v^\mu_{(1,2)} = \frac{dx^\mu_{(1,2)}}{d\tau} = \big(1 - \epsilon\Phi(x_{(1,2)}), 0, 0, 0\big) \tag{224}$$

so that $|v|^2 = -1$, and hence,

$$\frac{dt}{d\tau_{(1,2)}} \simeq 1 - \epsilon\,\Phi(x_{(1,2)}) \tag{225}$$

Hence we have,

$$\frac{d\tau_{(1)}}{d\tau_{(2)}} \simeq 1 + \epsilon\frac{\Phi(x_{(1)}) - \Phi(x_{(2)})}{c^2} \tag{226}$$

where I am being careful to include the necessary factors of c^2 if one does not use natural units.

Thus if particle one sends out radio pulses at a frequency $\omega_{(1)}$, then particle two will observe the frequency to be $\omega_{(2)}$ such that,

$$\frac{\omega_{(2)}}{\omega_{(1)}} \simeq \left(1 + \epsilon\,\frac{\Phi(x_{(1)}) - \Phi(x_{(2)})}{c^2}\right) \tag{227}$$

We see that if particle two is further from the gravity source, then $\Phi(x_{(1)}) - \Phi(x_{(2)}) < 0$, and the observed frequency is less than the emitted frequency. We call this *gravitational red shift.*

Pound-Rebka:

This effect was precisely measured in Harvard in 1959 using a $20m$ vertical separation of a gamma ray source (^{57}Fe) and absorber. Due to the gravity redshift the gamma rays reaching the source were redshifted, effecting their absorption. By vibrating the source (on a loudspeaker), a doppler shift was added so that now some gamma rays had the correct frequency and were absorbed. Thus the frequency shift due to redshift was deduced.

6.3 Light Bending

Light follows null geodesics in the Newtonian space-time. In the Ex Sheet 6 you calculate directly by solving null geodesics the deflection that a light ray experiences as it moves past a gravity point source, offset by distance R. One finds for a source at $x^i = (0, -R, 0)$, with potential,

$$\Phi = -\frac{G_N M}{\sqrt{x^2 + (y+R)^2 + z^2}} \tag{228}$$

then a light ray moving initially along the x-axis follows a curve,

$$x^\mu(\lambda) = (\lambda, \lambda, 0, 0) + \epsilon\, h^\mu(\lambda) \tag{229}$$

where $h^\mu(\lambda) \to 0$ as $\lambda \to -\infty$.

Solving the null geodesic equation to lowest order in ϵ determines $h^\mu(\lambda)$. One finds that asymptotically,

$$x(\lambda) \to \lambda \;, \quad y(\lambda) \to -\epsilon \frac{4 G_N M}{R} \lambda \tag{230}$$

and defining the deflection angle $\Delta\theta$ by,

$$\epsilon \Delta\theta = \lim_{\lambda \to \infty} \tan^{-1}\left(\frac{y(\lambda)}{x(\lambda)}\right) \tag{231}$$

gives,

$$\Delta\theta \simeq -\frac{4 G_N M}{c^2 R} \text{radians} \tag{232}$$

where I have reintroduced c, if one is not working in natural units.

Working in units where $c = 1$, then,

$$G_N = \frac{6.7 \times 10^{-11} kg^{-1} m^3 s^{-2}}{c^2} = 7.4 \times 10^{-28} kg^{-1} m \tag{233}$$

and one can calculate for the sun, radius $R \sim 7.0 \times 10^8 m$, mass $M = 2.0 \times 10^{30} kg$ that the deflection angle for a ray grazing the surface is,

$$\Delta\theta \simeq 8.5 \times 10^{-6}\, radians = 1.8\, \text{arcsec} \tag{234}$$

This was measured in May 1919 by Eddington, by observing stars positioned near the surface of the sun in the sky which were visible due to a solar eclipse. Their positions were distorted by the light bending and Eddington and his team claimed to measure this effect.

6.4 Conservation laws and symmetries in curved space time

Consider Minkowski space-time in Minkowski coordinates. Recall a timelike inertial particle follows the trajectory such that,

$$\frac{dv^\mu}{d\tau} = 0 \quad \Longrightarrow \quad p^\mu = m\, v^\mu = \text{const} \tag{235}$$

Its 4-momentum ie. its energy and 3-momentum are constant.

However, such non-local conservation laws depend strongly on the space-time having isometries. In general they do not apply, and hence we only have the local laws.

Isometries

Recall that if a vector field v^μ generates a symmetry or *isometry* of the metric then,

$$0 = \text{Lie}(v, g)_{\alpha\beta} = v^\mu \partial_\mu g_{\alpha\beta} + g_{\alpha\mu} \partial_\beta v^\mu + g_{\mu\beta} \partial_\alpha v^\mu \tag{236}$$

Recall from Ex. Sheet 4 that for the metric, this Lie derivative can actually be written in terms of the covariant derivative as;

$$\text{Lie}(v, g)_{\alpha\beta} = 2\nabla_{(\alpha} v_{\beta)} \tag{237}$$

A vector field generating an isometry is called a *Killing vector*, and satisfies,

$$\nabla_{(\alpha} v_{\beta)} = 0 \tag{238}$$

Energy conservation

Suppose we have a Killing vector field v^μ which is everywhere timelike and an inertial massive particle following a timelike geodesic with affine parameter τ and tangent $u^\mu = dx^\mu/d\tau$ so that $v^2 = -1$. Then as you showed in Ex Sheet 4;

$$\begin{aligned} \frac{d}{d\tau}\left(u^\mu v_\mu\right) &= u^\nu \partial_\nu \left(u^\mu v_\mu\right) = u^\nu \nabla_\nu \left(u^\mu v_\mu\right) \\ &= v_\mu u^\nu \nabla_\nu u^\mu + u^\mu u^\nu \nabla_\nu v_\mu = 0 \end{aligned} \tag{239}$$

the first term vanishing as u^μ is tangent to an affinely parameterized geodesic, and the second term vanishing for the Killing vector v^μ we have, $\nabla_{(\mu} v_{\nu)} = 0$ which implies $\nabla_\mu v_\nu = -\nabla_\nu v_\mu$. Hence $\nabla_\mu v_\nu$ is antisymmetric in its indices, but above is contracted with the symmetric $u^\mu u^\nu$.

Suppose the particle has rest mass m, and thus 4-momentum $p^\mu = m u^\mu$. Then,

$$\frac{d}{d\tau} \left(p^\mu v_\mu \right) = 0 \tag{240}$$

Hence we see that existence of the Killing vector implies that the momentum projected in the Killing direction is conserved.

Recall that a timelike observer with 4-velocity v^μ measures the particle's energy to be $E = -p^\mu v_\mu$. Thus observers (not necessarily inertial ones) moving with tangent v^μ all measure the same energy for the particle as it passes them. Thus energy of the particle is *conserved* as measured by *these* observers.

In general, timelike isometries give rise to energy conservation laws, and space like isometries give rise to (angular) momentum conservation laws.

Example: In the Newtonian space-time, co-ordinates $x^\mu = (t, x, y, z)$, with static potential, $\partial_t \Phi$, then the vector $v^\mu = (1, 0, 0, 0)$ is a Killing vector. Hence we saw the Newtonian energy $E = \frac{1}{2} \delta_{ij} v^i v^j + \Phi$ was conserved.

Example: In Newtonian space-time take a point source potential. This has a spherical symmetry, but not translations. In coordinates $x^\mu = (t, r, \theta, \phi)$ with the source at $r = 0$, then for example $v^\mu = (0, 0, 0, 1)$ is an isometry. The corresponding conserved quantity is angular momentum. With no translations, linear momentum is not conserved.

7 Curvature

We have seen that any geometry is locally Euclidean/Minkowski. The quantity that characterises the deviation away from this simple geometry is the curvature, or more precisely the Riemann curvature tensor.

7.1 The Riemann tensor

Suppose we have a function f. Then the commutator of two covariant derivatives vanishes,

$$\begin{aligned}[\nabla_\alpha, \nabla_\beta] f &\equiv \nabla_\alpha \partial_\beta f - \nabla_\beta \partial_\alpha f \\ &= \left(\partial_\alpha \partial_\beta f - \Gamma^\mu{}_{\alpha\beta} \partial_\mu f\right) - \left(\partial_\beta \partial_\alpha f - \Gamma^\mu{}_{\beta\alpha} \partial_\mu f\right) = 0 \end{aligned} \quad (241)$$

However for a covector field v_μ such a commutator does not generally vanish; in fact it can be written as,

$$[\nabla_\alpha, \nabla_\beta] v_\mu = R_{\alpha\beta\mu}{}^\nu v_\nu \quad (242)$$

where $R_{\alpha\beta\nu}{}^\mu$ is the *Riemann* $(1,3)$ tensor field.

We may explicitly evaluate the Riemann tensor in terms of the Christoffel symbol using;

$$\begin{aligned}\nabla_\alpha \nabla_\beta v_\mu &= \partial_\alpha \left(\nabla_\beta v_\mu\right) - \Gamma^\nu{}_{\alpha\beta} \left(\nabla_\nu v_\mu\right) - \Gamma^\nu{}_{\alpha\mu} \left(\nabla_\beta v_\nu\right) \\ &= \partial_\alpha \left(\partial_\beta v_\mu - \Gamma^\delta{}_{\beta\mu} v_\delta\right) - \Gamma^\nu{}_{\alpha\beta} \left(\nabla_\nu v_\mu\right) - \Gamma^\nu{}_{\alpha\mu} \left(\partial_\beta v_\nu - \Gamma^\delta{}_{\beta\nu} v_\delta\right) \\ &= -\left(\partial_\alpha \Gamma^\delta{}_{\beta\mu}\right) v_\delta + \Gamma^\nu{}_{\alpha\mu} \Gamma^\delta{}_{\beta\nu} v_\delta \\ &\quad + \partial_\alpha \partial_\beta v_\mu - \Gamma^\nu{}_{\alpha\beta} \left(\nabla_\nu v_\mu\right) - \Gamma^\nu{}_{\beta\mu} \partial_\alpha v_\nu - \Gamma^\nu{}_{\alpha\mu} \partial_\beta v_\nu \\ &= -\left(\partial_\alpha \Gamma^\delta{}_{\beta\mu}\right) v_\delta + \Gamma^\nu{}_{\alpha\mu} \Gamma^\delta{}_{\beta\nu} v_\delta + \ldots \end{aligned} \quad (243)$$

where the $\ldots$ terms are symmetric in $\alpha \leftrightarrow \beta$.

Taking the commutator projects out these symmetric terms, and we find,

$$[\nabla_\alpha, \nabla_\beta] v_\mu = \left(\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha\right) v_\mu = R_{\alpha\beta\mu}{}^\delta v_\delta \quad (244)$$

with,

$$R_{\alpha\beta\mu}{}^\delta = \partial_\beta \Gamma^\delta{}_{\alpha\mu} - \partial_\alpha \Gamma^\delta{}_{\beta\mu} + \Gamma^\nu{}_{\alpha\mu} \Gamma^\delta{}_{\beta\nu} - \Gamma^\nu{}_{\beta\mu} \Gamma^\delta{}_{\alpha\nu} \quad (245)$$

As we shall see, the Riemann tensor characterises the *curvature* of the geometry - ie. how it deviates from Euclidean/Minkowski space-time.

The commutator $[\nabla_\alpha, \nabla_\beta]$ on other tensors is also determined by Riemann.

Using the fact that $(v_\mu w^\mu)$ is a scalar, we find using the Liebnitz rule,

$$\begin{aligned} 0 = [\nabla_\alpha, \nabla_\beta](v_\mu w^\mu) &= v_\mu [\nabla_\alpha, \nabla_\beta] w^\mu + w^\mu [\nabla_\alpha, \nabla_\beta] v_\mu \\ &= v_\mu [\nabla_\alpha, \nabla_\beta] w^\mu + w^\mu R_{\alpha\beta\mu}{}^\nu v_\nu \end{aligned} \tag{246}$$

and as this is true for any v_μ and w^μ, this implies, for a vector field,

$$[\nabla_\alpha, \nabla_\beta] w^\mu = -R_{\alpha\beta\nu}{}^\mu w^\nu \tag{247}$$

Then repeating this logic for the scalar $(u^\mu v^\nu A_{\mu\nu})$ one concludes (Exercise);

$$[\nabla_\alpha, \nabla_\beta] A_{\mu\nu} = R_{\alpha\beta\mu}{}^\gamma A_{\gamma\nu} + R_{\alpha\beta\nu}{}^\gamma A_{\mu\gamma} \tag{248}$$

[*Comment:* Generally one finds for a tensor $T^{\mu_1 \ldots \mu_r}{}_{\nu_1 \ldots \nu_r}$ that,

$$\begin{aligned} [\nabla_\alpha, \nabla_\beta] T^{\mu_1 \ldots \mu_r}{}_{\nu_1 \ldots \nu_r} &= R_{\alpha\beta\nu_1}{}^\gamma T^{\mu_1 \ldots \mu_r}{}_{\gamma\nu_2 \ldots \nu_r} + R_{\alpha\beta\nu_2}{}^\gamma T^{\mu_1 \ldots \mu_r}{}_{\nu_1\gamma\nu_3 \ldots \nu_r} + \ldots \\ &\quad - R_{\alpha\beta\gamma}{}^{\mu_1} T^{\gamma\mu_2 \ldots \mu_r}{}_{\nu_1 \ldots \nu_r} - R_{\alpha\beta\gamma}{}^{\mu_2} T^{\mu_1\gamma\mu_2 \ldots \mu_r}{}_{\nu_1 \ldots \nu_r} \end{aligned}$$

]

7.2 Riemann in LIF coordinates:

Now recall that in LIF coordinates we have,

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + O(x^2) \ , \quad \Gamma^{\alpha}{}_{\mu\nu} = O(x) \tag{249}$$

Thus we note that at $x = 0$ we have $\Gamma^{\alpha}{}_{\mu\nu} = 0$, but we note that $\partial_\beta \Gamma^{\alpha}{}_{\mu\nu} \neq 0$.

Thus in LIF coordinates at $x = 0$ we have;

$$R_{\alpha\beta\mu}{}^{\delta}\big|_{x=0} \;=\; \partial_\beta \Gamma^{\delta}{}_{\alpha\mu} - \partial_\alpha \Gamma^{\delta}{}_{\beta\mu}\big|_{x=0} \tag{250}$$

Comment: For Minkowski or Euclidean space where the metric is constant and $\Gamma^{\mu}{}_{\alpha\beta} = 0$ everywhere then the Riemann tensor vanishes.

Generally we think of LIF coordinates as giving,

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + T_{\mu\nu\alpha\beta} x^\alpha x^\beta + O(x^3) \tag{251}$$

for some constants $T_{\mu\nu\alpha\beta} = T_{(\mu\nu)(\alpha\beta)}$. However, in fact one may find coordinates such that,

$$g_{\mu\nu}(x) = \eta_{\mu\nu} - \frac{1}{3}\, R_{\mu\alpha\nu\beta}\big|_{x=0}\, x^\alpha x^\beta + O(x^3) \tag{252}$$

which in fact fixes the coordinates entirely, up to the overall Lorentz transformation. Hence we see the Riemann tensor is the obstruction to the metric being simply Minkowski.

7.3 Symmetries of the Riemann tensor

Firstly, by definition it has the antisymmetry property; $\boxed{R_{\alpha\beta\mu}{}^{\nu} = -R_{\beta\alpha\mu}{}^{\nu}}$

Secondly, by direct calculation one can show; $\boxed{R_{[\alpha\beta\mu]}{}^{\nu} = 0}$

[Start with the fact;

$$R_{[\alpha\beta\mu]}{}^{\nu} v_\nu = \nabla_{[\alpha} \nabla_\beta v_{\mu]} \tag{253}$$

Then using,

$$\nabla_{[\beta} v_{\mu]} = \partial_{[\beta} v_{\mu]} - \Gamma^{\alpha}{}_{[\beta\mu]} v_\alpha = \partial_{[\beta} v_{\mu]} \tag{254}$$

one then finds,

$$\nabla_{[\alpha} \nabla_\beta v_{\mu]} = \nabla_{[\alpha} \partial_\beta v_{\mu]} = \partial_{[\alpha} \partial_\beta v_{\mu]} - \partial_\nu v_{[\mu} \Gamma^{\nu}{}_{\alpha\beta]} - \Gamma^{\nu}{}_{[\alpha\mu} \partial_{\beta]} v_\nu = 0 \tag{255}$$

and since this is true for any v we have, $R_{[\alpha\beta\mu]}{}^{\nu} = 0$.]

Thirdly, since ∇ is metric compatible; $\boxed{R_{\alpha\beta\mu\nu} = -R_{\alpha\beta\nu\mu}}$

[Since $\nabla_\alpha g_{\mu\nu} = 0$, then,

$$\begin{aligned} 0 = [\nabla_\alpha, \nabla_\beta] g_{\mu\nu} &= R_{\alpha\beta\mu}{}^{\gamma} g_{\gamma\nu} + R_{\alpha\beta\nu}{}^{\gamma} g_{\mu\gamma} \\ &= R_{\alpha\beta\mu\nu} + R_{\alpha\beta\nu\mu} \end{aligned} \tag{256}$$

as claimed.]

A consequence of the above symmetries is that; $\boxed{R_{\alpha\beta\mu\nu} = +R_{\mu\nu\alpha\beta}}$.

7.4 The Ricci tensor and scalar and the Einstein tensor

The natural contraction of the Riemann tensor gives the $(0,2)$ *Ricci tensor*;

$$R_{\mu\nu} = R_{\mu\alpha\nu}{}^{\alpha} \tag{257}$$

Due to the Riemann symmetry: $R_{\mu\alpha\nu\beta} = R_{\nu\beta\mu\alpha}$ we see the Ricci tensor is *symmetric*;

$$R_{\mu\nu} = R_{\nu\mu} \tag{258}$$

Taking a further trace we obtain the *Ricci scalar*;

$$R = R_{\mu}{}^{\mu} \tag{259}$$

Let us define the symmetric $(0,2)$ Einstein tensor in terms of the Ricci tensor and scalar;

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \tag{260}$$

This has precisely the same content as $R_{\mu\nu}$ (see Ex Sheet 7, Qu 1), but we have reorganised its trace a little. For a 4-dimensional spacetime,

$$R_{\mu\nu} \equiv G_{\mu\nu} - \frac{1}{2} g_{\mu\nu} G \tag{261}$$

where $G = G_{\mu\nu} g^{\mu\nu}$.

7.5 The Bianchi identity

The Bianchi identity for the Riemann tensor is;

$$\nabla_{[\mu} R_{\nu\alpha]\beta}{}^{\rho} = 0 \tag{262}$$

or equivalently;

$$\nabla_{\mu} R_{\nu\alpha\beta}{}^{\rho} + \nabla_{\nu} R_{\alpha\mu\beta}{}^{\rho} + \nabla_{\alpha} R_{\mu\nu\beta}{}^{\rho} = 0 \tag{263}$$

This is straightforwardly shown in LIF coordinates at a point $x = 0$;

$$\begin{aligned} \nabla_{\nu} R_{\alpha\beta\mu}{}^{\delta}|_{x=0} &= \partial_{\nu}\left(\partial_{\beta}\Gamma^{\delta}{}_{\alpha\mu} - \partial_{\alpha}\Gamma^{\delta}{}_{\beta\mu} + \Gamma^{\nu}{}_{\alpha\mu}\Gamma^{\delta}{}_{\beta\nu} - \Gamma^{\nu}{}_{\beta\mu}\Gamma^{\delta}{}_{\alpha\nu}\right)\Big|_{x=0} \\ &= \left(\partial_{\nu}\partial_{\beta}\Gamma^{\delta}{}_{\alpha\mu} - \partial_{\nu}\partial_{\alpha}\Gamma^{\delta}{}_{\beta\mu}\right)\Big|_{x=0} \end{aligned} \tag{264}$$

since $\Gamma^{\alpha}{}_{\mu\nu}|_{x=0} = 0$, and so,

$$\nabla_{[\nu} R_{\alpha\beta]\mu}{}^{\delta}|_{x=0} = \left(\partial_{[\nu}\partial_{\beta}\Gamma^{\delta}{}_{\alpha]\mu} - \partial_{[\nu}\partial_{\alpha}\Gamma^{\delta}{}_{\beta]\mu}\right)\Big|_{x=0} = 0 \tag{265}$$

since $\partial_{[\alpha}\partial_{\beta]} = 0$ as partial derivatives commute.

Contracting the Bianchi identity:

Contracting the above Bianchi identity as,

$$\nabla_{[\mu}R_{\nu\alpha]\beta}{}^{\mu} = 0 \tag{266}$$

yields;

$$\begin{aligned} 0 &= \nabla_\mu R_{\nu\alpha\beta}{}^{\mu} + \nabla_\nu R_{\alpha\mu\beta}{}^{\mu} + \nabla_\alpha R_{\mu\nu\beta}{}^{\mu} \\ &= \nabla_\mu R_{\nu\alpha\beta}{}^{\mu} + \nabla_\nu R_{\alpha\beta} - \nabla_\alpha R_{\nu\beta} \end{aligned} \tag{267}$$

Now a second contraction gives;

$$0 = \nabla_\mu R_{\nu\alpha}{}^{\nu\mu} + \nabla_\nu R_\alpha{}^{\nu} - \nabla_\alpha R_\nu{}^{\nu} = 2\nabla_\mu R_\alpha{}^{\mu} - \nabla_\alpha R_\nu{}^{\nu} \tag{268}$$

This identity;

$$\nabla_\mu R_\alpha{}^{\mu} - \frac{1}{2}\nabla_\alpha R = 0 \tag{269}$$

is physically of key importance.

The Einstein tensor $G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$ has the crucial property that it is *conserved*;

$$\begin{aligned} \nabla_\mu G^{\mu\nu} &= \nabla_\mu\left(R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R\right) \\ &= \nabla_\mu R^{\mu\nu} - \frac{1}{2}\nabla_\nu R = 0 \end{aligned} \tag{270}$$

from the contracted Bianchi identity.

7.6 Riemann tensor and parallel transport

The commutator $[\nabla_\mu, \nabla_\nu]$ can be thought of as arising when one parallel transports a vector around a closed curve.

Consider the square in the x, y plane, with corners at $(x, y) = (0,0), (0,\epsilon), (\epsilon,\epsilon), (\epsilon,0)$. Then consider the two curves,

$$\begin{aligned} \mathrm{C}_{(1)} : &\quad (0,0) \to (0,\epsilon) \to (\epsilon,\epsilon) \\ \mathrm{C}_{(2)} : &\quad (0,0) \to (\epsilon,0) \to (\epsilon,\epsilon) \end{aligned} \tag{271}$$

Consider a vector v^μ at $(0,0)$ which is parallel transported along the two curves $\mathrm{C}_{(1,2)}$ to yield two vectors $v^\mu_{(1,2)}$ at (ϵ,ϵ). The difference $v^\mu_{(2)} - v^\mu_{(1)}$ characterises the dependence of parallel transport on the path taken. Calculation (see Ex Sheet 7) shows;

$$v^\mu_{(2)} - v^\mu_{(1)} \quad = \quad \epsilon^2 \, R_{xy\nu}{}^{\mu} v^\nu \big|_{(0,0)} + O(\epsilon^3)$$

Hence we see that the Riemann curvature is the obstruction to parallel transport being path independent.

Comment: In Euclidean or Minkowski space-time where $R_{\mu\nu\alpha}{}^{\beta} = 0$, then parallel transport is path independent as we are familiar with.

8 The Einstein's equation

Recall that deforming Minkowski space to the Newtonian spacetime,

$$g_{\mu\nu} = \eta_{\mu\nu} - 2\epsilon\,\Phi(x)\delta_{\mu\nu} + O(\epsilon^2) \tag{272}$$

for coordinates $x^\mu = (t, x, y, z)$ yields Newton's law of gravity for slow moving objects in the Newtonian limit (governed by $\epsilon \to 0$), where the Newton gravity potential Φ obeys the usual,

$$\delta^{ij}\partial_i\partial_j\Phi = 4\pi G_N \rho \tag{273}$$

for mass density ρ. It also predicts gravitational red shift and light bending which are indeed observed. However, why should we pick this particular deformation of Minkowski? We will now discuss Einstein's equation that governs how the spacetime geometry depends on the matter present in it.

Recall that the stress energy tensor $T_{\mu\nu}$ gives a characterization of the matter in a spacetime. An important point is that it is conserved $\nabla_\mu T^{\mu\nu} = 0$. Consider the stress tensor for a perfect fluid;

$$T_{\mu\nu} = (\rho + P)\,u_\mu u_\nu + P g_{\mu\nu} \tag{274}$$

where ρ and P are energy density and pressure, and u^μ gives the local 4-velocity of the fluid, so $u^\mu u_\mu = -1$.

Consider Minkowski spacetime with $x^\mu = (t, x, y, z)$. Recall from Ex Sheet 5 that the Newtonian fluid equations are recovered in the limit of small 3-velocity $v^i \ll 1$ and $P \ll \rho$ (ie. the dominant contribution to the local energy comes from the local rest mass energy). Then the stress tensor to leading order is simply,

$$T_{tt} \simeq \rho\,, \quad T_{ti} \simeq 0\,, \quad T_{ij} \simeq 0 \tag{275}$$

Now any equation governing the geometry of spacetime must recover in the Newtonian case the equation $\delta^{ij}\partial_i\partial_j\Phi = 4\pi G_N \rho$ above.

Note that *two derivatives of the geometry are locally related to the energy density.* This strongly suggests we should try to equate local curvature with the stress tensor.

A first attempt at the Einstein equation:

We see that the stress tensor $T_{\mu\nu}$ provides a natural tensor to characterize the local matter content, giving just ρ in the Newtonian limit. In 1915 Einstein made a first attempt;

$$R_{\mu\nu} = \kappa T_{\mu\nu} \tag{276}$$

for some constant κ. This certainly has the correct flavour. Recall that the Ricci tensor involves two derivatives of the metric as we require.

However there is a serious problem with this equation. Note that since $T_{\mu\nu}$ is conserved we obtain,

$$\nabla^\mu R_{\mu\nu} = 0 \tag{277}$$

But by the Bianchi identity,

$$\nabla^\mu R_{\mu\nu} = \frac{1}{2}\nabla_\nu R \tag{278}$$

and hence this implies $\nabla_\nu R = 0$ and so R =constant. Taking the trace of the equation,

$$R = \kappa T \tag{279}$$

where $T = T_{\mu\nu}g^{\mu\nu}$. Thus we see this Einstein equation implies the trace of $T_{\mu\nu}$ =constant. In the Newtonian case above, $T \simeq \rho$, but this is far too restrictive.

Hence the key problem is reconciling the stress energy conservation with the Bianchi identity.

The Einstein equation:

Einstein realised that stress energy conservation was a crucial ingredient that must be consistent with the Einstein equation, rather than an additional restriction on it. The key observation is that the Einstein tensor;

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R \tag{280}$$

is conserved, so that,

$$\nabla^\mu G_{\mu\nu} = 0 \tag{281}$$

This is the unique conserved $(0, 2)$ tensor involving two derivatives of the metric. Hence we conclude the Einstein equation must take the form;

$$G_{\mu\nu} = \kappa \, T_{\mu\nu} \tag{282}$$

[*Comment:* It may seem mysterious that the Einstein curvature tensor obeys precisely a conservation equation. In fact the origin of the stress energy and Einstein tensor conservation is exactly the same, essentially due to the geometric nature of the theory, namely that tensor equations are coordinate invariant.]

Now consider the Newtonian spacetime. In Ex Sheet 7, Question 4, you computed that (to leading order in ϵ, which we then ignore);

$$\begin{aligned} G_{tt} &= 2\,\delta^{ab}\partial_a\partial_b\Phi \\ G_{ti} &= G_{it} = G_{ij} = 0 \end{aligned} \tag{283}$$

Further we have seen that in the Newtonian limit, the stress tensor on Minkowski spacetime takes the form,

$$\begin{aligned} T_{tt} &= \rho \\ T_{ti} &= T_{it} = T_{ij} = 0 \end{aligned} \tag{284}$$

To leading order in ϵ we will find the same result in the Newtonian spacetime (since to leading order it is simply Minkowski). To our relief, the ti and ij components of this Einstein equation are consistent, and the non-trivial tt component yields;

$$2\,\delta^{ab}\partial_a\partial_b\Phi = \kappa\rho \tag{285}$$

Recall in Newton's theory we expect $\delta^{ab}\partial_a\partial_b\Phi = 4\pi G_N\rho$. Hence we conclude that $\kappa = 8\pi G_N$, and hence we finally arrive that the celebrated Einstein equation (putting c back in);

$$G_{\mu\nu} = 8\pi G_N \, T_{\mu\nu} \tag{286}$$

which links the local matter to the local curvature.

Comment on units:

Choosing general rather than natural units so $c = 1$ we would obtain;

$$G_{\mu\nu} = \frac{8\pi G_N}{c^4} T_{\mu\nu} \tag{287}$$

Recall that,

$$[G_N] = \text{Mass}^{-1} \text{ Length}^3 \text{ Time}^{-2} \tag{288}$$

In units where $c = 1$ we derive time units from length (or vice versa).

We may go further and choose units such that $G_N = 1$ or $8\pi G_N = 1$. For instance, we may derive mass and time units from length. Such units are common in GR.

Indeed we may go only one step further (a step too far?) to also fix the quantum constant $\hbar = 1$, where we recall,

$$[\hbar] = \text{Mass Length}^2 \text{ Time}^{-1} \tag{289}$$

For instance we first derive mass and time from length using $G_N = 1$, and then choose length such that $h = 1$. These are called *Planck units*.

In SI units; $G_N = 6.7 \times 10^{-11} kg^{-1} m^3 s^{-2}$, $h = 6.6 \times 10^{-34} kgm^2 s^{-1}$

In $c = 1$ units; $G_N = 7.4 \times 10^{-28} kg^{-1} m$, $h = 2.2 \times 10^{-42} kgm$

In $c = 1$, $8\pi G_N = 1$ units; $\hbar = 6.8 \times 10^{-69} m^2$

Hence we see $1m = 2.0 \times 10^{-34}$ *Planck lengths.*

The Planck length units is the natural length/time units are the scales at which we expect quantum gravity to become important.

More remarks about Einstein's equation:

An important point is that the Ricci tensor is a contraction of Riemann, but does not contain all the information of Riemann. Hence the Einstein equation does not totally fix the space-time geometry.

An important example of this is the case of the *vacuum* Einstein equation, where $T_{\mu\nu} = 0$. Note that;

$$T_{\mu\nu} = 0 \quad \Longrightarrow \quad G_{\mu\nu} = 0 \quad \Longrightarrow \quad R_{\mu\nu} = 0 \tag{290}$$

Hence space-time in vacuum must the *Ricci flat*. However, that does not mean that Riemann must vanish - as we shall see, the Schwarzschild black hole is an example of this.

Furthermore to any solution of the Einstein equation one can add gravity waves - ripples in space-time - which encode the freedom in Riemann not constrained by the Ricci contraction.

It is also worth noting that these Einstein equations together with the matter can be thought of as a dynamical system, where one gives initial data - an initial 3-spatial geometry + its momentum, together with the matter and its momentum - and then evolves in a time direction to build up the full spacetime.

We have seen that Newtonian gravity arises from these Einstein equations. However, one can ask what other solutions they contain. We will spend the remainder of our time considering two crucial solutions - the cosmological solution (the FLRW cosmology) and the black hole space-time (the Schwarzschild metric).

9 The FLRW space-time and cosmology

The cosmological principle states;

There is no preferred point or direction in the universe (once we have smoothed out small scale features)

Geometrically we take this to require spatial *homogeneity* (no point is special) and *isotropy* (no direction is special). The *unique* homogeneous isotropic space-time geometry is the Friedmann-Lemaitre-Robertson-Walker space-time with metric;

$$ds^2 = -dt^2 + a(t)^2 h_{ij}(x) dx^i dx^j \tag{291}$$

where $x^\mu = (t, x, y, z)$, t is a time coordinate and $h_{ij}(x)$ is the metric on a homogeneous isotropic 3 dimensional spatial geometry. There are 3 possibilities; flat space, the round 3-sphere, and 3-hyperbolic space;

$$h_{ij}(x) dx^i dx^j = \begin{cases} dr^2 + r^2 d\Omega^2 \\ dr^2 + \sin^2 r d\Omega^2 \\ dr^2 + \sinh^2 r d\Omega^2 \end{cases} \tag{292}$$

with $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$ the round 2-sphere metric.

In fact these can be written in the unified manner;

$$h_{ij}(x) dx^i dx^j = \frac{d\rho^2}{1 - k\rho^2} + \rho^2 d\Omega^2 \tag{293}$$

where $k = +1$ for the sphere, $k = 0$ for flat space and $k = -1$ for hyperbolic space.

However, for simplicity, and since it appears to be the physically important case, we will focus on the flat case $k = 0$. Then,

$$h_{ij}(x) dx^i dx^j = dr^2 + r^2 d\Omega^2 = dx^2 + dy^2 + dz^2 \tag{294}$$

so,

$$ds^2 = -dt^2 + a(t)^2 \left(dx^2 + dy^2 + dz^2 \right) \tag{295}$$

The function $a(t)$ is called the *scale factor* and controls the expansion or contraction of the spatial geometry. When we say the universe is expanding, technically we mean $da/dt > 0$.

A measure of this expansion is the *Hubble constant* H defined as,

$$H \equiv \frac{1}{a}\frac{da}{dt} \tag{296}$$

which is not actually a constant.

Conformal time τ

Rather than using the time coordinate t - the proper time of the timeline curves above - the geometrically more natural coordinate is *conformal time* τ where,

$$ds^2 = a^2(\tau)\left(-d\tau^2 + dx^2 + dy^2 + dz^2\right) = a^2(\tau)\eta_{\mu\nu}dx^\mu dx^\nu \tag{297}$$

for $x^\mu = (\tau, x, y, z)$. Hence,

$$a^2(\tau)d\tau^2 = dt^2 \quad\implies\quad t = \int d\tau\, a(\tau) \tag{298}$$

and we see the metric is Minkowski up to an overall scaling (a *conformal factor*) given by the scale factor.

Note that in conformal time the Hubble constant is;

$$H = \frac{1}{a^2}\frac{da}{d\tau} \tag{299}$$

9.1 Geodesics of FLRW

We will now compute the geodesics of FLRW. Recall that an affinely parameterized geodesic, $x^\mu(\lambda)$, obeys the equation $v^\mu \nabla_\mu v^\nu = 0$ for $v^\mu = dx^\mu/d\lambda$, which can be written, $\frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu{}_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = 0$.

Recall also that the norm of the tangent, $|v|^2$, is conserved along the geodesic;

$$\begin{aligned}\frac{d}{d\lambda}\left(|v|^2\right) &= \frac{dx^\mu}{d\lambda}\frac{\partial}{\partial x^\mu}\left(g_{\alpha\beta}v^\alpha v^\beta\right) = v^\mu \nabla_\mu\left(g_{\alpha\beta}v^\alpha v^\beta\right) \\ &= 2v^\alpha v^\mu \nabla_\mu v_\alpha = 0 \end{aligned} \tag{300}$$

since $v^\mu \nabla_\mu v_\alpha = 0$ for the geodesic.

However, also recall that geodesics in affine parameterization extremize,

$$L = \int d\lambda \, g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \tag{301}$$

and in practice this is often the most convenient way to compute geodesics.

We will now use this variational method to compute the geodesics of FLRW. Please see the example sheet 8 Qu 2 for calculations with geodesics using the geodesic equation rather than the variational approach.

We write the geodesic curve in terms of an affine parameter λ as,

$$x^\mu(\lambda) = \left(\tau(\lambda), x(\lambda)^i\right) \ , \quad \frac{dx^\mu(\lambda)}{d\lambda} = \left(\tau'(\lambda), x'(\lambda)^i\right) \tag{302}$$

where $\tau'(\lambda) = d\tau/d\lambda$, so we have,

$$L = \int d\lambda \mathcal{L} = \int d\lambda \, a(\tau(\lambda))^2 \left(-\tau'(\lambda)^2 + \delta_{ij} x'(\lambda)^i x'(\lambda)^j\right) \tag{303}$$

The Euler-Lagrange equation for $x(\lambda)$ variations is;

$$\frac{d}{d\lambda} \left(\frac{\partial L}{\partial x'(\lambda)^i} \right) = \frac{\partial L}{\partial x(\lambda)^i} \tag{304}$$

giving,

$$\frac{d}{d\lambda} \left(2a^2(\tau(\lambda)) \delta_{ij} x'(\lambda)^j\right) = 0 \tag{305}$$

which implies

$$x'(\lambda)^i \quad = \quad \frac{v^i}{a^2(\tau(\lambda))} \tag{306}$$

where v^i are constants giving the spatial direction of the geodesic.

Anyhow we conclude that,

$$a^2(\tau(\lambda)) x'(\lambda)^i = \text{constant} \tag{307}$$

This is resulting from the fact that $\partial L/\partial x^i$ vanishes, and hence the metric is independent of coordinates x^i resulting in the spatial isometries generated by Killing vectors,

$$u^\mu_{(1)} = (0,1,0,0) \ , \quad u^\mu_{(2)} = (0,0,1,0) \ , \quad u^\mu_{(3)} = (0,0,0,1) \tag{308}$$

We see that as a result of this spatial translation invariance we recover the momentum conservation above, where we see the momentum of a particle in FLRW is $a^2(\tau)x'^i$.

The time variation gives;

$$\frac{d}{d\lambda}\left(\frac{\partial L}{\partial \tau'(\lambda)}\right) = \frac{\partial L}{\partial \tau(\lambda)} \tag{309}$$

which yields,

$$\frac{d}{d\lambda}\left(-2a^2(\tau(\lambda))\tau'(\lambda)\right) \ = \ 2a\frac{da(\tau)}{d\tau}\left(-\tau'(\lambda)^2 + \delta_{ij}x'(\lambda)^i x'(\lambda)^j\right) = \frac{2}{a}\frac{da(\tau)}{d\tau}\mathcal{L}$$

Timelike geodesics

We now consider a class of timeline geodesics, those where,

$$x(\lambda)^i = \text{constant} \quad \implies \quad x'(\lambda)^i = 0 \quad \implies \quad v^i = 0 \tag{310}$$

These are called *co-moving observers*. As we shall see later, since the metric is homogeneous and isotropic so is the matter, so these observers sit at rest with respect to the matter - the move with it, hence *co-moving*.

Hence the timeline geodesics (in conformal time coordinates) have tangent;

$$\frac{dx^\mu}{d\lambda} = \left(\frac{d\tau(\lambda)}{d\lambda}, 0, 0, 0\right) \tag{311}$$

Requiring λ to be the proper time, so that $|dx/d\lambda|^2 = -1$ then implies,

$$\frac{d\tau(\lambda)}{d\lambda} = \frac{1}{a(\tau(\lambda))} \tag{312}$$

Further note that this proper time λ of these observers is the *same* as the t coordinate in the form,

$$ds^2 = -dt^2 + a(t)^2 dx^i dx^i \tag{313}$$

which recall also obeyed,

$$\frac{d\tau(\lambda)}{dt} = \frac{1}{a} \tag{314}$$

Null geodesic

Note that the metric it *conformal* to Minkowski. In fact this generally means the null geodesics are the same as for Minkowski as we shall see.

We will consider the general null geodesic. We require the curve to be null so $\mathcal{L} = 0$. Then the τ variation (310) gives,

$$\frac{d}{d\lambda}\left(a^2(\tau(\lambda))\tau'(\lambda)\right) \;=\; 0 \tag{315}$$

and hence,

$$\tau'(\lambda) = \frac{\beta}{a^2} \tag{316}$$

for a constant β.

However, for $\mathcal{L} = 0$, then,

$$a^2\left(-\tau'^2 + \delta_{ij}x'^i x'^j\right) = 0 \tag{317}$$

and $x'^i = v^i/a^2$. Hence we see actually, $\beta^2 = \delta_{ij}v^i v^j$, so the tangent is,

$$\frac{dx^\mu}{d\lambda} = \left(\frac{\sqrt{\delta_{ij}v^i v^j}}{a^2}, \frac{v^i}{a^2}\right) \tag{318}$$

Again, one would integrate,

$$\frac{d\tau(\lambda)}{d\lambda} = \frac{\sqrt{\delta_{ij}v^i v^j}}{a^2(\tau)} \quad\implies\quad \lambda = \frac{1}{\sqrt{\delta_{ij}v^i v^j}}\int d\tau a^2(\tau) \tag{319}$$

to obtain the actual curve.

Note that the set of points the curve passes through is the same as a null geodesic in Minkowski spacetime - ie. a 45^o angle line! However the affine parameter λ evolves differently along the curve.

9.2 Cosmological redshift

Consider two comoving massive particles, on a source emitting pulses of radiation at a constant rate and the other receiving these pulses. The radiation pulses simply travel along 45^o lines in the conformal time coordinates (τ, x^i). The key point is that this remains true independent of the time of emission.

Suppose the emitter at (τ_e, x^i_e) emits signals every small proper time interval Δt_e. Corresponding to this proper time is the conformal time interval $\Delta\tau_e$;

$$(\Delta t_e)^2 = a(\tau_e)^2(\Delta\tau_e)^2 \tag{320}$$

An observer receives these signals at (τ_o, x^i_o) seeing a proper time interval Δt_o, and corresponding conformal time interval $\Delta\tau_o$;

$$(\Delta t_o)^2 = a(\tau_o)^2(\Delta\tau_o)^2 \tag{321}$$

Now the radiation then follows a null geodesic, a 45^o line in the (τ, x^i) coordinates between emitter and observer. This implies the crucial result that, independent of the emission time τ_e,

$$\Delta\tau_o = \Delta\tau_e \tag{322}$$

Hence we conclude the proper time intervals of emitter and receiver are related as;

$$\frac{\Delta t_o}{\Delta t_e} = \frac{a(\tau_o)}{a(\tau_e)} \tag{323}$$

Now in an expanding universe then $a(\tau_o) > a(\tau_e)$ and hence the observed time interval $\Delta t_o > \Delta t_e$, and hence the frequency is redshifted.

Thus observers in an expanding universe looking at signals emitted in their past see them redshifted. Let us expand the scale factor around the observers time as,

$$a(\tau) = a(\tau_o) + a(\tau_o)^2 H_o(\tau - \tau_o) + \dots \tag{324}$$

where $H_o > 0$ for an expanding universe and is the value of Hubble constant, H, at the observer time. (Recall $H = \frac{1}{a^2}\frac{da}{d\tau}$.)

Then write $\tau_0 = \tau_e + \Delta\tau$, and so we find,

$$\begin{aligned}
\frac{\Delta t_o}{\Delta t_e} &= \frac{a(\tau_o)}{a(\tau_e)} = \frac{a(\tau_o)}{a(\tau_o) + a(\tau_o)^2 H_o(\tau_e - \tau_o) + \dots} \\
&= \frac{1}{1 - a(\tau_o)H_o(\tau_o - \tau_e) + O((\tau_e - \tau_o)^2)} \\
&= 1 + a(\tau_o)H_o(\tau_o - \tau_e) + O((\tau_e - \tau_o)^2)
\end{aligned} \tag{325}$$

Now since null geodesics are 45^o lines in the(τ, x^i) coordinates, then,

$$|x_o - x_e| = \sqrt{\delta_{ij}(x_o^i - x_e^i)} = \tau_o - \tau_e \tag{326}$$

The proper distance, $D(\tau)$, between the observer and emitter in the spatial slice at constant τ is,

$$D(\tau) = a(\tau)|x_o - x_e| \tag{327}$$

Hence,

$$D_o \equiv D(\tau_0) = a(\tau_o)(\tau_o - \tau_e) \tag{328}$$

and we conclude that;

$$\frac{\Delta t_o}{\Delta t_e} = 1 + H_o\, D_o + O((\tau_e - \tau_o)^2) \tag{329}$$

We define the *Cosmological Redshift* Z as,

$$Z \equiv \frac{\Delta t_o}{\Delta t_e} - 1 \tag{330}$$

so $Z = 0$ implies no difference in clock speed, and $Z > 0$ means the observed signal is redshifted.

Then we find the *Hubble law*;

$$Z = H_o \, D_o \tag{331}$$

which states the redshift of any object is proportional to its distance from us.

Note: For an object in Minkowski space receding from us at velocity $v \ll 1$, then,

$$Z = \frac{\Delta t_o}{\Delta t_e} - 1 \simeq v \tag{332}$$

so for small redshift $Z \ll 1$, then Z gives the apparent recession velocity of the emitter. Thus the apparent velocity of recession of objects is proportional to their distance form us.

Hubble and latter observations found,

$$H_{now} \sim 70 (km/s)/Mpc \tag{333}$$

where $1 Mpc \sim 31 \times 10^{15} m$ or roughly 3 light years.

In fact this Hubble law is true for nearby objects only where $Z \ll 1$. In general one must use the full behaviour of $a(\tau)$ to compute the redshift - distance relation.

9.3 The Friedmann equation

In Question Sheet 8 you computed the Einstein tensor of the FLRW metric in conformal time coordinates to show;

$$\begin{aligned} G_{\tau\tau} &= 3\left(\frac{\dot{a}}{a}\right)^2 \\ G_{ij} &= \left(\frac{\dot{a}^2 - 2\,a\,\ddot{a}}{a^2}\right)\delta_{ij} \end{aligned} \tag{334}$$

where $\dot{a} = da/d\tau$.

We take the matter in the universe to be a perfect fluid so that,

$$T_{\mu\nu} = (\rho + P)\,u_\mu u_\nu + P g_{\mu\nu} \tag{335}$$

Then the *cosmological principle* states that the density ρ and pressure P should be homogeneous and isotropic - ie. not depend on the coordinates x^i at all. Further since there is no preferred direction then the 4-velocity $u^\mu = (\frac{1}{a}, 0, 0, 0)$ so that the 3-velocity vanishes and $u^\mu u_\mu = -1$.

Since, $u_\mu = (a, 0, 0, 0)$, this implies,

$$\begin{aligned} T_{\tau\tau} &= a^2(\tau)\rho(\tau) \\ T_{ij} &= a^2(\tau)P(\tau)\delta_{ij} \end{aligned} \tag{336}$$

Hence from the tt component of the Einstein equation,

$$G_{\mu\nu} = 8\pi G_N T_{\mu\nu} \tag{337}$$

we obtain the **Friedmann equation**;

$$H^2 = \left(\frac{1}{a^2}\frac{da}{d\tau}\right)^2 = \frac{\dot{a}^2}{a^4} \quad = \quad \frac{8\pi G_N}{3}\rho \tag{338}$$

In fact due to the contracted Bianchi identity we do not need to consider the ij component provided we ensure stress energy conservation. The fluid equations are given by $\nabla_\mu T^{\mu\nu} = 0$ and following the calculation you did in example sheet 8 one finds;

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + P) = 0 \tag{339}$$

For a perfect fluid the pressure P is determined in terms of the density ρ by the equation of state parameter w;

$$P = w\,\rho \tag{340}$$

The equation of state parameter takes the values;

$$w = \begin{cases} 0 & \text{cold matter} \\ +\frac{1}{3} & \text{radiation / hot matter} \\ -1 & \text{dark energy} \end{cases} \tag{341}$$

Note that for *dark energy* we have $T_{\mu\nu} = P(x) g_{\mu\nu}$. However (see Ex Sheet 7, Qu 2) this is simply a*cosmological constant*,

$$T_{\mu\nu} = \Lambda g_{\mu\nu} \tag{342}$$

where Λ is a constant. The Bianchi identities imply that $\partial_\mu P = 0$ so indeed $P = \Lambda$ is just a constant.

It is simple to derive from the fluid equations (see Ex Sheet 8, Qu 1) that;

$$\dot{\rho} + 3(1+w)\frac{\dot{a}}{a}\rho = 0 \quad \Longrightarrow \quad \rho(\tau) = \left(\frac{a_0}{a(\tau)}\right)^{3(1+w)} \tag{343}$$

Hence cold matter dilutes simply with the 3-volume ($a^3\rho$ =const), whereas hot matter/radiation dilutes with 4-volume ($a^4\rho$ =const). This latter result is because not only do the photons get diluted as the space expands, but they also become redshifted, which further decreases their energy density.

Now substituting these into the Friedmann equation one finds,

$$\frac{\dot{a}^2}{a^4} = \frac{8\pi G_N}{3}\left(\frac{a_0}{a(\tau)}\right)^{3(1+w)} \tag{344}$$

so that,

$$a(\tau)^{\frac{3}{2}(1+w)-2}\frac{da}{d\tau} = \sqrt{\frac{8\pi G_N a_0^{3(1+w)}}{3}} = \frac{1}{k} \tag{345}$$

and hence,

$$\tau - \tau_0 = k'\,a(\tau)^{\frac{3}{2}(1+w)-1} = k'\,a(\tau)^{\frac{1}{2}(1+3w)} \tag{346}$$

for constant of integration τ_0 and appropriate constant k'. So,

$$a(\tau) \quad = \quad k''(\tau - \tau_0)^{\frac{2}{(1+3w)}} \tag{347}$$

Hence for cold matter ($w = 0$) we have,

$$a \sim (\tau - \tau_0)^2 \tag{348}$$

and for hot matter/radiation ($w = 1/3$) we have,

$$a \sim (\tau - \tau_0) \tag{349}$$

Note: both hot and cold matter drive expansion for a flat universe. Also note that for hot matter, as one goes back in time the density and temperature increases.

We believe that early on in the universe, for approximately the first 10000 years the universe was hot. The time $\tau_0 = 0$ is the *Big Bang*!

For hot matter, $a^4\rho$ =constant, so if $a \sim (\tau - \tau_0)$ then,

$$\rho \sim (\tau - \tau_0)^{-1/4} \rightarrow +\infty \text{ as } \tau \rightarrow \tau_0 \tag{350}$$

However, note the proper time measured by a comoving observer after the big bang is *finite*!

$$t(\tau) = \int_{\tau_0}^{\tau} d\tau a(\tau) \simeq \int_{\tau_0}^{\tau} d\tau k(\tau - \tau_0) = \frac{k}{2}(\tau - \tau_0)^2 \tag{351}$$

(Interestingly the term 'Big Bang' was given by Fred Hoyle, who didn't believe in it!)

As the universe expands it cools, and at some point the matter becomes cold (around $Z \sim 1000$, a proper time ~ 100000 years after big bang), and the scale factor then goes as $a \sim \tau^2$. Note that cold matter dilutes more slowly than hot matter.

A decade ago a big surprise came when supernovae observations actually indicated that a new dark energy component seems to be dominating the expansion! For $w = -1$ one finds $H^2 \sim \Lambda =$ constant. So the expansion continues indefinitely - the cosmological constant *doesn't dilute*!

10 The Schwarschild spacetime

Already in 1918 Karl Schwarschild wrote down the space-time;

$$ds^2 = -\left(1 - \frac{R_S}{r}\right) dt^2 + \left(1 - \frac{R_S}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 \tag{352}$$

where $d\Omega^2 = d\theta^2 + \sin\theta^2 d\phi^2$ is the line element on the round 2-sphere. We will work with these Schwarschild coordinates $x^\mu = (t, r, \theta, \phi)$.

A calculation of the geometry of this yields;

$$R_{\mu\nu} = 0 \quad \Longrightarrow \quad G_{\mu\nu} = 0 \tag{353}$$

and hence this is a *vacuum* space-time - $T_{\mu\nu} = 0$.

It was later proven by Birkhoff (1923) that this is the *unique* vacuum space-time that is *spherically symmetric* - ie. it has the isometries of the 2-sphere.

An important point is that this space-time is *static*, meaning that $\partial_t g_{\mu\nu} = 0$ and hence,

$$v^\mu = (1, 0, 0, 0) \tag{354}$$

is a Killing vector generating the time translation isometry. Birkhoff's theorem means that spherical symmetry in vacuum implies a static metric.

Whilst Ricci vanishes, Riemann does not. An interesting geometric invariant is the *Kretchmann* scalar;

$$R_{\mu\nu\alpha\beta}R^{\mu\nu\alpha\beta} = \frac{12R_S^2}{r^6} \tag{355}$$

Since the Kretchmann scalar becomes infinite at $r = 0$, this is a physical singularity that cannot be removed by a coordinate transformation.

Something very funny happens at $r = R_S$ the *Schwarzschild radius.* For $r > R_S$ then t is a time coordinate and r is spatial. However for $r < R_S$ in fact t is spatial and r is timeline!

We may change radial coordinates;

$$r = \rho \left(1 + \frac{R_S}{4\rho}\right)^2 \quad (356)$$

to attain,

$$ds^2 = -\frac{\left(1 - \frac{R_S}{4\rho}\right)^2}{\left(1 + \frac{R_S}{4\rho}\right)^2} dt^2 + \left(1 + \frac{R_S}{4\rho}\right)^4 \left(d\rho^2 + \rho^2 d\Omega^2\right) \quad (357)$$

This are called *isotropic* coordinates. Expanding for $\rho \gg R_S$ we find,

$$ds^2 = -\left(1 - \frac{R_S}{\rho} + \ldots\right) dt^2 + \left(1 + \frac{R_S}{\rho} + \ldots\right) \left(d\rho^2 + \rho^2 d\Omega^2\right) \quad (358)$$

Changing from spherical coordinates (t, ρ, θ, ϕ) to Cartesian (t, x, y, z), so $\rho^2 = x^2 + y^2 + z^2$ we have,

$$\begin{aligned} ds^2 &= \left(\eta_{\mu\nu} + \frac{R_S}{\rho}\delta_{\mu\nu} + \ldots\right) dx^\mu dx^\nu \\ &= \left(\eta_{\mu\nu} - 2\Phi\delta_{\mu\nu} + \ldots\right) dx^\mu dx^\nu \end{aligned} \quad (359)$$

and hence *asymptotically* we may identify the space-time as having Newtonian form,

$$2\Phi = -\frac{R_S}{\rho} \quad (360)$$

and recall for a point mass M at $\rho = 0$ we expect,

$$\Phi = -\frac{G_N}{\rho} \quad \Longrightarrow \quad R_S = 2G_N M \quad (361)$$

Hence we conclude that the Schwarzcshild metric represents a spherically symmetric static space-time carrying a mass M given by R_S as above. It is often written as,

$$ds^2 = -\left(1 - \frac{2G_N M}{c^2 r}\right) dt^2 + \left(1 - \frac{2G_N M}{c^2 r}\right)^{-1} dr^2 + r^2 d\Omega^2 \quad (362)$$

Interpretation of the Schwarzschild solution:

Interpretation 1) - Stellar exterior, valid for $R > R_{star}$

We may view the Schwarschild metric as the *exterior* solution to any spherical matter distribution (static or otherwise!). For example, a star with mass M and sufficiently low density such that its radius is $> R_S$ will be Schwarschild in its exterior, with some static spheric metric in the interior which is *not vacuum* and hence not Schwarschild.

Example: For the sun, $M \simeq 2 \times 10^{30} kg$. The Schwarschild radius $R_S \simeq 3km$. Clearly the radius of the sun is much bigger being millions of km.

Example: Typical neutron stars has mass similar to the sun, but with radius of a few kilometres. Indeed they often rotate 1000 times per second!

Interpretation 2) - Black hole, valid except in the past

This is the end state of gravitational collapse - where a star burns its fuel and collapses below its Schwarzschild radius. This is a *black hole* space-time. The radius R_S is called the *horizon* of the black hole.

Now the space-time is taken for *all radii* R but since it is the end state of collapse, one should regard it as approximating the true physical space-time only after the point of collapse. In particular the far past of the Schwarzschild metric is not physical in this picture.

For this reason we should not allow any matter to be at a radius inside the horizon with an outward velocity.

10.1 Timelike geodesics of Schwarzschild

We write our space-time curve in terms of functions T, R, Θ, Φ in terms of an affine parameter λ.

$$x^\mu(\lambda) = (T(\lambda), R(\lambda), \Theta(\lambda), \Phi(\lambda)) \tag{363}$$

so the tangent is given by,

$$\frac{dx^\mu(\lambda)}{d\lambda} = \left(\dot{T}(\lambda), \dot{R}(\lambda), \dot{\Theta}(\lambda), \dot{\Phi}(\lambda)\right) \tag{364}$$

where $\dot{T} = dT/d\lambda$ and similarly for $\dot{R}, \dot{\Theta}, \dot{\Phi}$.

The define $\mathcal{L}$ the norm of the tangent for the curve. For an affine parameterization this will be constant; for a timeline geodesic $\mathcal{L} = -1$;

$$\mathcal{L} \equiv g_{\mu\nu} \frac{dx^\mu(\tau)}{d\tau} \frac{dx^\nu(\tau)}{d\tau} = -\left(1 - \frac{R_S}{R}\right) \dot{T}^2 + \left(1 - \frac{R_S}{R}\right)^{-1} \dot{R}^2 + R^2 \left(\dot{\Theta}^2 + \sin^2\Theta \dot{\Phi}^2\right)$$

Then we obtain the geodesics by varying this Lagrangian so that $L = \int d\lambda \mathcal{L}$ is extremized.

The Euler-Lagrange equations are;

$$\frac{d}{d\lambda}\left(\frac{\partial \mathcal{L}}{d\dot{T}}\right) = \frac{\partial \mathcal{L}}{dT} \tag{365}$$

and similarly for the other functions. In fact we need only vary T, Θ and Φ, and the remaining equation can be obtained from the condition $\mathcal{L} = -1$. In Ex Sheet 8 you explicitly performed these calculation to find,

$$\begin{aligned} \frac{\partial \mathcal{L}}{d\dot{T}} &= -2\left(1 - \frac{R_S}{R}\right)\dot{T} \\ \frac{\partial \mathcal{L}}{d\dot{\Theta}} &= 2R^2\dot{\Theta} \\ \frac{\partial \mathcal{L}}{d\dot{\Phi}} &= 2R^2 \sin^2\Theta\, \dot{\Phi} \end{aligned} \tag{366}$$

and,

$$\begin{aligned}
\frac{\partial \mathcal{L}}{dT} &= 0 \\
\frac{\partial \mathcal{L}}{d\Theta} &= 2R^2\dot{\Theta} \\
\frac{\partial \mathcal{L}}{d\Phi} &= 0
\end{aligned} \tag{367}$$

where the former and latter vanish due to the static symmetric and azimuthal symmetry of the metric. We expect these to yield energy and angular momentum conservation.

We then obtain the variational equations;

$$\begin{aligned}
T \text{ variation :} \qquad & \frac{d}{d\lambda}\left(\left(1-\frac{R_S}{R}\right)\dot{T}\right) = 0 \\
\Theta \text{ variation :} \qquad & \frac{d}{d\lambda}\left(R^2\dot{\Theta}\right) = R^2 \sin\Theta \cos\Theta \, \dot{\Phi}^2 \\
\Phi \text{ variation :} \qquad & \frac{d}{d\lambda}\left(R^2 \sin^2\Theta \, \dot{\Phi}\right) = 0
\end{aligned} \tag{368}$$

Consider motion in the equatorial plane $\Theta = \pi/2$. Hence $\dot{\Theta} = 0$ and $\sin\Theta = 1$, $\cos\Theta = 0$. We see this is consistent with the Θ equation of motion.

Note: By the spherical symmetry of the space-time we may always choose our coordinates so that the motion lies in this plane.

Now the first and last equations above simply imply conservation of energy E, and angular momentum L, where,

$$\begin{aligned}
E &= \frac{1}{2}\left(1-\frac{R_S}{R}\right)^2 \dot{T}^2 \\
L &= R^2\,\dot{\Phi}
\end{aligned} \tag{369}$$

for this equatorial plane motion. We must further supplement this with the equation $\mathcal{L} = \kappa$ where $\kappa = -1$ for a timeline geodesic (so $\lambda = \tau$ the proper time);

$$-1 = -\left(1-\frac{R_S}{R}\right)\dot{T}^2 + \left(1-\frac{R_S}{R}\right)^{-1}\dot{R}^2 + R^2\dot{\Phi}^2 \tag{370}$$

Now this may be written in terms of our conserved quantities;

$$-1 \quad = \quad -\left(1-\frac{R_S}{R}\right)^{-1} 2E + \left(1-\frac{R_S}{R}\right)^{-1} \dot{R}^2 + \frac{1}{R^2} L^2 \qquad (371)$$

and rearranging this gives;

$$E \quad = \quad \frac{1}{2}\dot{R}^2 + V(R) \qquad (372)$$

with,

$$V(R) \quad = \quad \frac{1}{2}\left(1-\frac{R_S}{R}\right)\left(1+\frac{L^2}{R^2}\right) \qquad (373)$$

Hence we see that motion is given by motion of a unit mass particle with energy E in the potential $V(R)$.

Properties of timeline geodesics:

The potential is then,

$$V(R) \quad = \quad \frac{1}{2} - \frac{R_S}{2}\frac{1}{R} + \frac{L^2}{2}\frac{1}{R^2} - \frac{L^2 R_S}{2}\frac{1}{R^3} \qquad (374)$$

which is the same as for the Newtonian theory except for the last term!

Newtonian motion occurs for large radius $R > R_S$ and low angular momentum $R > L$.

The extrema of this potential occur for;

$$\left(\frac{R}{L} - \frac{L}{R_S}\right)^2 = \frac{L^2}{R_S^2} - 3 \qquad (375)$$

Hence for low angular momentum $L < 3R_S$ there are no extrema. For $L > 3R_S$ there are two roots, $R_+ > R_-$, which you computed in Ex Sheet 8;

$$\frac{3}{2}R_S < R_- < 3R_S < R_+ \qquad (376)$$

where R_+ is a stable minimum giving a stable circular orbit, and R_- is unstable.

Hence the innermost stable circular orbit (ISCO) is $R = 3R_S = 6G_N M/c^2$.

Note that unlike the Newtonian theory, the potential is attractive near $R = 0$. Anyone geodesic passing R_- with $\dot{R} < 0$ will inevitably reach $R = 0$.

Perihelion advance:

The classic result of the Newtonian potential is that one obtains closed elliptical orbits. However, this is a very special property of the $1/R$ potential.

Let us expand the potential $V(R)$ about the stable circular orbit R_+ as,

$$V(R) = V(R_+) + \frac{1}{2}(R - R_+)^2 V''(R_+) + O\left((R - R_+)^3\right) \tag{377}$$

As you show in the Ex Sheet 8, one finds,

$$V''(R_+) = \frac{L^2}{R_+^4}\left(1 - 3\frac{R_S}{R_+}\right) \tag{378}$$

Hence for a nearly circular orbit (ie. $|R - R_+| \ll R_S$ for the duration of the orbit) we see the radius $R(\tau)$ performs simple harmonic motion about R_+ with a period, T_{radial}, given by,

$$T_{radial} = \frac{2\pi}{\sqrt{V''(R_+)}} \tag{379}$$

Recall from the definition of L we have,

$$\frac{d\Phi}{d\tau} = \frac{L}{R^2} \simeq \frac{L}{R_+^2} + O\left((R - R_+)\right) \tag{380}$$

and hence for a nearly circular motion to traverse 2π radians in Φ takes time $T_{angular}$, where,

$$T_{angular} \simeq \frac{2\pi R_+^2}{L} \tag{381}$$

Hence for nearly circular orbits,

$$\frac{T_{radial}}{T_{angular}} = \frac{1}{\sqrt{1 - 3\frac{R_S}{R_+}}} \simeq 1 + \frac{3}{2}\frac{R_S}{R_+} + \ldots \tag{382}$$

Thus in the Newtonian limit we recover closed orbits - $T_{radial} = T_{angular}$.

However, there is a small *advance* of the perihelion each orbit, with angle of advance;

$$\alpha \simeq 2\pi \left(\frac{T_{radial} - T_{angular}}{T_{radial}} \right) \text{radians/orbit} = \frac{\pi R_S}{R_+} \text{radians/orbit} \tag{383}$$

For Mercury about the sun; the solar mass $M_\odot = 2.0 \times 10^{30} kg$ and the orbital radius is $R \simeq 50 \times 10^9 m$. It orbits 4 times a year, so over 100 years the angle of advance;

$$\alpha|_{100yrs} \simeq 43 \, arcsec \tag{384}$$

Note; the moon in the sky subtends $\sim 2000 arcsec$.

10.2 Black holes

Recall the peculiar property that $r = 0$ is strongly attractive (unlike for Newtonian gravity), and further more r becomes a timeline coordinate for $r < R_S$.

Consider a massive particle with proper time τ, and position,

$$x^\mu(\tau) = (T(\tau), R(\tau), \Theta(\tau), \Phi(\tau)) \tag{385}$$

as above. Now we will not assume it is inertial (non-accelerated).

Consider the particle to be equipped with a rocket so that it may sit at constant R, Θ, Φ. Then its 4-velocity of the form,

$$v^\mu = \frac{dx^\mu}{d\tau} = \left(\frac{1}{\sqrt{1 - \frac{R_S}{R}}}, 0, 0, 0 \right) \tag{386}$$

so that $v^\mu v_\mu = -1$. We may compute the 4-acceleration of this particle,

$$a^\mu = v^\nu \nabla_\nu v^\mu \tag{387}$$

and find that,

$$a^\mu = \left(0, \frac{R_S}{2R^2}, 0, 0\right) \quad \Longrightarrow \quad |a| = \sqrt{a^\mu a_\mu} = \frac{R_S}{2R^2\sqrt{1-\frac{R_S}{R}}} \tag{388}$$

(see Ex Sheet 8 for the details of this calculation). Recall that this magnitude is a physical quantity and measures the usual acceleration experienced in the instantaneous LIF. The components of a^μ do not really measure anything physical themselves as they depend on the coordinates.

Hence we see that the force required to hold the particle at the fixed coordinate location becomes infinite at the horizon $r = R_S$. The horizon is the "radius of infinite attraction"!

The horizon is also a surface of infinite gravitational redshift. Consider the particle above emits a signal at a constant frequency, and we are very far from the black hole (ie. at $r \ll R_S$) observing it. Our proper time is approximately given by the coordinate t. However from the 4-velocity of the particle we see,

$$\frac{d\tau}{dt} = \frac{1}{\sqrt{1-\frac{R_S}{R}}} \tag{389}$$

An important point: Since the space-time is static the gravity redshift is just given by the ration of our proper time t to the particles.
Hence the frequency we observe ω_o is related to that emitted ω_e as,

$$\frac{\omega_o}{\omega_e} = \sqrt{1-\frac{R_S}{R}} \tag{390}$$

Thus for a particle approaching the horizon this redshift becomes infinite.

Falling in to a black hole:
Now consider the particle to have a general (accelerated) timeline trajectory so;

$$-1 = -\left(1-\frac{R_S}{R}\right)\dot{T}^2 + \left(1-\frac{R_S}{R}\right)^{-1}\dot{R}^2 + R^2\left(\dot{\Theta}^2 + \sin^2\Theta\dot{\Phi}^2\right) \tag{391}$$

Note that since r is time and t is spatial inside the horizon $1 - \frac{R_S}{R} < 0$. Thus for $R < R_S$ we may rearrange to;

$$+\left(\frac{R_S}{R} - 1\right)^{-1} \dot{R}^2 = 1 + \left(\frac{R_S}{R} - 1\right) \dot{T}^2 + R^2 \left(\dot{\Theta}^2 + \sin^2 \Theta \dot{\Phi}^2\right) \geq 1$$

and hence inside the horizon we have,

$$\dot{R}^2 \geq \frac{R_S}{R} - 1 \tag{392}$$

Now consider a particle that has just entered the horizon, so that $R < R_S$ and $\dot{R} < 0$. Thus R is decreasing, but for $R < R_S$ we cannot have $\dot{R} = 0$ and hence R will continue to decrease! This shows its *timeline* nature - you cannot reverse time.

Suppose this (unfortunate) particle falls in at time τ_i. Then,

$$\begin{aligned} -\sqrt{\frac{R_S}{R} - 1} \geq \dot{R} \qquad &\Longrightarrow \qquad d\tau \geq -\sqrt{\frac{R}{R_S - R}} dR \\ &\Longrightarrow \qquad \tau(R) - \tau_i \geq \int_{R_S}^{R} \sqrt{\frac{R}{R_S - R}} dR \\ &\Longrightarrow \qquad \tau(R) - \tau_i \leq \int_{R}^{R_S} \sqrt{\frac{R}{R_S - R}} dR \end{aligned} \tag{393}$$

Hence we see the time to reach $R = 0$ after falling in, $\Delta\tau = \tau(0) - \tau_i$ is bounded *from above* as;

$$\Delta\tau \leq \int_0^{R_S} \sqrt{\frac{R}{R_S - R}} dR = \frac{\pi R_S}{2} \tag{394}$$

Putting in the 'c' explicitly;

$$\Delta\tau \leq \frac{\pi R_S}{2c} \tag{395}$$

This is approximately the time it takes light to travel a distance R_S.

Example: For a solar mass black hole, $R_S \sim 3km$ which implies, $\Delta\tau \sim 10^{-5}s$.

Example: For a super massive black hole, $M \sim 10^9 M_\odot$ so $R_S \sim 10^{12} m$ which implies, $\Delta\tau \sim 10^4 s$.

The horizon

Comment: To the particle falling into a black hole *nothing* special happens when they pass the horizon - this is simply because for them, locally space-time is Minkowski!

The metric looks *singular* at the horizon with $g_{tt} \to 0$ and $g_{rr} \to \infty$. However this is a *coordinate singularity*, and is analogous to the coordinate singularity that occurs at the origin of polar coordinates;

$$ds^2_{Euc} = dr^2 + r^2 d\Omega^2 \tag{396}$$

where the determinant of the metric vanishes at $r = 0$.

Just as for Euclidean space, we know there is no problem with $r = 0$ since we may go to Cartesian coordinates where the metric is perfectly regular there, the same is true for the horizon.

We define an 'ingoing' time coordinate v as,

$$v = t + r^* = t + r + R_S \ln \left| \frac{r}{R_S} - 1 \right| \tag{397}$$

where we note the *tortoise* radial coordinate r^* obeys;

$$\frac{dr^*}{dr} = \frac{1}{1 - \frac{R_S}{r}} \tag{398}$$

In the coordinates (v, r, θ, ϕ) the metric takes the **ingoing Eddington-Finklestein** form;

$$ds^2 = -\left(1 - \frac{R_S}{r}\right) dv^2 + 2dv\, dr + r^2 d\Omega^2 \tag{399}$$

which is perfectly fine at $r = R_S$. In fact one finds the eigenvalues,

$$\{-1, 1, R_S^2, R_S^2 \sin^2\theta\} \tag{400}$$

there. These coordinates cover all the physically interesting region of the black hole - ie. the bit after the collapse.

The crucial point is that in these coordinates, nothing particularly dramatic happens at the horizon.

The point $r = 0$ is, however, a physical singularity - a *mysterious* point of infinite curvature.