Tensor Calculus and General Relativity (2022-2023)

Christopher Ford

Contents

1 Special Relativity

Special Relativity is developed with emphasis on aspects useful in General Relativity.

1.1 Inertial Frames

Inertial frames are systems of coordinates where the laws of Physics take a simple form

Although this is rather informal, we will develop it into a precise definition. An inertial frame K has coordinates (t, x, y, z) where t is time and x, y, z are position coordinates. Newton's First Law applies

A particle moves with constant velocity until it is disturbed by an external force

Inertial frames are not unique: consider K' with coordinates (t', x', y', z')

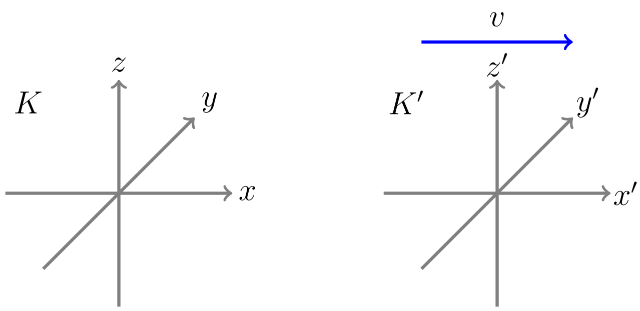
$$t' = t + A, \quad x' = x + B, \quad y' = y + c, \quad , z' = z + D,$$

where A, B, C, D are constants. This is just a translation (or redefinition of the origin) of the original coordinates. If K is inertial so is K'. The same idea works for rotations

$$t' = t, \quad z' = z, \quad x' = x\cos\theta + y\sin\theta, \quad y' = -x\sin\theta + y\cos\theta.$$

(t', x', y', z') is inertial if K is.

Now we consider the idea of a *boost.* Suppose K is inertial, and K' is uniformly translating with respect to K



Here K' is moving with a constant speed v in the x-direction. The *Galilean Transformation* representing this boost is

$$x' = x - vt, \quad y' = y, \quad z' = z. \tag{1}$$

Note that K' is *not* an inertial frame if K is. This is because the transformation does not preserve the speed of light.

The Special Principle of Relativity states

The laws of physics are the same in all inertial frames

An important law of physics is that the speed of light, c, is constant[1]

$$c = 299792458 \text{ metres per second.} \tag{2}$$

Note that the Galilean transformation preserves distances. The distance between two points P_1 and P_2 with respective coordinates (x_1, y_1, z_1) and (x_2, y_2, z_2) is given through the Pythagoras formula

$$d(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}.$$

Under a Galilean transformation $x_1 - x_2 = x_1' - x_2'$, $y_1 - y_2 = y_1' - y_2'$, $z_1 - z_2 = z_1' - z_2'$ so that the same distance is obtained on using the coordinates of K or K'. The speed of light is not preserved by the Galilean transformation.

We will now consider the *Lorentz transformation* which will be used to represent boosts. This transformation preserves the speed of light but does not preserve lengths (as we shall see lengths depend on which inertial frame is used). In addition the Lorentz transformation involves a redefinition of time whereas the Galilean transformation treats time as universal. The Lorentz transformation for a boost in the positive x direction is

$$x' = \gamma(x - vt), \quad t' = \gamma\left(t - \frac{vx}{c^2}\right), \quad y' = y, \quad z' = z, \tag{3}$$

where

$$\gamma = \left(1 - \frac{v^2}{c^2}\right)^{-1/2}.$$

For small v ($v^2 \ll c^2$)

$$\gamma \approx 1 + \frac{v^2}{2c^2} + \ldots$$

and the Galilean transformation is a good approximation to the Lorentz transformation.

To check that the Lorentz transformation preserves the speed of light consider

$$\frac{dx'}{dt'} = \frac{\gamma(dx - vdt)}{\gamma\left(dt - \frac{v}{c^2}dx\right)} = \frac{\dot{x} - v}{\left(1 - \frac{v\dot{x}}{c^2}\right)},$$

where the dot denotes the derivative with respect to t. Similarly,

$$\frac{dy'}{dt'} = \frac{\dot{y}}{\gamma\left(1 - \frac{v\dot{x}}{c^2}\right)}, \quad \frac{dz'}{dt'} = \frac{\dot{z}}{\gamma\left(1 - \frac{v\dot{x}}{c^2}\right)}.$$

[1] This now defines the metre.

Accordingly,

$$\left(\frac{dx'}{dt'}\right)^2+\left(\frac{dy'}{dt'}\right)^2+\left(\frac{dz'}{dt'}\right)^2=\frac{1}{\left(1-\frac{v\dot{x}}{c^2}\right)^2}\left[(\dot{x}-v)^2+\frac{1}{\gamma^2}(\dot{y}^2+\dot{z}^2)\right].$$

The square bracket can be rewritten as $[...]=c^2(1-v\dot{x}/c^2)^2+\gamma^{-2}(\dot{x}^2+\dot{y}^2+\dot{z}^2-c^2)$. Therefore

$$\left(\frac{dx'}{dt'}\right)^2+\left(\frac{dy'}{dt'}\right)^2+\left(\frac{dz'}{dt'}\right)^2=c^2,$$

if $\dot{x}^2+\dot{y}^2+\dot{z}^2=c^2$. This calculation is rather clumsy. We will later give a more general formulation of Lorentz transformations where the constancy of the speed of light is implicit.

It is straightforward to invert the Lorentz transformation

$$x=\gamma(x'+vt'),\quad t=\gamma\left(t'+\frac{vx'}{c^2}\right),\quad y=y',\quad z=z'.$$

This is the same as the standard boost formula with primed and unprimed coordinates exchanged and the sign of v reversed.

We have seen that the Lorentz transformation preserves the speed of light. Is the transformation unique? Consider the transformation

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

where λ is a constant. This is a Lorentz boost combined with a scaling of all four coordinates by λ. Clearly this transformation preserves the speed of light even if $\lambda\neq 1$. We will assume $\lambda=1$ - this is because the laws of physics are not scale invariant.

Note that the Lorentz transformation is linear. While there are some non-linear transformations that preserve the speed of light[2] these do not preserve Newton's first law.

Time and the notion of length are not invariant under Lorentz boosts. Two remarkable consequences of this are *length contraction* and *time-dilation* :

Length Contraction: a moving object becomes smaller!

Time Dilation: Moving clocks run slow!

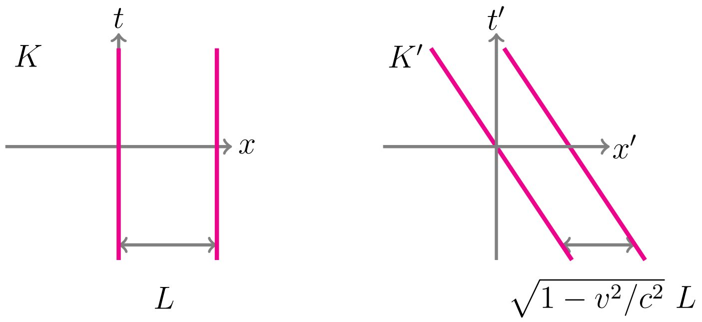
When we move on to General Relativity time dilation is more important than length contraction. Let us consider the two phenomena in turn.

1.2 Length Contraction

Consider a ruler of length L which is stationary in frame K; take the end points of the ruler to be at $x=0$ and $x=L$. The *worldline* of the two ends can be represented as vertical lines

[2]Special conformal transformations.

in the xt plane. In the frame K' the ruler is moving in the negative x' direction with speed v. Plotting the worldline of the ends in the $x't'$ frame one can see that the length of the moving ruler is $\sqrt{1-v^2/c^2}\ L$



For the end at $x = 0$

$$t' = \gamma\left(t - \frac{vx}{c^2}\right) = \gamma t, \quad x' = \gamma(x - vt) = -\gamma vt = -vt'$$

For the end at $x = L$

$$t' = \gamma\left(t - \frac{vx}{c^2}\right) = \gamma\left(t - \frac{vL}{c^2}\right) \quad \text{or} \quad \gamma t = \left(t' + \frac{\gamma vL}{c^2}\right),$$

and

$$x' = \gamma(x - vt) = \gamma(L - vt) = \gamma L - v\gamma t = \gamma L - v\left(t' + \frac{\gamma vL}{c^2}\right)$$

$$= -vt' + \gamma\left(1 - \frac{v^2}{c^2}\right) L = -vt' + \sqrt{1 - \frac{v^2}{c^2}}\ L.$$

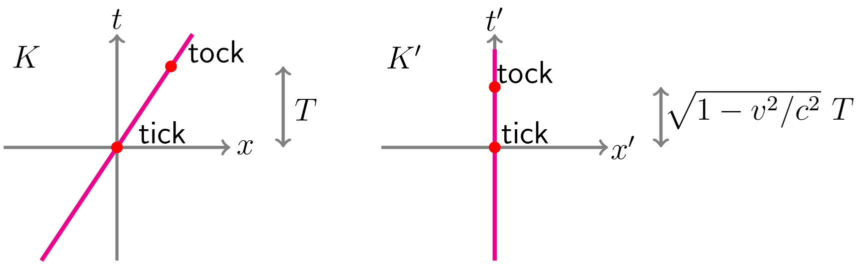
1.3 Time Dilation

Consider an object (a clock) moving at speed v in the positive x direction. The position of the clock is $x = vt$ which is represented by a line of slope $1/v$ in the xt plane. In the frame K' the object is stationary and is represented by a vertical line in the $x't'$ plane. The time interval between a 'tick' and 'tock' is greater in K (where the clock is moving) than in K' (where the clock is stationary); see the diagram below.

Here

$$t' = \gamma\left(t - \frac{vx}{c^2}\right) = \gamma\left(t - \frac{v^2 t}{c^2}\right) = \sqrt{1 - \frac{v^2}{c^2}}t,$$

is the time in a frame where the object (or clock) is stationary. This can be interpreted as



the time, τ, measured by the clock. This is called the *proper time* of the clock

$$\tau = \sqrt{1 - \frac{v^2}{c^2}}\ t.$$

Here we assumed that the clock had constant velocity in K. Now assume that for non-uniform motion the formula can be applied for a small time-interval, or in differential form

$$d\tau = \sqrt{1 - \frac{v^2}{c^2}}\ dt.$$

The proper time of a non-uniformly moving clock can be expressed as the integral

$$\tau_2 - \tau_1 = \int_{t_1}^{t_2} \sqrt{1 - \frac{v^2}{c^2}}\ dt. \tag{4}$$

One can avoid the awkward square root by working with $d\tau^2$ instead of $d\tau$

$$d\tau^2 = \left(1 - \frac{v^2}{c^2}\right) dt^2.$$

As

$$v^2 = \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2,$$

it follows that

$$c^2 d\tau^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2. \tag{5}$$

This is Lorentz invariant in that

$$c^2 d\tau^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = c^2 dt'^2 - dx'^2 - dy'^2 - dz'^2.$$

One can verify this using the explicit formula for the Lorentz transformation

$$dt' = \gamma\left(dt - \frac{v dx}{c^2}\right), \quad dx' = \gamma(dx - v dt), \quad dy' = dy, \quad dz' = dz.$$

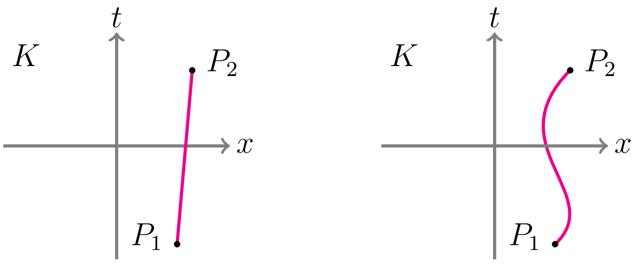
Note that if $v^2 = c^2$, $d\tau = 0$. That is, time is frozen for a clock moving at the speed of light. We can use the proper time formula to give a precise definition of an inertial frame:

An inertial frame K is a system of coordinates (t, x, y, z) where
$c^2 d\tau^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2$ describes the time of any moving clock

$d\tau^2$ is Lorentz invariant and also invariant under translations and spatial rotations. $c^2 d\tau^2$ is *not* invariant under a rescaling of the coordinates $t' = \lambda t$, $x' = \lambda x$, $y' = \lambda y$, $z' = \lambda z$.

One can also reformulate Newton's First Law as *The Principle of Maximal Ageing*.

Suppose a particle starts at P_1 and is later at P_2.[3] According to Newton's First Law the worldline should be a straight line joining P_1 and P_2. This trajectory maximises $\tau_2 - \tau_1$.



In the diagram above, the particle ages less for the non-uniform motion shown on the right.

The Principle of Maximal Ageing applies to the famous twin paradox. There are two twins. Twin 1 remains on Earth while twin 2 embarks on a high speed journey through space (v near c) and returns to Earth to reunite with twin 1. Twin 1 is much older than twin 2 when they meet.

1.4 Index Notation

We now introduce index notation which is useful in Special and General Relativity. Instead of using t, x, y, z define

$$x^0 = ct, \quad x^1 = x, \quad x^2 = y, \quad x^3 = z.$$

All four coordinates have the dimensions of length. The proper time formula (5) is

$$c^2 d\tau^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2.$$

We can absorb c into the definition of proper time through

$$s = c\tau$$

which also has the dimensions of length, so that

$$ds^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2.$$

[3] P_1 and P_2 are points in K. They are sometimes called *events* rather than points.

Alternatively,

$$ds^2 = \sum_{\mu=0}^{3}\sum_{\nu=0}^{3} \eta_{\mu\nu} dx^\mu dx^\nu,$$

where the *metric* $\eta_{\mu\nu}$ is an array of 16 numbers defined through

$$\eta_{00} = 1, \quad \eta_{11} = \eta_{22} = \eta_{33} = -1, \quad \eta_{\mu\nu} = 0 \text{ if } \mu \neq \nu.$$

$\eta_{\mu\nu}$ can be viewed as the entries of a 4×4 matrix

$$\eta = \text{diag}(1, -1, -1, -1) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Lorentz transformations can also be formulated in the same way:

$$(x^\mu)' = \sum_{\nu=0}^{3} \Lambda^{\mu'}_{\nu} x^\nu,$$

where $\Lambda^{\mu'}_{\nu}$ can be viewed as the entries of the 4×4 matrix Λ. For the standard Lorentz boost

$$\Lambda = \begin{pmatrix} \gamma & -\gamma v/c & 0 & 0 \\ -\gamma v/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The index notation can be streamlined through *Einstein's summation convention*. Here it is understood that repeated upper and lower indices are summed over without explicitly writing the summation signs. Thus we can can write

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu, \tag{6}$$

and

$$x^{\mu'} = \Lambda^{\mu'}_{\nu} x^\nu, \tag{7}$$

without the clumsy summation signs. In (7) the notation is further streamlined by writing $x^{\mu'}$ instead of $(x^\mu)'$. This convention of placing the primes on the indices is also useful in General Relativity.

1.5 Lorentz Group

The Lorentz boost discussed above is a specific Lorentz transformation. Now consider all linear transformations which preserve the form of ds^2. This includes all boosts, rotations and combinations thereof. Consider a linear transformation

$$x^{\mu'} = \Lambda^{\mu'}_{\nu} x^{\nu},$$

and require

$$ds^2 = \eta_{\mu'\nu'} dx^{\mu'} dx^{\nu'} = \eta_{\mu\nu} dx^{\mu} dx^{\nu},$$

which gives

$$\eta_{\mu'\nu'} \Lambda^{\mu'}_{\alpha} dx^{\alpha} \Lambda^{\nu'}_{\beta} dx^{\beta} = \eta_{\mu\nu} dx^{\mu} dx^{\nu}.$$

This is satisfied if

$$\Lambda^{\mu'}_{\alpha} \eta_{\mu'\nu'} \Lambda^{\nu'}_{\beta} = \eta_{\alpha\beta}.$$

This unwieldy equation is much simpler in matrix language

$$\Lambda^T \eta \Lambda = \eta, \tag{8}$$

where T is the transpose. The *Lorentz Group* is the set of 4×4 matrices which satisfy (8). This includes all boosts and spatial rotations. Consider

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \mathsf{R} & \\ 0 & & & \end{pmatrix}, \quad \eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & -\mathsf{I} & \\ 0 & & & \end{pmatrix},$$

where R is a 3×3 matrix and I is the 3×3 identity matrix. This gives

$$\Lambda^T \eta \Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & -\mathsf{R}^T\mathsf{R} & \\ 0 & & & \end{pmatrix} = \eta$$

so that $\mathsf{R}^T\mathsf{R} = \mathsf{I}$. That is R is an orthogonal matrix (or $\mathsf{R}^T = \mathsf{R}^{-1}$ which is a property of rotation matrices). For any member of the Lorentz group $\det \Lambda = \pm 1$. Standard boosts and rotations have $\det \Lambda = 1$ and $\Lambda^{0'}_{0} \geq 1$. Lorentz transformations with unit determinant are called *proper*. Improper transformations include parity $P = \text{diag}(1, -1, -1, -1)$ and time-reversal $T = \text{diag}(-1, 1, 1, 1)$. Transformations where $\Lambda^{0'}_{0} \geq 1$ are called *orthochronous*. $PT = \text{diag}(-1, -1, -1, -1)$ is proper but not orthochronous.

Now include translations in addition to (general) Lorentz transformations.

$$x^{\mu'} = \Lambda^{\mu'}_{\nu} x^{\nu} + a^{\mu'},$$

where the $a^{\mu'}$ are constants ($\mu' = 0, 1, 2, 3$). These *Poincaré transformations* define the *Poincaré Group*.

Claim: any transformation from one inertial frame to another is a Poincaré transformation.

1.6 Four-vectors

An *event* x can be identified with four coordinates x^{μ} ($\mu = 0, 1, 2, 3$). These coordinates change under a Lorentz transformation $x^{\mu'} = \Lambda^{\mu'}_{\nu} x^{\nu}$ with $\Lambda^{T} \eta \Lambda = \eta$.

A four-vector v^{μ} is four numbers v^{μ} ($\mu = 0, 1, 2, 3$) which transform in the same way as the coordinates under a Lorentz transformation

$$v^{\mu'} = \Lambda^{\mu'}_{\nu} v^{\nu} \quad \textit{where} \quad x^{\mu'} = \Lambda^{\mu'}_{\nu} x^{\nu}.$$

We will sometimes write $v = (v^0, v^1, v^2, v^3)$ much as $x = (x^0, x^1, x^2, x^3)$. Clearly x with components x^{μ} is trivially a four-vector.

An important example of a four-vector is the *four-velocity* u defined through

$$u^{\mu} = \frac{dx^{\mu}}{d\tau}. \tag{9}$$

where τ is the proper time of the particle. Now $d\tau$ is Lorentz invariant and since $dx^{\mu'} = \Lambda^{\mu'}_{\nu} dx^{\nu}$ it follows that $u^{\mu'} = \Lambda^{\mu'}_{\nu} u^{\nu}$. How are the components of the four-velocity related to the usual velocity components? In fact

$$u = \gamma(c, \mathbf{v}),$$

where $\gamma = (1 - v^2/c^2)^{-1/2}$ with v being the speed of the particle (here γ is not a boost parameter). Here $\mathbf{v} = (\dot{x}, \dot{y}, \dot{z})$ is the usual velocity vector. For example,

$$u^0 = \frac{dx^0}{d\tau} = c\frac{dt}{d\tau} = \frac{c}{\sqrt{1 - \frac{v^2}{c^2}}} = c\gamma.$$

The four-acceleration, a, is defined through

$$a^{\mu} = \frac{du^{\mu}}{d\tau} = \frac{d^2 x^{\mu}}{d\tau^2}. \tag{10}$$

Newton's first law is the statement that the four-acceleration of a free particle is zero ($a^\mu = 0$).

The four-momentum, p, of a particle has the components

$$p^\mu = mu^\mu,$$

where m is the mass of the particle. Using the explicit formula for u

$$p = \left(\frac{mc}{\sqrt{1-\frac{v^2}{c^2}}}, \frac{m\mathbf{v}}{\sqrt{1-\frac{v^2}{c^2}}} \right).$$

We can identify the components of p with the energy and linear momenta of the particle

$$p = \left(\frac{E}{c},\ \mathbf{p} \right),$$

so that

$$E = \frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}}, \quad \mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1-\frac{v^2}{c^2}}}.$$

If $v = 0$, $E = mc^2$.

1.7 Inner Product

An inner product of two four-vectors v and w is defined through

$$v \cdot w = \eta_{\mu\nu} v^\mu w^\nu = v^0 w^0 - v^1 w^1 - v^2 w^2 - v^3 w^3. \tag{11}$$

This is Lorentz invariant.

Note that in contrast to the dot product in three dimensions, $v \cdot v$ can be negative or zero.

If $v \cdot v > 0$, v is called *time-like*

If $v \cdot v = 0$ and $v \neq 0$, v is called *null*

If $v \cdot v < 0$, v is called *space-like*

For example $(1,0,0,0)$ is time-like, $(1,1,0,0)$ is null and $(0,1,0,0)$ is space-like.

The four-velocity $u = \gamma(c, \mathbf{v})$ is time-like since $u \cdot u = \gamma^2(c^2 - v^2) = c^2$. The four-momentum is time-like for a massive particle.

What about a massless particle such as a photon? For a massive particle $p \cdot p = m^2c^2$ so the four-momentum of a massless particle should be null. If $p = (E/c, \mathbf{p})$ is null then $E = c|\mathbf{p}|$ (assuming E is positive). The four-momentum of a photon of energy E moving in the positive x^1 direction is

$$p = \left(\frac{E}{c}, \frac{E}{c}, 0, 0 \right).$$

Note that the four-acceleration is orthogonal to the four-velocity, i.e. $a \cdot u = 0$. This follows by differentiating $u \cdot u = c^2$ with respect to proper time.

The total four-momentum of an isolated system is conserved

This result does not survive the transition to General Relativity. Instead the conservation of four-momentum can be formulated locally. This local form can be extended to General Relativity.

1.8 Vector Calculus in Special Relativity

A scalar field, ϕ or $\phi(x)$, is a function of the coordinates x^μ. A vector field, $v(x)$, with components $v^\mu(x)$ which are functions of the coordinates that transform like x^μ

$$v^{\mu'} = \Lambda^{\mu'}_{\nu} v^\nu \quad \text{where} \quad x^{\mu'} = \Lambda^{\mu'}_{\nu} x^\nu.$$

In three dimensions the gradient is the vector operator

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z},$$

or in index notation, the gradient of ϕ has components $\partial_i \phi$ where

$$\partial_i = \frac{\partial}{\partial x^i} \qquad (i = 1, 2, 3).$$

In Special Relativity a four dimensional version of the gradient can be defined through

$$\partial_\mu = \frac{\partial}{\partial x^\mu} \qquad (\mu = 0, 1, 2, 3).$$

Consider $\partial_\mu \phi = (\partial_0\phi, \partial_1\phi, \partial_2\phi, \partial_3\phi)$. This is *not* a vector field as it does not transform like x^μ. The transformation rule is

$$\partial_{\mu'} = \Lambda^{\nu}_{\mu'} \partial_\nu,$$

where $\Lambda^{\nu}_{\mu'}$ are the entries of the 4×4 matrix $\tilde{\Lambda} = (\Lambda^T)^{-1}$, where as before Λ is the matrix representing the Lorentz transformation ($x^{\mu'} = \Lambda^{\mu'}_{\nu} x^\nu$, $\Lambda^{\mu'}_{\nu}$ being the entries of Λ). To see this use

$$\partial_\mu x^\nu = \delta^\nu_\mu,$$

where δ^ν_μ is the Kronecker delta defined by $\delta^\nu_\mu = 1$ if $\mu = \nu$ and $\delta^\nu_\mu = 0$ if $\mu \neq \nu$. Consider

$$\delta^{\nu'}_{\mu'} = \partial_{\mu'} x^{\nu'} = \Lambda^{\alpha}_{\mu'} \partial_\alpha \Lambda^{\nu'}_{\beta} x^\beta = \Lambda^{\alpha}_{\mu'} \Lambda^{\nu'}_{\beta} \partial_\alpha x^\beta = \Lambda^{\alpha}_{\mu'} \Lambda^{\nu'}_{\beta} \delta^\beta_\alpha = \Lambda^{\alpha}_{\mu'} \Lambda^{\nu'}_{\alpha}.$$

In matrix language this is $I = \tilde{\Lambda}\Lambda^T$ where I is the 4×4 identity matrix. Accordingly,

$\tilde{\Lambda} = (\Lambda^T)^{-1}$. For pure rotations

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \mathsf{R} & \\ 0 & & & \end{pmatrix},$$

$\tilde{\Lambda} = \Lambda$.

As the four-gradient does not transform as a vector we have two distinct transformation properties!

Define a *contravariant* four vector v^μ as four numbers with the same transformation properties as the coordinates x^μ

$$v^{\mu'} = \Lambda^{\mu'}_\nu v^\nu \quad \text{under} \quad x^{\mu'} = \Lambda^{\mu'}_\nu x^\nu.$$

Define a *covariant* four-vector v_μ as four numbers with the same transformation properties as ∂_μ

$$v_{\mu'} = \Lambda^\nu_{\mu'} v_\nu,$$

where $\partial_{\mu'} = \Lambda^\nu_{\mu'}\partial_\nu$ and the $\Lambda^\nu_{\mu'}$ are the entries of the matrix $\tilde{\Lambda} = (\Lambda^T)^{-1}$.

If v^μ is a contravariant four-vector then

$$v_\mu = \eta_{\mu\nu} v^\nu,$$

is a covariant four-vector. Here we have used the metric $\eta_{\mu\nu}$ to 'lower' the index μ producing a covariant vector. This process can be reversed. Consider

$$\eta^{\mu\nu} = \text{diag}(1, -1, -1, -1),$$

which we call the inverse metric. If v_μ is covariant then the index can be 'raised' using the inverse metric to produce a contravariant vector

$$v^\mu = \eta^{\mu\nu} v_\nu.$$

The inner product $v \cdot w = \eta_{\mu\nu} v^\mu w^\nu$ can be rewritten as $v^\mu w_\mu$ or $v_\mu w^\mu$. Similarly $v \cdot v = v^\mu v_\mu$.

These ideas can be applied to the gradient $\partial_\mu = (\partial_0, \partial_1, \partial_2, \partial_3)$. A contravariant version of the gradient is $\partial^\mu = \eta^{\mu\nu}\partial_\nu = (\partial_0, -\partial_1, -\partial_2, -\partial_3)$. A four dimensional Laplacian or 'box' operator is

$$\Box = \partial^\mu \partial_\mu = \partial_0^2 - \partial_1^2 - \partial_2^2 - \partial_3^2 = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2.$$

$\Box\phi = 0$ is the wave equation (the wave speed is c).

2 Tensors in Special Relativity

In this chapter Special Relativity is formulated using tensors. In particular the conservation of four momentum is formulated as a tensor equation. The metric symbol introduced in chapter 1 is interpreted as an isotropic tensor. Basis vectors are bypassed by defining tensors through the transformation properties of their components.

2.1 Tensors

A tensor of type (p, q) is an array of 4^{p+q} numbers

$$T^{\mu_1\mu_2\ldots\mu_p}_{\nu_1\nu_2\ldots\nu_q},$$

with a specific transformation property under Lorentz transformations. Here there are p contravariant indices (μ_1, μ_2,..., μ_p) and q covariant indices (ν_1, ν_2,..., ν_q). Each index takes the values $0, 1, 2$ or 3.

A scalar is a tensor of type $(0, 0)$. This is Lorentz invariant.

A contravariant vector, v^μ, is a tensor of type $(1, 0)$. This transforms the same way as the coordinates, x^μ,

$$v^{\mu'} = \Lambda^{\mu'}_\nu v^\nu,$$

under the Lorentz transformation, $x^{\mu'} = \Lambda^{\mu'}_\nu x^\nu$.

A covariant vector, v_μ, is a tensor of type $(0, 1)$. This transforms in the same way as the gradient, ∂_μ, under a Lorentz transformation

$$v_{\mu'} = \Lambda^\nu_{\mu'} v_\nu.$$

A tensor of type $(2, 0)$ is 16 numbers, $T^{\mu\nu}$, where $\mu = 0, 1, 2, 3$ and $\nu = 0, 1, 2, 3$. Under a Lorentz transformation, $x^{\mu'} = \Lambda^{\mu'}_\nu x^\nu$, the numbers change according to the rule

$$T^{\mu'\nu'} = \Lambda^{\mu'}_\alpha \Lambda^{\nu'}_\beta T^{\alpha\beta}.$$

A tensor of type $(1, 1)$ is 16 numbers $T^{\mu\nu}$ with the transformation rule

$$T^{\mu'}_{\nu'} = \Lambda^{\mu'}_\alpha \Lambda^\beta_{\nu'} T^\alpha_\beta.$$

A tensor of type $(0, 2)$, $T_{\mu\nu}$, has the transformation rule

$$T_{\mu'\nu'} = \Lambda^\alpha_{\mu'} \Lambda^\beta_{\nu'} T_{\alpha\beta}.$$

The following examples are all isotropic, meaning their components are the same in all inertial frames

The metric, $\eta_{\mu\nu}$, is a tensor of type $(0,2)$. While the metric 'looks' like a type $(0,2)$ tensor in that it has two lower indices it is important that it satisfies the transformation rule

$$\eta_{\mu'\nu'} = \Lambda^{\alpha}_{\mu'}\Lambda^{\beta}_{\nu'}\eta_{\alpha\beta}.$$

In matrix language this is[4]

$$\eta = \tilde{\Lambda}\eta\tilde{\Lambda}^T,$$

where as in chapter 1, $\Lambda^{\nu}_{\mu'}$ are the entries of $\tilde{\Lambda} = (\Lambda^T)^{-1}$. This is equivalent to $\Lambda^T\eta\Lambda = \eta$ which defines Lorentz transformations.

The inverse metric, $\eta^{\mu\nu}$ is a tensor of type $(2,0)$.

The Kronecker delta δ^{μ}_{ν} is a tensor of type $(1,1)$.

A tensor of type (p,q) has the transformation rule

$$T^{\mu'_1\ldots\mu'_p}_{\nu'_1\ldots\nu'_q} = \Lambda^{\mu'_1}_{\alpha_1}\ldots\Lambda^{\mu'_p}_{\alpha_p}\,\Lambda^{\beta_1}_{\nu'_1}\ldots\Lambda^{\beta_q}_{\nu'_q}\,T^{\alpha_1\ldots\alpha_p}_{\beta_1\ldots\beta_q}. \tag{12}$$

Note that $\Lambda^{\mu'}_{\nu}$ and $\Lambda^{\nu}_{\mu'}$ are not tensors.

2.2 Tensor Algebra

Tensors of the same type can be added.

Tensors of any type can be multiplied: multiplying a type (p,q) tensor by a type (r,s) gives a type $(p+r,q+s)$ tensor.

For example

$$T^{\mu\nu}_{\alpha}S^{\beta\gamma} = W^{\mu\nu\beta\gamma}_{\alpha},$$

multiplies a type $(2,1)$ tensor with a type $(2,0)$ tensor to produce a type $(4,1)$ tensor.

Vectors can be multiplied to yield tensors. The product of v^μ and w^ν, written $v^\mu w^\nu$, is a type $(2,0)$ tensor (called the *outer product* of v and w).

Type $(2,0)$ and type $(0,2)$ tensors can be split into symmetric an anti-symmetric parts (much as for square matrices)

$$T^{\mu\nu} = S^{\mu\nu} + A^{\mu\nu},$$

where $S^{\mu\nu} = S^{\nu\mu}$ (symmetric) and $A^{\mu\nu} = -A^{\nu\mu}$ (antisymmetric). Here

$$S^{\mu\nu} = \frac{T^{\mu\nu}+T^{\nu\mu}}{2}, \qquad A^{\mu\nu} = \frac{T^{\mu\nu}-T^{\nu\mu}}{2}.$$

[4] If A and B are square matrices $A_{ij}B_{jk} = (AB)_{ik}$ and $A_{ji}B_{jk} = (A^TB)_{ik}$.

Note that $S^{\mu\nu}$ has 10 independent components. $A^{\mu\nu}$ has 6 as the diagonal entries are zero and

$$A^{12} = -A^{21}, \quad A^{23} = -A^{32}, \quad A^{31} = -A^{13}.$$

2.3 Contraction

Given a tensor of type (p, q) one can obtain a tensor of type $(p-1, q-1)$ by summing over a pair of contravariant and covariant indices. For example a type $(2, 1)$ tensor, $T^{\mu\nu}_{\alpha}$, can be contracted in two ways

$$T^{\mu\nu}_{\mu}, \quad \text{or} \quad T^{\mu\nu}_{\nu},$$

which are of type $(1, 0)$, that is a contravariant vector.

Multiple contractions are possible. For example $\eta_{\mu\nu} v^{\alpha} w^{\beta}$ can be contracted twice to yield $\eta_{\alpha\beta} v^{\alpha} w^{\beta} = v \cdot w$ a tensor of type $(0, 0)$, a scalar. Using the metric (or inverse metric) to lower (or raise) indices is an example of contraction. $v_{\mu} = \eta_{\mu\nu} v^{\nu}$ is a contraction of $\eta_{\mu\nu} v^{\alpha}$.

Note that contracting the Kronecker delta yields the number 4 not 1

$$\delta^{\mu}_{\mu} = \delta^{0}_{0} + \delta^{1}_{1} + \delta^{2}_{2} + \delta^{3}_{3} = 4.$$

Double contracting a symmetric type $(2, 0)$ tensor with an anti-symmetric type $(0, 2)$ tensor gives zero

$$S^{\mu\nu} A_{\mu\nu} = 0,$$

as the term involving, say A_{01}, cancels the term involving A_{10}.

The metric (and inverse metric) can also be used to lower (and raise) tensor indices. For example $T^{\mu}_{\ \nu} = \eta_{\nu\alpha} T^{\mu\alpha}$

2.4 Differentiation of Tensors

Suppose $T^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q}$ is a tensor of type (p, q) then its derivative,

$$\partial_{\alpha} T^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q},$$

is a tensor of type $(p, q+1)$. For example if $T^{\mu\nu}$ is a tensor of type $(2, 0)$ then $\partial_{\alpha} T^{\mu\nu}$ is a tensor of type $(2, 1)$. By definition ∂_{α} transforms like a covariant vector. That is, ∂_{α} is a tensor operator of type $(0, 1)$.

Let v^{μ} be of type $(1, 0)$, then $\partial_{\alpha} v^{\mu}$ is of type $(1, 1)$. Contracting this yields the scalar

$$\partial_{\mu} v^{\mu} = \partial_0 v^0 + \partial_1 v^1 + \partial_2 v^2 + \partial_3 v^3,$$

which is a four dimensional version of the divergence. The Kronecker delta can also be expressed as a derivative

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

A tensor can be differentiated more than once, e.g. $\partial_\mu \partial_\nu \phi$ is a tensor of type $(0, 2)$ obtained by differentiating the type $(0, 0)$ scalar ϕ twice.

As the metric is isotropic

$$\partial_\alpha \eta_{\mu\nu} = 0.$$

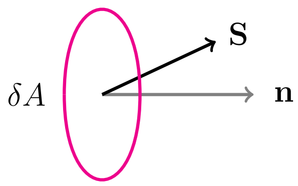
Although this is apparently trivial, the question of how this equation extends to non-inertial frames is key to the development of General Relativity.

2.5 Energy Momentum Tensor

The total four-momentum of an isolated system is conserved. In Special Relativity it is useful to reformulate this in a local form. In General Relativity this is essential. Start with the conservation of energy (or conservation of the zeroth component of four momentum).

Assume that energy is localised

Let $u(x)$ be the energy density (energy per unit volume) and assume that there is a three vector field, $\mathbf{S}(x)$ representing the flux of energy. That is the flow of energy across a small surface element of area δA is given by $\mathbf{S} \cdot \mathbf{n}\, \delta A$ ($\mathbf{n}$ is normal to the surface element).



Consider a bounded three-dimensional region R (assumed fixed). The energy in R is given by the integral

$$E(t) = \int_R u \, dV. \tag{13}$$

This is time-dependent as the energy density u is time-dependent. As the region does not change with time

$$\frac{dE}{dt} = \int_R \frac{\partial u}{\partial t} \, dV.$$

If energy is conserved this should be equal to the flow of energy into R

$$-\oint_{\partial R} \mathbf{S} \cdot \mathbf{n} \, dA,$$

where ∂R is the boundary of R, and $\mathbf{n}$ is the outward normal to the surface ∂R. The divergence theorem,

$$\oint_{\partial R} \mathbf{S} \cdot \mathbf{n}\, dA = \int_R \nabla \cdot \mathbf{S}\, dV,$$

gives

$$\int_R \left(\frac{\partial u}{\partial t} + \nabla \cdot \mathbf{S} \right) dV = 0.$$

As R is arbitrary

$$\frac{\partial u}{\partial t} + \nabla \cdot \mathbf{S} = 0. \tag{14}$$

Note that the derivation is similar to that for continuity equations in fluid mechanics.

This equation, along with similar equations for the local conservation of momentum can be combined into a single tensor equation. The energy momentum tensor $T^{\mu\nu}$ is defined through

$$T^{00} = u, \quad T^{0i} = \frac{S_i}{c}, \quad T^{i0} = cg_i, \tag{15}$$

where S_i is the ith component of $\mathbf{S}$, g_i is the ith component of $\mathbf{g}$ which is the momentum density. T^{ij} is the flux of the ith component of momentum in the x^j direction. The local conservation of energy and momentum can be written as a single tensor equation

$$\partial_\nu T^{\mu\nu} = 0. \tag{16}$$

The left hand side is a tensor of type $(1, 0)$ so this is four equations (conservation of energy and the three components of momentum). For $\mu = 0$ this is $\partial_0 T^{00} + \partial_1 T^{01} + \partial_2 T^{02} + \partial_3 T^{03} = 0$. Inserting $T^{00} = u$ and $T^{0i} = S_i/c$ yields (14).

Claim: The energy momentum tensor is symmetric

It is not possible to 'prove' this assertion as there is no general definition of energy and momentum. It is possible to verify it in particular cases such as electromagnetic theory. A simpler example is as follows. Consider a swarm of particles all with mass m and velocity $\mathbf{v}$[5]. The energy and momentum of each particle is

$$E = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad \mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}},$$

respectively. Therefore $\mathbf{p} = E\mathbf{v}/c^2$. Multiplying by ρ_n, the number of particles per unit volume, gives $\rho_n \mathbf{p} = \rho_n E \mathbf{v}/c^2$. Now $\rho_n \mathbf{p}$ and $\rho_n E$ are the momentum density, $\mathbf{g}$, and the energy density, u, respectively. Accordingly,

$$\mathbf{g} = u\mathbf{v}/c^2.$$

[5]This is based on the discussion in section 27-6, Volume 2, of the Feynman Lectures on Physics.

Interpreting $u\mathbf{v}$ as the energy flux $\mathbf{S}$ gives $\mathbf{g} = \mathbf{S}/c^2$ which is equivalent to $T^{i0} = T^{0i}$.

An example of an energy momentum tensor is

$$T^{\mu\nu} = \left(\rho + \frac{p}{c^2}\right) u^\mu u^\nu - p\eta^{\mu\nu},$$

which is for a perfect relativistic fluid. Here p is the pressure, u^μ is the four-velocity of the fluid and ρ is the *proper density*. The proper density at a point is defined as the density (mass per unit volume) in an inertial frame where the fluid is stationary ($u^\mu = (c, 0, 0, 0)$) at the given point; the proper density is Lorentz invariant. The equation $\partial_\nu T^{\mu\nu} = 0$ gives the equations of motion for a relativistic fluid (see Problem Sheet 2).

The energy formula (13) can be generalised to a formula for the total four-momentum

$$P^\mu = \frac{1}{c}\int T^{\mu 0}\, dV. \tag{17}$$

2.6 Geometrical Formulation

In this chapter tensors have been defined through their components. That is as arrays of numbers with certain transformation properties under Lorentz transformations. In three dimensions a vector $\mathbf{V}$ may be understood as an object with magnitude and direction. For actual calculations it is convenient to express vectors using a set of basis vectors $\mathbf{e}_i$ with $i = 1, 2, 3$, that is

$$\mathbf{V} = V_1\mathbf{e}_1 + V_2\mathbf{e}_2 + V_3\mathbf{e}_3,$$

or $\mathbf{V} = V_i\mathbf{e}_i$ using the summation convention. Under a change of basis vectors the components will change

$$(V_i)' = R_{ij}V_j.$$

If the old and new bases are orthonormal R_{ij} is an orthogonal matrix. Note that under such a transformation the vector $\mathbf{V}$ does not change.

The same ideas apply to four vectors. A four vector can be understood as $v = v^\mu e_\mu$, where the components, v^μ, transform like the coordinates under a Lorentz transformation. Tensors of type $(p, 0)$ can be understood in a similar fashion; for $p = 2$

$$T = T^{\mu\nu} e_\mu \otimes e_\nu.$$

The tensor T can be viewed as a geometric object which does not change under a Lorentz transformation. In this module tensors will be developed through the transformation properties of their components which avoids the use of basis vectors. An overview of the geometrical approach to vectors and tensors is given in Chapter 6.

3 Tensors

In chapter 2 tensors were developed for inertial frames. In this chapter tensors are developed in general coordinate systems. This extension is rather straightforward except for the differentiation of tensors. Covariant derivatives and the related notion of parallel transport are introduced. The Levi-Civita connection is considered in detail. As in chapter 2, basis vectors are not used.

3.1 General Coordinate Systems

In Special Relativity contravariant vectors v^μ have the transformation property

$$v^{\mu'} = \Lambda^{\mu'}_{\nu} v^\nu,$$

under Lorentz transformations $x^{\mu'} = \Lambda^{\mu'}_{\nu} x^\nu$. Differentiating the Lorentz transformation gives $\partial x^{\mu'}/\partial x^\alpha = \Lambda^{\mu'}_{\nu}\delta^\nu_\alpha = \Lambda^{\mu'}_{\alpha}$, that is

$$\Lambda^{\mu'}_{\nu} = \frac{\partial x^{\mu'}}{\partial x^\nu}.$$

Covariant four-vectors transform according to $v_{\mu'} = \Lambda^{\nu}_{\mu'} v_\nu$, where

$$\Lambda^{\nu}_{\mu'} = \frac{\partial x^\nu}{\partial x^{\mu'}}.$$

That is the constant matrices entering the Lorentz transformation rules for contravariant and covariant vectors are Jacobian matrices.

We will use (non-constant) Jacobians to give a more general definition of contravariant and covariant vector fields. Let x^a $(a = 1, 2, ..., N)$ be coordinates of an N dimensional space (in this chapter N is not necessarily 4). Consider an arbitrary change of coordinates

$$x^{a'} = x^{a'}(x).$$

The new coordinates $x^{a'}$ are N functions of the old coordinates x^b. The chain rule gives

$$dx^{a'} = \frac{\partial x^{a'}}{\partial x^b} dx^b, \tag{18}$$

using the summation convention. Define

$$\partial_a = \frac{\partial}{\partial x^a}.$$

The chain rule gives

$$\partial_{a'} = \frac{\partial x^b}{\partial x^{a'}} \frac{\partial}{\partial x^b} = \frac{\partial x^b}{\partial x^{a'}} \partial_b.$$

A contravariant vector field is N functions, $v^a(x)$, with the transformation property

$$v^{a'} = \frac{\partial x^{a'}}{\partial x^b} v^b,$$

under a general change of coordinates.

A covariant vector field is N functions, $v_a(x)$, with the transformation property

$$v_{a'} = \frac{\partial x^b}{\partial x^{a'}} v_b.$$

In other words

v^a transforms like dx^a,

v_a transforms like ∂_a.

Example: in the plane ($N = 2$) take the 'old' coordinates to be polar coordinates (r, θ) and the 'new' coordinates to be cartesian (x, y)

$$x = r\cos\theta, \quad y = r\sin\theta.$$

Suppose $v^r = 0$, $v^\theta = 1/r$. The differentials of the new coordinates are

$$dx = \frac{\partial x}{\partial r}dr + \frac{\partial x}{\partial \theta}d\theta = \cos\theta dr - r\sin\theta d\theta,$$

$$dy = \frac{\partial y}{\partial r}dr + \frac{\partial y}{\partial \theta}d\theta = \sin\theta dr + r\cos\theta d\theta.$$

Accordingly

$$v^x = \cos\theta v^r - r\sin\theta v^\theta = -\sin\theta,$$

$$v^y = \sin\theta v^r + r\cos\theta v^\theta = \cos\theta.$$

See problem sheet 3 for a similar problem transforming covariant vectors.

Now that contravariant and covariant vectors have been introduced it is straightforward to define tensors. A tensor field of type (p, q) is N^{p+q} functions

$$T^{a_1 a_2 \ldots a_p}{}_{b_1 b_2 \ldots b_q},$$

with the transformation property

$$T^{a'_1 a'_2 \ldots a'_p}{}_{b'_1 b'_2 \ldots b'_q} = \frac{\partial x^{a'_1}}{\partial x^{c_1}} \cdots \frac{\partial x^{a'_p}}{\partial x^{c_p}} \frac{\partial x^{d_1}}{\partial x^{b'_1}} \cdots \frac{\partial x^{d_q}}{\partial x^{b'_q}} T^{c_1 c_2 \ldots c_p}{}_{d_1 d_2 \ldots d_q}. \tag{19}$$

The rules of addition, multiplication and contraction are as in chapter 2. Note that vectors or tensors at different points cannot be added. This is because the partial derivatives,

$\partial x^{a'}/\partial x^b$, are not constants. In other words tensor algebra is local. As in chapter 2 the Kronecker delta

$$\delta^a_b = \frac{\partial x^a}{\partial x^b}$$

is a tensor of type $(1,1)$.

3.2 Metric Tensor

The metric $g_{ab}(x)$ gives the distance squared[6] between neighbouring points through

$$ds^2 = g_{ab}(x)dx^a dx^b. \tag{20}$$

g_{ab} is a symmetric tensor of type $(0,2)$.

The inverse metric g^{ab} is a symmetric tensor of type $(2,0)$ defined through

$$g^{ab}(x)g_{bc}(x) = \delta^a_c.$$

In matrix language g^{ab} is the inverse of the metric.

Example: in 2d polars $ds^2 = dr^2 + r^2 d\theta^2$, so that $g_{rr} = 1, \quad g_{\theta\theta} = r^2, \quad g_{r\theta} = g_{\theta r} = 0$. The components of the inverse metric are $g^{rr} = 1, \quad g^{\theta\theta} = 1/r^2, \quad g^{r\theta} = g^{\theta r} = 0$.

As in Special Relativity one can use the metric (inverse metric) to lower (raise) tensor indices. For example if T^{ab} is type $(2,0)$ define $T^a_b = g_{bc}T^{ac}$. To raise indices use the inverse metric $T^a_b = g^{ac}T_{cb}$.

A 'contravariant derivative' is defined through

$$\partial^a = g^{ab}\partial_b.$$

For the 2d polar example

$$\partial^r = g^{rb}\partial_b = g^{rr}\partial_r + g^{r\theta}\partial_\theta = g^{rr}\partial_r = \partial_r.$$

Similarly,

$$\partial^\theta = g^{\theta b}\partial_b = g^{\theta\theta}\partial_\theta = \frac{1}{r^2}\partial_\theta.$$

An instructive example is to consider spherical polar coordinates. In three dimensional space with cartesian coordinates (x,y,z) the standard metric is

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

[6] In Special and General Relativity ds^2 can be negative. The results of this chapter apply equally to spaces where ds^2 is positive definite and space-times where it can be positive, zero or negative.

Spherical polar coordinates (r, θ, ϕ) are defined through

$$x = r \sin\theta \cos\phi, \quad y = r \sin\theta \sin\phi, \quad z = r\cos\theta.$$

In this coordinate system the metric has the form

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta \; d\phi^2. \tag{21}$$

The components can be read off from (21)

$$g_{rr} = 1, \quad g_{\theta\theta} = r^2, \quad g_{\phi\phi} = r^2 \sin^2\theta,$$

and the off-diagonal components are zero. Setting r to be a constant (say $r = 1$) gives a sphere which is a curved space (see chapter 4) with metric

$$ds^2 = d\theta^2 + \sin^2\theta \; d\phi^2. \tag{22}$$

As in Chapter 2 an inner product of two vectors v^a and w^a is defined through

$$v \cdot w = g_{ab} v^a w^b.$$

This is a scalar, that is invariant under general coordinate transformations. See Problem Sheet 3 for examples.

3.3 Differentiation of Tensors

In Special Relativity the derivative of a tensor of type (p, q) is a tensor of type $(p, q + 1)$. For example, start with $v^\mu(x)$ type $(1, 0)$, its derivative is $\partial_\nu v^\mu(x)$ type $(1, 1)$. Differentiating again gives $\partial_\lambda \partial_\nu v^\mu(x)$ type $(1, 2)$.

In general coordinates, if ϕ is a scalar - type $(0, 0)$, then $\partial_a \phi$ is a covariant vector field - type $(0, 1)$

$$\partial_{a'}\phi = \frac{\partial x^b}{\partial x^{a'}} \frac{\partial \phi}{\partial x^b} = \frac{\partial x^b}{\partial x^{a'}} \partial_b \phi.$$

However, if $v^a(x)$ is a contravariant vector field, $\partial_c v^a(x)$ is *not* a tensor of type $(1, 1)$ as

$$\partial_{c'} v^{a'} = \frac{\partial x^b}{\partial x^{c'}} \frac{\partial}{\partial x^b} \frac{\partial x^{a'}}{\partial x^d} v^d = \frac{\partial x^b}{\partial x^{c'}} \frac{\partial x^{a'}}{\partial x^d} \partial_b v^d + \frac{\partial x^b}{\partial x^{c'}} \frac{\partial}{\partial x^b} \left(\frac{\partial x^{a'}}{\partial x^d} \right) v^d.$$

For $\partial_c v^a$ to be a tensor the second term should be absent. For linear transformations

$$\frac{\partial}{\partial x^b} \left(\frac{\partial x^{a'}}{\partial x^d} \right)$$

is zero. This is why the issue does not arise in Special Relativity as

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

are constants for Lorentz (and Poincaré) transformations.

3.4 Covariant Derivative

In general, the derivative of a type (p, q) tensor is not a tensor unless $p = q = 0$. A way round this is to use an alternative definition of differentiation. Replace the derivative $\partial_a = \partial/\partial x^a$ with the *covariant derivative*[7] ∇_a. The definition of the covariant derivative depends on what it is acting on. For a scalar field ϕ, $\nabla_a \phi = \partial_a \phi$, that is the covariant derivative of a scalar is the same as the partial derivative. The covariant derivative of a tensor of type (p, q),

$$\nabla_c T^{a_1 \ldots a_p}_{ b_1 \ldots b_q},$$

is a tensor of type $(p, q+1)$. The covariant derivative of tensor products satisfies the Leibniz property[8]

$$\nabla_c T^{a_1 \ldots a_p}_{ b_1 \ldots b_q} U^{d_1 \ldots d_r}_{ e_1 \ldots e_s} = (\nabla_c T^{a_1 \ldots a_p}_{ b_1 \ldots b_q}) U^{d_1 \ldots d_r}_{ e_1 \ldots e_s} + T^{a_1 \ldots a_p}_{ b_1 \ldots b_q} \nabla_c U^{d_1 \ldots d_r}_{ e_1 \ldots e_s}. \tag{23}$$

Consider a contravariant vector field $v^a(x)$. The covariant derivative is defined through

$$\nabla_c v^a = \partial_c v^a + \Gamma^a_{bc} v^b. \tag{24}$$

Γ^a_{bc} is the *connection*. It is not a tensor; it transforms in such a way that $\nabla_c v^a$ is a tensor of type $(1, 1)$. As we shall see Γ^a_{bc} is not unique. In General Relativity a particular connection, the *Levi-Civita* connection, is used. This connection is defined through

$$\Gamma^a_{bc} = \frac{1}{2} g^{ad} \left(\partial_b g_{dc} + \partial_c g_{bd} - \partial_d g_{bc} \right). \tag{25}$$

Γ^a_{bc} defined as in (25) is also called a Christoffel symbol. If the components of the metric are constant (as for inertial frames in Special Relativity) the Levi-Civita connection is zero and the covariant derivative reduces to the partial derivative.

For a covariant vector field

$$\nabla_c v_b = \partial_c v_b - \Gamma^a_{bc} v_a,$$

which is a tensor of type $(0, 2)$. Note the minus sign! This ensures that the covariant derivative of a scalar obtained by contracting a contravariant and covariant vector is just a

[7] The terminology can be confusing as the word covariant is used for specific vectors and a derivative.

[8] Suppressing the indices leads to a more compact form $\nabla_c TU = (\nabla_c T)U + T\nabla_c U$.

partial derivative due to a cancellation of connection terms

$$\nabla_c u^a v_a = (\nabla_c u^a) v_a + u^a (\nabla_c v_a)$$

$$= (\partial_c u^a + \Gamma^a_{bc} u^b) v_a + u^a (\partial_c v_a - \Gamma^d_{ac} v_d) = (\partial_c u^a) v_a + u^a (\partial_c v_a) = \partial_c (u^a v_a).$$

The covariant vector of a tensor T (suppressing the indices) of type (p, q) has the form $\nabla_c T = \partial_c T$ plus connection terms for each contravariant and covariant index. For a tensor of type $(2, 0)$

$$\nabla_c T^{ab} = \partial_c T^{ab} + \Gamma^a_{dc} T^{db} + \Gamma^b_{dc} T^{ad},$$

and for a tensor of type $(0, 2)$

$$\nabla_c T_{ab} = \partial_c T_{ab} - \Gamma^d_{ac} T_{db} - \Gamma^d_{bc} T_{ad}.$$

.

Example: 2d polar coordinates. The metric is $ds^2 = dr^2 + r^2 d\theta^2$ with components $g_{rr} = 1$, $g_{\theta\theta} = r^2$, $g_{r\theta} = g_{\theta r} = 0$. As there are two coordinates and Γ^a_{bc} has three indices, Γ^a_{bc} has 8 components. Using (25) and setting $a = r$

$$\Gamma^r_{bc} = \frac{1}{2} g^{rd} \left(\partial_b g_{dc} + \partial_c g_{bd} - \partial_d g_{bc} \right).$$

Only $d = r$ contributes to the sum over d

$$\Gamma^r_{bc} = \frac{1}{2} g^{rr} \left(\partial_b g_{rc} + \partial_c g_{br} - \partial_r g_{bc} \right) = \frac{1}{2} (0 + 0 - \partial_r g_{bc}).$$

This is zero unless $b = c = \theta$. Using $g_{\theta\theta} = r^2$, $\Gamma^r_{\theta\theta} = -r$. Setting $a = \theta$ in (25) only $d = \theta$ contributes

$$\Gamma^\theta_{bc} = \frac{1}{2} g^{\theta\theta} \left(\partial_b g_{\theta c} + \partial_c g_{b\theta} - \partial_\theta g_{bc} \right).$$

Using $g^{\theta\theta} = 1/r^2$ this gives $\Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = 1/r$. Note that 5 components of Γ^a_{bc} are zero.

Another example is the unit sphere with metric $ds^2 = d\theta^2 + \sin^2\theta d\phi^2$ as in the previous example Γ^a_{bc} has eight components (see Problem Sheet 3).

3.5 Parallel Transport

We have seen that the derivative of a tensor is not a tensor and have outlined a 'fix' for this problem which is to replace partial derivative ∂_c with a covariant derivative ∇_c. Here we consider the 'failure' of partial differentiation in detail and develop the idea of connections and covariant derivatives from the notion of *parallel transport*.

What's wrong with partial differentiation? The partial derivative of a scalar field ϕ is defined through

$$\partial_c \phi = \lim_{\epsilon \to 0} \frac{\phi(x^r + \epsilon \delta^r_c) - \phi(x^r)}{\epsilon}.$$

For a contravariant vector field v^a

$$\partial_c v^a = \lim_{\epsilon \to 0} \frac{v^a(x^r + \epsilon \delta^r_c) - v^a(x^r)}{\epsilon}.$$

For finite ϵ the numerator is the difference of vectors at different points. As the transformation rule for contravariant vectors,

$$v^{a'} = \frac{\partial x^{a'}}{\partial x^b} v^b,$$

is position-dependent, adding (or subtracting) vectors at different points does not give a vector. The $\epsilon \to 0$ limit does not yield a tensor.

It does not make sense to add a vector defined at the point x_A to a vector defined at another point x_B. Now 'transport' a vector defined at a point x_A along a curve C to the point x_B.

$$P_C(x_A, x_B)^a_b v^b$$

is a vector defined at x_B with the transformation properties of a vector defined at x_B. $P_C(x_A, x_B)^a_b$ is a linear transformation which turns a vector defined at one point into a vector defined at another; it depends on the curve joining the two points. Now take C to be a small line segment joining neighbouring points x_A and x_B and write

$$P(x_A, x_B)^a_b = \delta^a_b - \Gamma^a_{bc} \delta x^c,$$

where $\delta x^c = x^c_B - x^c_A$. Γ^a_{bc} is called the connection. The change in the vector due to transportation along the segment is

$$\delta v^a = -\Gamma^a_{bc} \delta x^c v^b.$$

This can also be written as a differential equation. Parametrise C via $x^r = x^r(\lambda)$ where λ is a parameter. The *equation of parallel transport* is

$$\frac{dv^a}{d\lambda} + \Gamma^a_{bc} \frac{dx^c}{d\lambda} v^b = 0. \tag{26}$$

Example: In the plane with standard metric $ds^2 = dx^2 + dy^2 = dr^2 + r^2 d\theta^2$. In cartesian coordinates the metric components are constant so that $\Gamma^a_{bc} = 0$ meaning parallel transport is trivial. In polar coordinates

$$\Gamma^r_{\theta\theta} = -r, \quad \Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = \frac{1}{r},$$

with the other five components zero. Accordingly

$$\frac{dv^r}{d\lambda} + \Gamma^r_{bc}\frac{dx^c}{d\lambda}v^b = \frac{dv^r}{d\lambda} + \Gamma^r_{\theta\theta}\frac{d\theta}{d\lambda}v^\theta = \frac{dv^r}{d\lambda} - r\frac{d\theta}{d\lambda}v^\theta = 0.$$

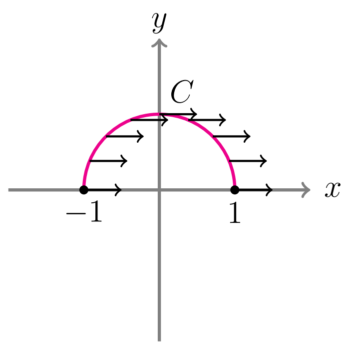
Similarly

$$\frac{dv^\theta}{d\lambda} + \frac{1}{r}\left(\frac{d\theta}{d\lambda}v^r + \frac{dr}{d\lambda}v^\theta\right) = 0.$$

Start at $(x, y) = (1, 0)$ that is $r = 1$ and $\theta = 0$. Take $r = 1$ and $\lambda = \theta$ so that C is the unit circle

$$\frac{dv^r}{d\theta} - v^\theta = 0, \quad \frac{dv^\theta}{d\theta} + v^r = 0,$$

which is easy to solve. For example with the initial values $v^x = 1$, $v^y = 0$ or $v^r = 1$, $v^\theta = 0$. The solution of the coupled ODEs is $v^r = \cos\theta$, $v^\theta = -\sin\theta$. In the diagram below the vector is transported along a semi-circle to $(x, y) = (-1, 0)$ or $\theta = \pi$.



Returning to differentiation

$$\partial_c v^a = \lim_{\epsilon\to 0}\frac{v^a(x^r + \epsilon\delta^r_c) - v^a(x^r)}{\epsilon}$$

is not a tensor. Now consider

$$\frac{v^a(x^r + \epsilon\delta^r_c) - P^a_b v^b(x^r)}{\epsilon}$$

where P^a_b transports a vector from x^r to $x^r + \epsilon\delta^r_c$

$$P^a_b = \delta^a_b - \Gamma^a_{br}\delta x^r = \delta^a_b - \Gamma^a_{br}\epsilon\delta^r_c = \delta^a_b - \epsilon\Gamma^a_{bc}.$$

Using this and taking the $\epsilon \to 0$ limit yields the covariant derivative

$$\nabla_c v^a = \partial_c v^a + \Gamma^a_{bc}v^b.$$

Parallel transport and covariant derivatives are based on the connection Γ^a_{bc}. In General Relativity the Levi-Civita form of the connection (25) is used. This formula follows from two assumptions about the connection

(i) the connection is *metric*.

(ii) the connection is *torsion free*.

The metric $g_{ab}(x)$ depends on x. Assume that the dependence on x of $g_{ab}(x)$ can be accounted for by parallel transport. Suppose that $g_{ab}(x_A)$ is known at the point x_A. Then transport g_{ab} along a curve C to another point x_B. If the transported g_{ab} is the same as the actual metric at x_B for any x_B the connection is called metric. If the connection is metric

$$\nabla_c g_{ab} = 0. \tag{27}$$

More explicitly

$$\nabla_c g_{ab} = \partial_c g_{ab} - \Gamma^d_{ac} g_{db} - \Gamma^d_{bc} g_{ad} = 0. \tag{28}$$

Although Γ^a_{bc} is not a tensor it is useful to lower the upper index via the definition

$$\Gamma_{abc} = g_{ad}\Gamma^d_{bc},$$

which can be inverted in the usual way through $\Gamma^a_{bc} = g^{ad}\Gamma_{dbc}$. With this notation (28) becomes

$$\nabla_c g_{ab} = \partial_c g_{ab} - \Gamma_{bac} - \Gamma_{abc} = 0. \tag{29}$$

Assumption (ii) is the torsion-free condition

$$\Gamma^a_{bc} = \Gamma^a_{cb},$$

or

$$\Gamma_{abc} = \Gamma_{acb}.$$

This together with (29) leads to

$$\Gamma_{abc} = \frac{1}{2}\left(\partial_b g_{ac} + \partial_c g_{ab} - \partial_a g_{bc}\right),$$

which is equivalent to the Levi-Civita connection (25).

4 Geodesics and Curvature

This chapter introduces the idea of geodesics which generalises the elementary notion of straight lines. The property of geodesic deviation leads to the definition of the curvature tensor (or Riemann tensor). The properties of this type $(1,3)$ tensor and related objects are developed.

4.1 Geodesics

Consider a curve C with form $x^a = x^a(\lambda)$ where λ is a parameter. The parameter can be taken to be the distance s along the curve so that $x^a = x^a(s)$. Define

$$u^a(s) = \frac{dx^a}{ds}.$$

Using $ds^2 = g_{ab}(x)dx^a dx^b$ it follows that[9]

$$g_{ab}u^a u^b = 1. \tag{30}$$

Now consider the elementary idea of straight lines. In the plane with cartesian coordinates $u = (dx/ds, dy/ds)$, and for straight lines the components dx/ds and dy/ds are constants and so

$$\frac{du^a}{ds} = 0.$$

This characterisation of straight lines is specific to the plane with the standard metric $ds^2 = dx^2 + dy^2$. The result is not coordinate independent. For example, straight lines do not satisfy $du^r/ds = 0$ and $du^\theta/ds = 0$ where r and θ are polar coordinates. More precisely, du^a/ds is not a contravariant vector as the numerator in the limit

$$\lim_{\epsilon\to 0} \frac{u^a(s+\epsilon) - u^a(s)}{\epsilon},$$

is the difference of vectors at different points. In general coordinates use

$$\frac{Du^a}{ds} = 0,$$

where

$$\frac{Dv^a}{ds} = \lim_{\epsilon\to 0} \frac{v^a(s+\epsilon) - P^a_b v^b(s)}{\epsilon},$$

and $P^a_b = \delta^a_b - \Gamma^a_{bc}\delta x^c = \delta^a_b - \epsilon\Gamma^a_{bc}u^c$ parallel transports vectors from $x^a(s)$ to $x^a(s+\epsilon)$, so

[9] This mimics the four velocity in Special Relativity where $u \cdot u = \eta_{\mu\nu}u^\mu u^\nu = c^2$ since $u^\mu = dx^\mu/d\tau = c dx^\mu/ds$.

that

$$\frac{Dv^a}{ds} = \frac{dv^a}{ds} + \Gamma^a_{bc} v^b u^c. \tag{31}$$

$Du^a/ds = 0$, yields the *geodesic equation*

$$\frac{d^2 x^a}{ds^2} + \Gamma^a_{bc} \frac{dx^b}{ds}\frac{dx^c}{ds} = 0. \tag{32}$$

The solutions called *geodesics* generalise the notion of straight lines.

Consider again the plane using polar coordinates. From Chapter 3 the non-zero Christoffel symbols are

$$\Gamma^r_{\theta\theta} = -r, \quad \Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = \frac{1}{r}.$$

Setting $a = r$ in (32) gives

$$\frac{d^2 r}{ds^2} + \Gamma^r_{bc} \frac{dx^b}{ds}\frac{dx^c}{ds} = 0.$$

Only $b = c = \theta$ contributes to the double sum giving

$$\frac{d^2 r}{ds^2} - r\left(\frac{d\theta}{ds}\right)^2 = 0.$$

Setting $a = \theta$ in (32)

$$\frac{d^2 \theta}{ds^2} + \Gamma^\theta_{bc} \frac{dx^b}{ds}\frac{dx^c}{ds} = 0,$$

which gives

$$\frac{d^2 \theta}{ds^2} + \frac{2}{r}\frac{dr}{ds}\frac{d\theta}{ds} = 0,$$

as there are contributions from $b = r$, $c = \theta$ and $b = \theta$, $\theta = r$. Another example is to obtain the geodesics on a sphere with metric $ds^2 = d\theta^2 + \sin^2\theta d\phi^2$ (see question 2 on Problem Sheet 4).

4.2 Solving the Geodesic Equation

The geodesic equation is N coupled second order non-linear ODEs. To solve these it is useful to identify quantities that are constant along geodesics. Equation (30) always provides one such constant. This is useful in tackling question 5 on Problem Sheet 4. Suppose that $g_{ab}(x)$ does not depend on one coordinate x^p (p is a fixed number between 1 and N) then $g_{pb}\frac{dx^b}{ds}$, is constant along geodesics, that is

$$\frac{d}{ds}\left(g_{pb}\frac{dx^b}{ds}\right) = 0. \tag{33}$$

See question 4 on Problem Sheet 4.

For example, consider the paraboloid $ds^2 = (1+4\rho^2)d\rho^2 + \rho^2 d\theta^2$ considered in question 6 of

Problem Sheet 3. As the metric does not depend on θ, $h = g_{\theta b}dx^b/ds = g_{\theta\theta}d\theta/ds = \rho^2 d\theta/ds$ is constant, 'solves' the geodesic equation for θ. One can extract four of the Christoffel symbols from this result as $dh/ds = 0$ can be written as

$$\frac{d^2\theta}{ds^2} + \frac{2}{\rho}\frac{d\rho}{ds}\frac{d\theta}{ds} = 0,$$

so that

$$\Gamma^\theta_{\rho\theta} = \Gamma^\theta_{\theta\rho} = \frac{1}{\rho}, \quad \Gamma^\theta_{\rho\rho} = \Gamma^\theta_{\theta\theta} = 0.$$

The result (33) is useful in solving geodesic equations. What if you are in the 'wrong' coordinate system? For example, the metric for the paraboloid in cartesian coordinates is $ds^2 = (1+4x^2)dx^2 + (1+4y^2)dy^2 + 4xydxdy$. g_{ab} depends on both x and y but $h = \rho^2 d\theta/ds$ is still constant!

The result can be stated in a coordinate independent form. Suppose that the covariant vector field $v_a(x)$ satisfies the *Killing equation*

$$\nabla_a v_b + \nabla_b v_a = 0, \tag{34}$$

then

$$v_a \frac{dx^a}{ds},$$

is constant along geodesics (see question 4 of Problem Sheet 4).

4.3 Curvature

In the plane using cartesian coordinates the metric components g_{ab} are constants so that the solutions of the geodesic equation have the form u^a =constant. For the sphere there are no coordinates for which the metric is constant. This is a feature of a *curved space*. Curvature is defined through a type $(1,3)$ tensor

$$R^a_{bcd} = \partial_c \Gamma^a_{bd} - \partial_d \Gamma^a_{bc} + \Gamma^a_{ec}\Gamma^e_{bd} - \Gamma^a_{ed}\Gamma^e_{bc}, \tag{35}$$

which is known as the *Riemann tensor*.

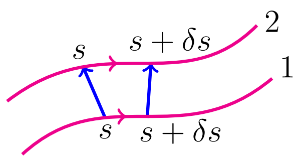
It is not obvious from the above formula that R^a_{bcd} is a tensor as the definition includes Christoffel symbols (not tensors) and partial derivatives. A direct proof that it is a tensor follows from its interpretation as a commutator of covariant derivatives (see later in this chapter).

It is clear from the definition that if the metric is constant the curvature is zero. As R^a_{bcd} is a tensor this will also hold if the coordinates are transformed to a different coordinate system; if a tensor is zero in one coordinate system it is zero in any other coordinate system.

4.4 Geodesic Deviation

In General Relativity Newton's first law is replaced by the geodesic equation. The geodesic equation also replaces Newton's second law and Newton's Universal Law of Gravitation as in General Relativity gravity is not a 'force'. Tidal forces can be understood through *geodesic deviation* which is the relative acceleration of particles on neighbouring geodesics.

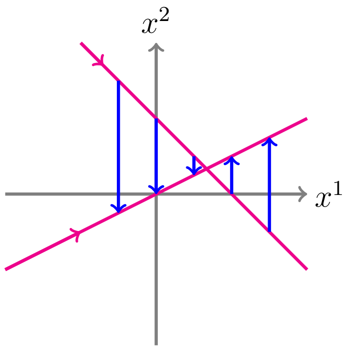
Consider two 'neighbouring geodesics' and the separation, $w^a(s) = x_2^a(s) - x_1^a(s)$, of 'neighbouring points' on the two geodesics (1 and 2 in the diagram below). Here there is some arbitrariness in synchronising the clocks (or defining $s = 0$) on the two geodesics.



If the metric components g_{ab} are constant, $\Gamma^a_{bc} = 0$, and geodesics are straight lines. In this case

$$\frac{d^2 w^a}{ds^2} = 0, \tag{36}$$

as $d^2x^a/ds^2 = 0$ for any geodesic (with $\Gamma^a_{bc} = 0$).



On changing coordinates so that g_{ab} is not constant (36) becomes

$$\frac{D^2 w^a}{ds^2} = 0, \tag{37}$$

where w^a is small or infinitesimal. One can make this more precise by considering a family of geodesics parametrised through $x^a = x^a(s,t)$ where s is the distance along a geodesic and the parameter t specifies the geodesic. Then, rather than considering w^a as small, it can be defined as a derivative

$$w^a = \frac{\partial x^a}{\partial t}.$$

However, (37) is not generally true. For each t the geodesic equation is

$$\frac{\partial^2 x^a}{\partial s^2} + \Gamma^a_{bc}\frac{\partial x^b}{\partial s}\frac{\partial x^c}{\partial s} = 0.$$

Partially differentiating this equation with respect to t leads to the *equation of geodesic deviation*

$$\frac{D^2 w^a}{\partial s^2} = -R^a_{bcd} u^b w^c u^d. \tag{38}$$

See Problem Sheet 4. Equation (38) is usually written in the form

$$\frac{D^2 w^a}{ds^2} = -R^a_{bcd} u^b w^c u^d. \tag{39}$$

If the Riemann tensor is non-zero the space is called curved; there is no coordinate system for which the components of the metric are constants. In General Relativity the equation of geodesic deviation describes tidal forces (see Chapter 5). If the Riemann tensor is zero everywhere then the theory reduces to Special Relativity (possibly in non-inertial coordinates).

Geodesic deviation gives a very concrete interpretation of the Riemann tensor. There are other ways of interpreting the Riemann tensor:

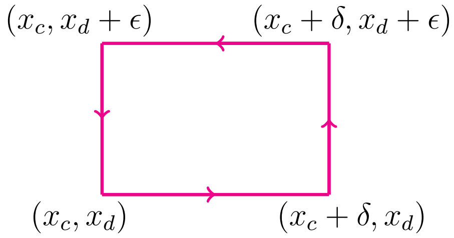
(a) as a commutator of covariant derivatives. The following formula

$$[\nabla_c, \nabla_d] v^a = R^a_{bcd} v^b, \tag{40}$$

shows that the Riemann tensor represents the action of the commutator, $[\nabla_c, \nabla_d] = \nabla_c \nabla_d - \nabla_d \nabla_c$, of covariant derivatives on a contravariant vector field v^a. As the left hand side is a tensor and v^b on the right hand side is arbitrary then R^a_{bcd} must be a tensor. What is the action of the commutator on a general tensor?

(b) representing parallel transport round a small rectangle. Recall that the Connection Γ^a_{bc} defines parallel transport along a small line segment in the x^c direction or $\delta v^a = -\Gamma^a_{bc} v^b \epsilon$ where ϵ is the increase in x^c along the segment. Transporting a vector around a small rectangle in the $x^c x^d$ plane as depicted in the diagram below gives

$$\delta v^a = -R^a_{bcd} v^b \epsilon \delta.$$



If $R^a_{bcd} = 0$ parallel transporting a vector around any small rectangle has no effect. In general if $R^a_{bcd} = 0$ everywhere parallel transport around a closed curve has no effect [10]

[10] During transport around a closed curve in flat space, the components of the vector change but on reaching the starting point the vector is unchanged.

4.5 Properties of the Riemann Tensor

The Riemann tensor has the following symmetry properties

$$R^a_{bcd} = -R^a_{bdc}. \tag{41}$$

$$R^a_{bcd} + R^a_{cdb} + R^a_{dbc} = 0. \tag{42}$$

$$R_{abcd} = R_{cdab}. \tag{43}$$

From (41) and (43) it follows that

$$R_{abcd} = -R_{bacd}. \tag{44}$$

Equation (41) is the property that the Riemann tensor is anti-symmetric in the last two covariant indices which is obvious from the definition. The property is also clear from (40). The symmetry is incorporated into the more compact formula

$$R^a_{bcd} = \partial_c \Gamma^a_{bd} + \Gamma^a_{ec}\Gamma^e_{bd} - c \leftrightarrow d.$$

Equation (42) is called the algebraic Bianchi identity.

Equation (43) concerns the type $(0,4)$ tensor $R_{abcd} = g_{ae}R^e_{bcd}$ obtained by lowering the contravariant index in the Riemann tensor. It states that the tensor is unchanged on exchanging the first and second pair of indices. The derivation is straightforward - the algebra can be reduced by taking Γ^a_{bc} to be zero at a given point (more later).

For $N = 4$ (as in General Relativity) the Riemann tensor has 256 components. The two asymmetry conditions (41) and (44) reduce the number of independent components from 256 to 36. Property (43) reduces the number of independent components to 21 and the algebraic Bianchi identity reduces the number to 20.

For $N = 2$ only one of the 16 components of R^a_{bcd} is independent. How many independent components does the Riemann tensor have in three dimensions?

4.6 Bianchi Identity

A key property of the Riemann tensor is the Bianchi identity

$$\nabla_e R^a_{bcd} + \nabla_c R^a_{bde} + \nabla_d R^a_{bec} = 0. \tag{45}$$

This can be derived from the Jacobi identity

$$[\nabla_e, [\nabla_c, \nabla_d]] + [\nabla_c, [\nabla_d, \nabla_e]] + [\nabla_d, [\nabla_e, \nabla_c]] = 0,$$

and the properties of the Riemann tensor including (40) and its generalisations.

4.7 Associated Tensors

The *Ricci tensor* is defined by the following contraction of the Riemann tensor

$$R_{bc} = -R^a_{bca} = R^a_{bac}. \tag{46}$$

Using the symmetry properties of the Riemann tensor it can be shown that the Ricci tensor is symmetric.

The *scalar curvature* is defined as a contraction of the Ricci tensor with the inverse metric

$$\mathcal{R} = g^{ab}R_{ab}. \tag{47}$$

It can also be written in the form $\mathcal{R} = R^a_a$ where $R^a_b = g^{ac}R_{cb}$.

4.8 Contracted Bianchi Identity

By contracting the Bianchi identity (45) it follows that

$$2\nabla_c R^c_b - \nabla_b \mathcal{R} = 0. \tag{48}$$

To derive (48) contract a with d in (45) and use the definition of the Ricci tensor which gives

$$-\nabla_e R_{bc} + \nabla_c R_{be} + \nabla_a R^a_{bec} = 0.$$

Multiplying by g^{be} gives[11]

$$-\nabla_e R^e_c + \nabla_c \mathcal{R} - \nabla_a R^a_c = 0,$$

which is the same as (48). Multiplying the contracted Bianchi identity by g^{cb} gives $2\nabla_a R^{ca} - \nabla_c g^{cb}\mathcal{R} = 0$ or

$$\nabla_a G^{ab} = 0, \tag{49}$$

where G^{ab} is the Einstein tensor is defined by

$$G^{ab} = R^{ab} - \frac{1}{2}\mathcal{R}g^{ab}.$$

Equation (49) is similar to $\partial_\nu T^{\mu\nu} = 0$ in Special Relativity. This idea will be developed further in chapter 5.

[11] Note that $g^{be}R^a_{bec} = -R^a_c$ and that the inverse metric commutes with the covariant derivative.

4.9 Normal Coordinates

If $R^a_{bcd} = 0$ everywhere (flat space) there are coordinates such that $\Gamma^a_{bc} = 0$ everywhere. In general $R^a_{bcd} \neq 0$ (curved space) but it is still possible to find *normal coordinates* so that $\Gamma^a_{bc} = 0$ at a given point - at the same point one can bring the metric to a diagonal form with entries ± 1, or zero. Suppose $\Gamma^a_{bc} \neq 0$ at some point (take this to be the origin $x^a = 0$). Consider the new coordinates $y^a = x^a + \frac{1}{2}\Gamma^a_{bc}(x^r = 0)x^b x^c$. For any geodesic passing through the origin

$$\frac{d^2y^a}{ds^2} = \frac{d^2x^a}{ds^2} + \Gamma^a_{bc}(x^r = 0)\frac{dx^b}{ds}\frac{dx^c}{ds} + \text{ terms vanishing at } x^r = 0$$

$$= -\Gamma^a_{bc}\frac{dx^b}{ds}\frac{dx^c}{ds} + \Gamma^a_{bc}(x^r = 0)\frac{dx^b}{ds}\frac{dx^c}{ds} + \text{ terms vanishing at } x^r = 0$$

which is zero at the origin. The geodesic equation in the new coordinates is

$$\frac{d^2y^a}{ds^2} + \tilde{\Gamma}^a_{bc}\frac{dy^b}{ds}\frac{dy^c}{ds} = 0,$$

where $\tilde{\Gamma}^a_{bc}$ is the Christoffel symbol for the new y^a coordinates. Therefore $\tilde{\Gamma}^a_{bc}$ vanishes at the origin.

The preceding discussion shows that a coordinate system can be found for which Γ^a_{bc} is zero at one point. At the same point the metric can be transformed into a diagonal form with entries ± 1 or zero. To see this consider the transformation property of the metric

$$g_{a'b'} = \frac{\partial x^c}{\partial x^{a'}}\frac{\partial x^d}{\partial x^{b'}}g_{cd}.$$

As a matrix equation this can be written as

$$g' = JgJ^T,$$

where J is the Jacobian matrix with entries $J^c_{a'} = \partial x^c/\partial x^{a'}$; as in Chapter 2 the primed index denotes the row and the unprimed index denotes the column. For a linear transformation J is a constant matrix. For any point J can be chosen so that g' is diagonal and has entries ± 1 or zero. As g is symmetric it can be diagonalised by an orthogonal transformation. A further (in general non-orthogonal) transformation is required to bring g into the required form. For example, suppose that the metric has been brought to the diagonal form

$$g = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.$$

The transformation $g' = JgJ^T$ where

$$J = J^T = \begin{pmatrix} \frac{1}{\sqrt{|a|}} & 0 \\ 0 & \frac{1}{\sqrt{|b|}} \end{pmatrix} \qquad (a, b \neq 0)$$

yields a diagonal matrix with entries ± 1 only. Such a transformation does not change the sign of the eigenvalues. This is Sylvester's law of inertia. The number of zero, positive and negative eigenvalues is called he *signature* of the metric. In General Relativity metrics with one positive and three negative eigenvalues are relevant.

4.10 Geodesics and the Calculus of Variations

In this chapter geodesics have been introduced as generalisations of straight lines through

$$\frac{Du^a}{ds} = 0.$$

Here two alternative characterisations of geodesics are considered:

(a) As curves of minimum (or maximum) length joining two fixed points.

Consider a curve C joining two fixed points x_A and x_B. This can be defined through $x^a = x^a(\lambda)$ where $\lambda \in [a, b]$ is a parameter. The length of the curve can be expressed as an integral[12]

$$L = \int_a^b \sqrt{g_{cd} \frac{dx^c}{d\lambda} \frac{dx^d}{d\lambda}} d\lambda.$$

Geodesics may be defined as curves that minimise (or maximise) L for fixed end points x_A and x_B. The integral is awkward to work with due to the parametrisation ambiguity of the integral.

(b) As solutions of a mechanics problem.[13]

Consider a particle constrained to a surface with metric g_{ab}. Solve Newton's second law assuming no force apart from any constraint force (assuming the surface is embedded in a higher dimensional flat space, e.g. the unit sphere embedded in three-dimensional space). Here the particle moves along geodesics.

The kinetic energy of the particle

$$T = \frac{1}{2} m v^2 = \frac{1}{2} m g_{ab} \frac{dx^a}{dt} \frac{dx^b}{dt},$$

[12] In Special or General Relativity the 'length' is c times the proper time.

[13] This part which uses ideas from the Classical Dynamics module is not examinable.

is constant. As there is no external force the potential energy is zero. The equations of motion can be obtained through Lagrangian mechanics (see Problem Sheet 4). The Lagrangian is

$$L = T = \frac{1}{2} m g_{ab} \frac{dx^a}{dt}\frac{dx^b}{dt}.$$

Here the Euler-Lagrange equations are equivalent to the geodesic equation except that the distance s is replaced with t. The forms are equivalent as $s = vt$ and the speed v is a constant of the motion.

For example on a unit sphere $ds^2 = d\theta^2 + \sin^2\theta d\phi^2$. Taking $m = 1$

$$L = \frac{1}{2}\left(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2\right).$$

The Euler-Lagrange equations are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{\theta}}\right) - \frac{\partial L}{\partial\theta} = \frac{d}{dt}\dot{\theta} - \sin\theta\cos\theta\dot{\phi}^2 = 0.$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{\phi}}\right) - \frac{\partial L}{\partial\phi} = \frac{d}{dt}\left(\sin^2\theta\dot{\phi}\right) = 0.$$

The Christoffel symbols can be read off from the equations of motion. This is an alternative to using the explicit formula for Γ^a_{bc}.

We have seen that if the metric g_{ab} does not depend on x^p then $g_{pb}dx^b/ds$ is constant along geodesics. In the mechanics approach this is the result that if x^p is cyclic then the momentum $\partial L/\partial\dot{x}^p$ is a constant of the motion. The more general result regarding Killing vectors is Noether's theorem in the Lagrangian approach.

5 General Relativity

This chapter uses the ideas of the previous four chapters to develop General Relativity. The relationship between General Relativity and Newtonian gravity is discussed. Schwarzschild's metric and applications, including the deflection of light, are considered.

5.1 Special Relativity in Inertial and Non-inertial Frames

An inertial frame is four coordinates x^μ $(\mu = 0, 1, 2, 3)$ such that the proper time of any particle, τ, satisfies

$$c^2 d\tau^2 = ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu, \tag{50}$$

where $\eta_{\mu\nu} = \text{diag}(1, -1, -1, -1)$ is the metric. Newton's first law can be written in the form

$$\frac{d^2 x^\alpha}{d\tau^2} = 0, \tag{51}$$

which is the equation of motion of a free particle.

The conservation of energy and momentum can be formulated locally through

$$\partial_\nu T^{\mu\nu} = 0. \tag{52}$$

The components of the metric are constant

$$\partial_\alpha \eta_{\mu\nu} = 0. \tag{53}$$

Suppose that we start off with an inertial frame and then transform to another non-inertial set of coordinates. For example, we could transform from an inertial frame (t, x, y, z) to (t, r, θ, ϕ) where r,θ,ϕ are spherical polar coordinates. The metric can be written

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = c^2 dt^2 - dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta d\phi^2. \tag{54}$$

In the original inertial frame parallel transport is trivial and the Levi-Civita connection is zero. Therefore equations (51) and (52) can be written as

$$\frac{d^2 x^\alpha}{d\tau^2} + \Gamma^\alpha_{\beta\gamma} \frac{dx^\beta}{d\tau} \frac{dx^\gamma}{d\tau} = 0, \tag{55}$$

and

$$\nabla_\nu T^{\mu\nu} = 0, \tag{56}$$

respectively. As (55) and (56) transform as vectors they also hold on transforming to non-inertial coordinates. In addition, the trivial formula (53) becomes

$$\nabla_\alpha g_{\mu\nu} = 0. \tag{57}$$

In the original inertial frame the Riemann tensor is zero. It will remain zero under an arbitrary change of coordinates

$$R^\alpha_{\beta\gamma\delta} = 0. \tag{58}$$

The connection has the Levi-Civita form[14]

$$\Gamma^\alpha_{\beta\gamma} = \frac{1}{2} g^{\alpha\delta} \left(\partial_\beta g_{\delta\gamma} + \partial_\gamma g_{\beta\delta} - \partial_\delta g_{\beta\gamma} \right), \tag{59}$$

5.2 General Relativity

General Relativity comprises all the structure seen in the discussion of Special Relativity in non-inertial coordinates except that the zero curvature condition is dropped. In addition the form of the curvature is constrained by *Einstein's field equations*.

In General Relativity we start with a general space-time with metric

$$ds^2 = g_{\mu\nu}(x) dx^\mu dx^\nu.$$

Here $g_{\mu\nu}(x)$ is assumed to have one positive and three negative eigenvalues at each point in space-time. The connection is assumed to be torsion free and metric leading to the Levi-Civita connection (59). The motion of a massive particle is described by the geodesic equation (55). This can also be written as $a^\mu = 0$ where the acceleration is defined through

$$a^\mu = \frac{Du^\mu}{d\tau} = \frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\beta\gamma} \frac{dx^\beta}{d\tau} \frac{dx^\gamma}{d\tau}.$$

The density and flow of energy and momentum is described by the energy-momentum tensor $T^{\mu\nu}$. The local conservation of energy and momentum is encoded in the equation $\nabla_\nu T^{\mu\nu} = 0$.

In General Relativity, the Riemann tensor is not neccesarily zero and the equation of geodesic deviation applies

$$\frac{D^2 w^\alpha}{ds^2} = -R^\alpha_{\beta\gamma\delta} \frac{dx^\beta}{ds} w^\gamma \frac{dx^\delta}{ds}. \tag{60}$$

If the curvature is zero it is possible to transform to coordinates where the metric takes the Special Relativity form $\eta_{\mu\nu}$. If the curvature is non-zero it is possible via normal coordinates

[14] In this case it has to be the Levi-Civita connection as assuming trivial parallel transport in the original inertial frame implies the connection is torsion free in the transformed (non-inertial) coordinate system. Equation (57) means the connection is metric. As in Problem 8 of Problem Sheet 3 these conditions lead to the Levi-Civita form.

to eliminate the connection at a point and bring the metric to the Special Relativity form at the same point. At such a point $\partial_\nu T^{\mu\nu} = 0$ and a free particle satisfies $d^2x^\alpha/ds^2 = 0$. In other words Special Relativity holds locally.

The presence of curvature means that the geodesics are non-trivial. How is this related to gravity? In Newton's theory mass is the source of gravitational forces. In the light of the formula $E = mc^2$ it is natural to suspect that energy will play a key role in a relativistic theory of gravitation. In Special Relativity the energy density is a component of the energy momentum tensor. From Chapter 4 the Einstein tensor

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2}\mathcal{R}g^{\mu\nu},$$

has the property $\nabla_\nu G^{\mu\nu} = 0$ which resembles the local conservation law $\nabla_\nu T^{\mu\nu} = 0$. Einstein argued that they are proportional and proposed the field equations

$$G^{\mu\nu} = \frac{8\pi G}{c^4}T^{\mu\nu}. \tag{61}$$

The constant of proportionality includes Newton's Gravitational constant G. The precise form is chosen so that the geodesics resemble trajectories in Newtonian gravity.

Einstein's field equations are actually ten equations as the symmetric tensors $G^{\mu\nu}$ and $T^{\mu\nu}$ have ten independent components. In the presence of matter $T^{\mu\nu}$ and hence $G^{\mu\nu}$ is non-zero which in turn implies non-zero curvature. As we shall see the converse is not true; it is possible for the Riemann tensor to be non-zero even if $G^{\mu\nu}$ and $T^{\mu\nu}$ are zero.

5.3 Newtonian Gravity

It is far from obvious that General Relativity has anything to do with Newtonian gravity except for the appearance of the physical constant G in Einstein's field equations. We will demonstrate that for non-relativistic velocities and for spaces where the metric, $g_{\mu\nu}$, is 'close' to the flat space-metric, $\eta_{\mu\nu}$, that Newton's theory is an approximation to Einstein's theory.

In Newtonian gravity all matter is subject to an attractive force; for two particles of mass m_1 and m_2 the magnitude of the force is

$$\frac{Gm_1m_2}{r^2},$$

where r is the distance between the particles. Consider the force due to a number of particles; the force on a particle with mass m can be written in the form

$$\mathbf{F} = -m\nabla\phi,$$

where

$$\phi = -G\sum_{i=1}^{N}\frac{m_i}{|\mathbf{r}-\mathbf{r}_i|},$$

is the gravitational potential. Here $\mathbf{r}_i$ and m_i is the position and mass of the ith particle, respectively. Newton's second law states that the force on a particle of mass m is given by $\mathbf{F} = m\mathbf{a}$ where $\mathbf{a}$ is the acceleration vector. Accordingly, the acceleration of a particle is $\mathbf{a} = -\nabla\phi$ with ϕ as above. For a continuous distribution of matter

$$\phi(\mathbf{r}) = -G\int\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d^3r',$$

where $\rho(\mathbf{r})$ is the mass density. It follows from the above formula that[15] $\nabla^2\phi = 4\pi G\rho$ (Poisson's equation). Therefore Newtonian gravity can be formulated through two equations

$$\mathbf{a} = -\nabla\phi, \tag{62}$$

and

$$\nabla^2\phi = 4\pi G\rho. \tag{63}$$

Note that (62) implies that the acceleration of a particle due to gravitational forces is independent of the mass of the particle. This mass-independence is also a feature of General Relativity (mass does not enter the geodesic equation).

Tidal forces can be understood as a relative acceleration of neighbouring points. Let $\mathbf{w} = \mathbf{r}_2 - \mathbf{r}_1$ be the separation of two particles. Now

$$\ddot{w}^i = a^i(\mathbf{r}_2) - a^i(\mathbf{r}_1) = -\partial_i\phi(\mathbf{r}_2) + \partial_i\phi(\mathbf{r}_1),$$

where w^i is the ith component of $\mathbf{w}$. Using Taylor's theorem (expanding about $\mathbf{r} = \mathbf{r}_1$)

$$\partial_i\phi(\mathbf{r}_2) = \partial_i\phi(\mathbf{r}_1) + \partial_i\partial_j\phi(x_2^j - x_1^j) + \text{higher order terms.}$$

For small $\mathbf{w}$ this yields

$$\frac{d^2w^i}{dt^2} = -w^j\partial_i\partial_j\phi. \tag{64}$$

That is the matrix of second partial derivatives of the gravitational potential describes tidal forces in Newtonian gravity.

[15] By applying the divergence theorem to $\nabla\phi$ or using the Green's function $\nabla^2|\mathbf{r}-\mathbf{r}'|^{-1} = -4\pi\delta^3(\mathbf{r}-\mathbf{r}')$.

5.4 The Newtonian Limit of General Relativity

The acceleration formula (62) *is a limit of the geodesic equation*

Poisson's equation (63) *is a limit of Einstein's equation*

The tidal equation (64) *is a limit of the equation of geodesic deviation*

To justify these claims write the metric in the form

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\mu\nu}(x), \tag{65}$$

where $\eta_{\mu\nu}$ is the metric of Special Relativity and $h_{\mu\nu}(x)$ is a 'small' space-time dependent contribution. In the following analysis quadratic terms in $h_{\mu\nu}$ will be neglected.

Consider the geodesic equation

$$\frac{d^2x^\alpha}{ds^2} + \Gamma^\alpha_{\beta\gamma}\frac{dx^\beta}{ds}\frac{dx^\gamma}{ds} = 0,$$

where the velocities dx^i/dt are small compared to the speed of light. Then dx^0/ds is approximately 1 ($s = c\tau$ and in Special Relativity $dx^0/d\tau \approx c$ for non-relativistic speeds) and dx^i/ds is small by assumption. In this approximation the spatial components of the geodesic equation reduce to

$$\frac{d^2x^i}{ds^2} + \Gamma^i_{00} = 0,$$

or using $s = c\tau \approx ct$

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

Now

$$\Gamma^i_{00} = \frac{1}{2}g^{i\delta}\left(\partial_0 g_{\delta 0} + \partial_0 g_{0\delta} - \partial_\delta g_{00}\right) = \frac{1}{2}g^{i\delta}\left(\partial_0 h_{\delta 0} + \partial_0 h_{0\delta} - \partial_\delta h_{00}\right).$$

using (65). Replacing $g^{i\delta}$ with $\eta^{i\delta} = -1$ for $\delta = i$ gives

$$\Gamma^i_{00} = -\frac{1}{2}\left(2\partial_0 h_{i0} - \partial_i h_{00}\right).$$

Ignoring the time-dependence of $h_{\mu\nu}$ gives the approximation

$$\Gamma^i_{00} = \frac{1}{2}\partial_i h_{00}.$$

With the various approximations we obtain $a^i = -c^2\Gamma^i_{00} = -\frac{1}{2}c^2\partial_i h_{00}$ which is identical to the Newtonian acceleration formula (62) if

$$\phi = \frac{1}{2}c^2 h_{00}. \tag{66}$$

We now use Einstein's equation to show that ϕ satisfies the Poisson equation (63). Einstein's equations (61) can be rewritten in the form (see Problem Sheet 5)

$$R^{\mu\nu} = \frac{8\pi G}{c^4}\left(T^{\mu\nu} - \frac{1}{2}\mathcal{T} g^{\mu\nu}\right), \tag{67}$$

where $\mathcal{T} = g_{\mu\nu}T^{\mu\nu}$ is the 'trace' of the energy-momentum tensor.

Poisson's equation involves the mass density which for slowly moving matter should be proportional to the energy density (in view of $E = mc^2$). As the energy density is T^{00}, consider the 00 component of Einstein's equation

$$R^{00} = \frac{8\pi G}{c^2}\left(T^{00} - \frac{1}{2}\mathcal{T} g^{00}\right). \tag{68}$$

Now approximate $\mathcal{T} \approx T^{00}$ (trace dominated by the energy density). As in the discussion of the geodesic equation we use the approximation $g^{00} \approx \eta^{00} = 1$, so that

$$R^{00} \approx \frac{4\pi G}{c^4}T^{00} \approx \frac{4\pi G\rho}{c^2}, \tag{69}$$

using $T^{00} \approx \rho c^2$. Now

$$R^{00} \approx R_{00} = R^i_{0i0} = \partial_i \Gamma^i_{00} - \partial_0 \Gamma^i_{0i} + \text{quadratic terms.}$$

Ignore the quadratic terms as they are quadratic in $h_{\mu\nu}$ and neglect $\partial_0\Gamma^i_{0i}$, giving

$$R^{00} = \partial_i\Gamma^i_{00} = \frac{1}{2}\nabla^2 h_{00},$$

using $\Gamma^i_{00} = \frac{1}{2}\partial_i h_{00}$. Using $\phi = \frac{1}{2}c^2 h_{00}$, (68) reduces to Poisson's equation (63).

It can be shown that the Newtonian tidal formula (64) is a limit of the equation of geodesic deviation (60).

5.5 The Schwarzschild Metric

How does one solve Einstein's equations? Any space-time metric, $g_{\mu\nu}$, 'solves' the equations provided the energy-momentum tensor has the form $c^4 G^{\mu\nu}/8\pi G$. It is much more difficult to solve the equations when the form of the energy momentum tensor is restricted in some way. In fact, most known solutions of Einstein's equations have

$$T^{\mu\nu} = 0,$$

which is appropriate for empty space! Equation (67) implies that $R^{\mu\nu} = 0$ if $T^{\mu\nu} = 0$, in other words empty space-time is Ricci-flat. Note that this does not mean that the Riemann tensor is zero.

The first exact solution of this type was found by Schwarzschild shortly after the publication of Einstein's equations. The Schwarzschild metric is a simple modification of (54) which is the Special Relativity metric transformed to spherical polar coordinates. The Schwarzschild metric is[16]

$$ds^2 = c^2\left(1-\frac{R}{r}\right)dt^2 - \frac{dr^2}{1-\frac{R}{r}} - r^2(d\theta^2 + \sin^2\theta d\phi^2), \tag{70}$$

where R is a positive constant (the Schwarzschild radius). Setting $R = 0$ gives the flat-space metric (54). The Schwarzschild metric is Ricci-flat but the curvature is non-zero. The metric describes the space-time surrounding a massive spherically-symmetric body. [17] It also incorporates a black hole and a white hole. The Schwarzschild radius can be written as

$$R = \frac{2GM}{c^2},$$

where M is the 'mass' of the spherically symmetric object. This identification is justified by comparing the geodesics for $r \gg R$ with Newtonian orbits around a spherically-symmetric object of mass M.

The geodesics for the Schwarzschild metric resemble the orbits found in Newtonian mechanics. However, the non-trivial form of the geodesics is ascribed to the non-trivial geometry rather than the action of a gravitational force. Just to write down the four geodesic equations appears to be a non-trivial undertaking as there are 40 Christoffel symbols to calculate! Fortunately, the symmetries effectively reduce the number of geodesic equations from four to one.

The geodesic equation for r is

$$\ddot{r} + \frac{c^2R}{2r^2}\left(1-\frac{R}{r}\right)\dot{t}^2 - \left(1-\frac{R}{r}\right)^{-1}\frac{R}{2r^2}\dot{r}^2 - r\left(1-\frac{R}{r}\right)(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2) = 0, \tag{71}$$

where the dot refers to differentiation with respect to s. This can be obtained through a (tedious) calculation of the Christoffel symbols $\Gamma^r_{\beta\gamma}$; the non-zero components are

$$\Gamma^r_{tt} = \frac{c^2R}{2r^2}\left(1-\frac{R}{r}\right), \quad \Gamma^r_{rr} = -\left(1-\frac{R}{r}\right)^{-1}\frac{R}{2r^2},$$

$$\Gamma^r_{\theta\theta} = -r\left(1-\frac{R}{r}\right), \quad \Gamma^r_{\phi\phi} = -r\left(1-\frac{R}{r}\right)\sin^2\theta.$$

Alternatively, the geodesic equation (71) can be obtained using the the (non-examinable) Lagrangian approach outlined at the end of Chapter 4; a suitable Lagrangian is

[16] Here R is not the scalar curvature $\mathcal{R}$ (which is zero as the Ricci tensor is zero).

[17] Remarkably, the Schwarzschild metric is the unique spherically symmetric solution of $R^{\mu\nu} = 0$ (Birkhoff's theorem). This means that the space-time outside any spherically symmetric body (including a collapsing star) is described by the Schwarzschild metric.

$$L = \frac{1}{2} g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} = \frac{c^2}{2}\left(1 - \frac{R}{r}\right)\left(\frac{dt}{d\lambda}\right)^2 - \frac{1}{2\left(1 - \frac{R}{r}\right)}\left(\frac{dr}{d\lambda}\right)^2 - \frac{r^2}{2}\left[\left(\frac{d\theta}{d\lambda}\right)^2 + \sin^2\theta\left(\frac{d\phi}{d\lambda}\right)^2\right].$$

Here the 'time' used at the end of Chapter 4 is denoted λ to avoid confusion with the coordinate t in the Schwarzschild metric.

The geodesic equations for t and ϕ are much easier to obtain as the components of the metric do not depend on t and ϕ. The t-independence implies that

$$c^2 k = g_{t\beta} \frac{dx^\beta}{ds} = c^2\left(1 - \frac{R}{r}\right)\frac{dt}{ds},$$

is constant along geodesics. A similar argument for ϕ shows that

$$h = r^2 \sin^2\theta \frac{d\phi}{ds},$$

is constant along geodesics. The geodesic equation for θ is the same as for Special Relativity in spherical polar coordinates, i.e. R does not enter in to the equation. However, the equation can be bypassed by noting that it has the trivial solution $\theta = \pi/2$. This is just the statement that the geodesics are planar; every geodesic lies in a plane through the origin. Without loss of generality this can be taken to be the 'equatorial plane' $\theta = \pi/2$. In this plane $h = r^2 d\phi/ds$ which looks like the angular momentum. The geodesic equation for r in the equatorial plane is simpler. It follows from this equation that there are circular orbits (see Problem Sheet 5). Such orbits are unphysical for $r < \frac{3}{2}R$ as the proper time is undefined (or imaginary). Circular orbits with $r = \frac{3}{2}R$ represent the orbits of photons ($r = \frac{3}{2}R$ defines the 'photon sphere'). Circular orbits for $r > 3R$ are stable as in the Newtonian case (see the 2020 examination).

The geodesic equation for r can be directly integrated using

$$g_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} = c^2\left(1 - \frac{R}{r}\right)\left(\frac{dt}{ds}\right)^2 - \left(1 - \frac{R}{r}\right)^{-1}\left(\frac{dr}{ds}\right)^2 - r^2\left(\frac{d\phi}{ds}\right)^2 = 1, \tag{72}$$

which can be written in the form

$$c^2 k^2 \left(1 - \frac{R}{r}\right)^{-1} - \left(1 - \frac{R}{r}\right)^{-1}\left(\frac{dr}{ds}\right)^2 - \frac{h^2}{r^2} = 1.$$

Multiplying by $1 - R/r$ gives

$$\left(\frac{dr}{ds}\right)^2 + \frac{h^2}{r^2}\left(1 - \frac{R}{r}\right) - \frac{R}{r} = c^2 k^2 - 1 = \text{constant}. \tag{73}$$

One can integrate this using separation of variables to find r as a function of s, and then use the constants k and h to determine t and ϕ. Alternatively, it is possible to determine r as a function of ϕ by using the new coordinate

$$u = \frac{1}{r}.$$

Now

$$\frac{dr}{ds} = -\frac{1}{u^2}\frac{du}{ds} = -\frac{1}{u^2}\frac{d\phi}{ds}\frac{du}{d\phi} = -h\frac{du}{d\phi}.$$

Accordingly,

$$h^2\left(\frac{du}{d\phi}\right)^2 + h^2(u^2 - Ru^3) - Ru = \text{constant}.$$

Differentiating with respect to ϕ gives the orbit equation

$$h^2\frac{du}{d\phi}\left(2\frac{d^2u}{d\phi^2} + 2u - 3Ru^2\right) = R\frac{du}{d\phi},$$

so that $du/d\phi = 0$ (a circular orbit) or

$$\frac{d^2u}{d\phi^2} + u - \frac{3}{2}Ru^2 = \frac{R}{2h^2}. \tag{74}$$

The third term on the left hand side (74) is the 'General Relativity term' as dropping it gives the Newtonian orbit equation[18]

$$\frac{d^2u}{d\phi^2} + u = \frac{R}{2h^2}. \tag{75}$$

The general solution of this linear ODE is

$$\frac{1}{r} = u = A\cos(\phi + \beta) + \frac{R}{2h^2},$$

where A and β are arbitrary constants. An alternative form is

$$u = \frac{R}{2h^2}\left[1 + e\cos(\phi + \beta)\right],$$

where e is the eccentricity (the solutions are conic sections). Going back to the full relativistic orbit equation (74) there are no elliptical orbits (apart from circular ones). If uR is small the orbits are approximately elliptical but not closed. In this case the following approximation can be used

$$u = \frac{R}{2h^2}[1 + e\cos(\phi + \beta)] + u_1,$$

where

$$\frac{d^2u_1}{d\phi^2} + u_1 = \frac{3R}{2}u^2 \approx \frac{3R^3}{8h^4}[1 + e\cos(\phi + \beta)]^2.$$

Why are the solutions non-periodic?

[18] This can be derived from the conservation energy $\frac{1}{2}mv^2 + V = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2) - GMm/r = \frac{1}{2}m(\dot{r}^2 + h^2c^2/r - Rc^2/r)$ with $ch = r^2 d\phi/dt$ and $u = 1/r$.

5.6 Deflection of Light

Two early successes of General Relativity were the explanation of the anomalous perihelion shift of Mercury and the *prediction* of the deflection of light by the Sun. The dramatic verification of Einstein's prediction by Dyson and Eddington's twin expeditions to Principe and Sobral to observe an eclipse on 29 May 1919 was a key moment in the history of science.[19]

Here the deflection formula is obtained directly from the Schwarzschild metric. Before we do this a loose end from Chapter 4 is addressed. Following the line of argument in chapter 4 the geodesic equation takes the form (55) where the geodesics are parametrised using the proper time τ derived from the metric. This form of the geodesic equation does not work for photons as proper time is meaningless (or frozen). By using the parameter $\lambda = s/b$ where b is a positive constant (55) can be written in the form

$$\frac{d^2x^\alpha}{d\lambda^2} + \Gamma^\alpha_{\beta\gamma}\frac{dx^\beta}{d\lambda}\frac{dx^\gamma}{d\lambda} = 0, \tag{76}$$

where

$$g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = b^2. \tag{77}$$

Photon trajectories or *null geodesics* can be defined as the $b \to 0$ limit so that (76) holds and

$$g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = 0. \tag{78}$$

One can also define space-like geodesics by setting b^2 negative in (77). An example of a space-like geodesic is a circular obit in the Schwarzschild space-time with a radius less than $\frac{3}{2}R$. Note that space-like geodesics do not represent the trajectory of any known particles as they would be travelling faster than the speed of light.

Retracing the derivation of the orbit equation using (78) equation (72) becomes

$$g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = c^2\left(1-\frac{R}{r}\right)\left(\frac{dt}{d\lambda}\right)^2 - \left(1-\frac{R}{r}\right)^{-1}\left(\frac{dr}{d\lambda}\right)^2 - r^2\left(\frac{d\phi}{d\lambda}\right)^2 = 0.$$

Therefore

$$c^2\tilde{k}^2\left(1-\frac{R}{r}\right)^{-1} - \left(1-\frac{R}{r}\right)^{-1}\left(\frac{dr}{d\lambda}\right)^2 - \frac{\tilde{h}^2}{r^2} = 0,$$

where $\tilde{k} = (1-R/r)dt/d\lambda$ and $\tilde{h} = r^2 d\phi/d\lambda$ are constants. Much as in the derivation of the time-like orbit equation (74)

$$\frac{d^2u}{d\phi^2} + u = \frac{3}{2}Ru^2. \tag{79}$$

One can also understand this equation as the $h^2 \to \infty$ limit of (74).

To obtain Einstein's formula for the deflection of light treat the right hand side of (79) as

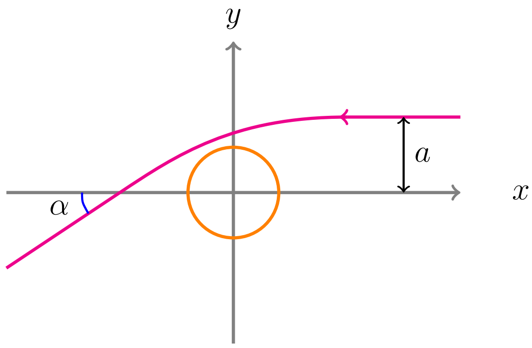
[19] For a modern account see 'No Shadow of Doubt The 1919 Eclipse That Confirmed Einstein's Theory of Relativity' by Daniel Kennefick, Princeton University Press, 2019.

a small quantity. Neglecting it all together a solution is

$$\frac{1}{r} = u = \frac{1}{a}\sin\phi,$$

or $r = a/\sin\phi$ which represents a straight line ($y = a$) which is a good approximation to a light ray! Inserting $u = a^{-1}\sin\phi$ into the right hand side of (79) yields

$$\frac{d^2u}{d\phi^2} + u = \frac{3}{2}Ru^2 = \frac{3R}{4a^2}(1 - \cos 2\phi).$$



The solution

$$u = \frac{R}{4a^2}\cos 2\phi + \frac{3R}{4a^2} + \frac{1}{a}\sin\phi - \frac{R}{a^2}\cos\phi,$$

satisfies the boundary conditions $u(\phi = 0) = 0$ and $du/d\phi(\phi = 0) = 1/a$. Consider u where ϕ is near π:

$$u = \frac{R}{4a^2} + \frac{3R}{4a^2} - \frac{1}{a}\sin(\phi - \pi) + \frac{R}{a^2}.$$

Look for a solution of $u(\phi) = 0$ near $\phi = \pi$:

$$\frac{1}{a}\sin(\phi - \pi) = \frac{2R}{a^2}.$$

The deflection angle $\alpha = \phi - \pi$ is

$$\alpha = \frac{2R}{a}. \tag{80}$$

This key prediction of General Relativity was first observed in the eclipse of 1919. This formula can also be used to calculate the focal length of a gravitational lens. What is the focal length of the Sun? The radius of the Sun is approximately 700000km and the Schwarzschild radius of the Sun is approximately 3km.

If the right hand side of (79) is not small one must solve the full (non-linear) problem. Here it is easier to work with the 'integrated' form

$$\left(\frac{du}{d\phi}\right)^2 + u^2 - Ru^3 = \text{constant}.$$

See Problem Sheet 5.

5.7 Falling into a Black Hole

The Schwarzschild metric is singular at $r = R$ as $g_{tt} \to 0$ and $g_{rr} \to -\infty$ as $r \to R^+$. To study time-like geodesics use (72). Note this also holds for non-geodesic trajectories. Suppose $r > R$ (the exterior of the black hole). Clearly

$$c^2 \left(1 - \frac{R}{r}\right) \left(\frac{dt}{ds}\right)^2 \geq 1,$$

so that dt/ds is never zero and dt/ds does not change sign. For positive energy particles dt/ds is positive. For $r < R$ the second term in (72) is positive so that

$$-\left(1 - \frac{R}{r}\right)^{-1} \left(\frac{dr}{ds}\right)^2 \geq 1$$

and dr/ds is never zero and doesn't change sign. For $dr/ds < 0$ the particle is falling towards $r = 0$ (the centre of the black hole).

Consider a particle initially at $r = R_1 > R$ which falls towards $r = R$. Then $t \to \infty$ as $t \to R^+$. To see this multiplying (72) by $(ds/dr)^2$

$$c^2 \left(1 - \frac{R}{r}\right) \left(\frac{dt}{dr}\right)^2 = \frac{1}{1 - \frac{R}{r}} + r^2 \left(\frac{d\phi}{dr}\right)^2 + \left(\frac{ds}{dr}\right)^2 \geq \frac{1}{1 - \frac{R}{r}},$$

or

$$c^2 \left(\frac{dt}{dr}\right)^2 \geq \frac{1}{(1 - \frac{R}{r})^2}, \qquad -c\frac{dt}{dr} \geq \frac{1}{1 - \frac{R}{r}}.$$

Integrating from $r = R_1$ to $r = R$ yields

$$c\Delta t \geq \int_R^{R_1} \frac{dr}{1 - \frac{R}{r}}$$

which is divergent. That is it takes an infinite amount of coordinate time to reach $r = R$. However the proper time is finite.

Now consider a geodesic for a particle falling towards $r = R$. For simplicity set $h = 0$ in (73) and take the constant on the right hand side of (73) to be zero

$$\left(\frac{dr}{ds}\right)^2 = \frac{R}{r},$$

or $dr/ds = -\sqrt{R}/\sqrt{r}$ which integrates to

$$R^{1/2}\Delta s = \frac{2}{3}\left(R_1^{3/2} - R^{3/2}\right).$$

As you fall towards the event horizon coordinate time fails ($t \to \infty$) but the proper time taken to reach the horizon is finite

The singularity at $r = R$ reflects 'bad' coordinates rather than a real singularity. Consider the Riemann tensor. It can be shown that

$$R^{\alpha}_{\beta\gamma\delta}R_{\alpha}^{\beta\gamma\delta} = \frac{12R^2}{r^6}.$$

The left hand side is a scalar which is independent of the choice of coordinates. This is non-singular at $r = R$. There is a genuine singularity at $r = 0$.

5.8 Kruskal Coordinates

Kruskal coordinates do not lead to an apparent singularity at $r = R$. They are defined by replacing r and t with new coordinates X and T:

$$X = \left(\frac{r}{R} - 1\right)^{\frac{1}{2}} e^{\frac{1}{2}r/R}\cosh(\tfrac{1}{2}ct/R)$$

$$T = \left(\frac{r}{R} - 1\right)^{\frac{1}{2}} e^{\frac{1}{2}r/R}\sinh(\tfrac{1}{2}ct/R),$$

for $r > R$, and

$$X = \left(1 - \frac{r}{R}\right)^{\frac{1}{2}} e^{\frac{1}{2}r/R}\sinh(\tfrac{1}{2}ct/R)$$

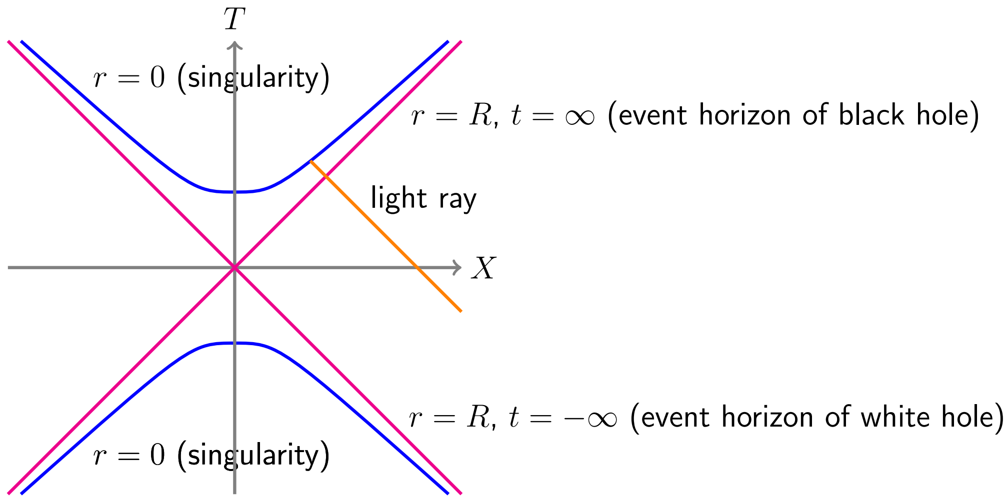
$$T = \left(1 - \frac{r}{R}\right)^{\frac{1}{2}} e^{\frac{1}{2}r/R}\cosh(\tfrac{1}{2}ct/R),$$

for $r < R$. It follows that (see Problem Sheet 5) that

$$c^2\left(1 - \frac{R}{r}\right)dt^2 - \frac{dr^2}{1 - \frac{R}{r}} = \frac{4R^3}{r}e^{-r/R}(dT^2 - dX^2),$$

ignoring the ϕ coordinate. This form of the metric is non-singular at $r = R$. The Kruskal coordinates defined above are defined separately for $r > R$ (the 'universe' outside the black hole) and $r < R$ (inside the black hole). In fact, X and T do not cover the whole plane and two further regions can be added corresponding to the interior of a white hole and 'another universe'. In the XT plane null geodesics are straight lines at 45 degrees to the axes.

Kruskal coordinates are non-singular at the event horizon. They also make clear that the universe is connected to 'another universe'. In fact, this striking feature of the Schwarzschild



spacetime was discovered by Einstein and Rosen before Kruskal coordinates were invented. Einstein and Rosen considered the simple change of variables

$$u^2 = r - R.$$

Allowing u to be both positive and negative describes a space time with two universes joined by an 'Einstein-Rosen bridge' which was the first example of a wormhole (see Problem Sheet 5). It is important to note that the bridge is non-traversable in that no time-like curve joins the two universes.

5.9 Linearised Gravity and Gravitational Waves (not examinable)

In the discussion of the Newtonian limit we considered a space-time 'close' to flat space-time through (65). We also ignored time-dependence and disregarded all components of $h_{\mu\nu}$ except h_{00} (which was the identified with the gravitational potential in Newtonian gravity). Here we consider 'weak' gravity through $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ with $h_{\mu\nu}$ small without the further assumptions used to extract the Newtonian limit.

Neglecting terms quadratic and higher in $h_{\mu\nu}$ it is straightforward to obtain the approximation

$$\Gamma_{\alpha\beta\gamma} = \frac{1}{2}\left(\partial_\beta h_{\alpha\gamma} + \partial_\gamma h_{\alpha\beta} - \partial_\alpha h_{\beta\gamma}\right),$$

which gives

$$R_{\alpha\beta\gamma\delta} = \partial_\gamma \Gamma_{\alpha\beta\delta} - \partial_\delta \Gamma_{\alpha\beta\gamma} = \frac{1}{2}\left(\partial_\gamma\partial_\beta h_{\alpha\delta} - \partial_\gamma\partial_\alpha h_{\beta\delta} - \partial_\delta\partial_\beta h_{\alpha\gamma} + \partial_\delta\partial_\alpha h_{\beta\gamma}\right).$$

The Ricci tensor in this approximation is

$$R_{\beta\delta} = \eta^{\alpha\gamma} R_{\alpha\beta\gamma\delta} = \frac{1}{2}\left(\partial_\beta\partial^\alpha h_{\alpha\delta} - \Box h_{\beta\delta} - \partial_\delta\partial_\beta h + \partial_\delta\partial^\gamma h_{\beta\gamma}\right).$$

Here $\Box = \partial^\mu \partial_\mu$ or 'box' and $h = \eta^{\mu\nu} h_{\mu\nu}$. This form of the Ricci tensor is complicated. However, it is possible to change the $h_{\mu\nu}$ through a change of coordinates in a way that provides a more useful formula for $R_{\mu\nu}$. Consider the *de Donder gauge condition*[20]

$$\partial^\mu h_{\mu\nu} = \frac{1}{2} \partial_\nu h. \tag{81}$$

If (81) holds $R_{\beta\delta} = -\frac{1}{2} \Box h_{\beta\delta}$ so that in empty space

$$\Box h_{\mu\nu} = 0, \tag{82}$$

a wave equation!

To justify (81) it is useful to consider a general transformation whereby the coordinates x^μ are replaced with $x^\mu - \xi^\mu(x)$. Here the ξ^μ are four arbitrary functions of the original x^μ coordinates which are treated as small in a linearised approximation. Under such a replacement

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\nu, \tag{83}$$

where $\xi_\mu = \eta_{\mu\alpha} \xi^\alpha$. This is called a *gauge transformation*. One now shows that there exist ξ^μ (or a gauge transformation) such that (81) holds. Assume that it does not hold then write $\partial^\mu h_{\mu\nu} - \frac{1}{2} \partial_\nu h = f_\nu$. Under a gauge transformation this becomes $\partial^\mu h_{\mu\nu} - \frac{1}{2} \partial_\nu h \to f_\nu + \Box \xi_\nu$. Requiring $\Box \xi_\nu = -f_\nu$, a linear wave equation, provides the required gauge transformation to achieve (81).

Consider a plane wave solution to (82)

$$h_{\mu\nu} = a_{\mu\nu} \cos(k \cdot x), \tag{84}$$

where $a_{\mu\nu}$ is constant and symmetric and k is a null four-vector. $a_{\mu\nu}$ has 10 independent components. However, the gauge freedom can be used to fix $h = 0$ (that is $h_{\mu\nu}$ traceless) and $h_{i0} = h_{0i} = 0$ for $i = 1, 2, 3$. Four of the remaining 6 independent constants are fixed by the de Donder condition (81) which reduces to $k^\mu a_{\mu\nu} = 0$. Suppose $k = (\kappa, 0, 0, \kappa)$ representing a plane wave in the positive x^3-direction. The above equation fixes a_{00}, a_{33}, $a_{13} = a_{31}$, $a_{23} = a_{32}$ to be zero. $a_{\mu\nu}$ can be written in the matrix form

$$a = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & a_+ & a_\times & 0 \\ 0 & a_\times & -a_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

[20] Also known as the harmonic gauge. This is similar to the Lorenz gauge in electrodynamics.

For $a_\times = 0$ the metric is

$$ds^2 = c^2 dt^2 - (1 - a_+ \cos(c\kappa t - z))dx^2 - (1 + a_+ \cos(c\kappa t - z))dy^2 - dz^2.$$

representing a periodic stretching and squeezing in the x and y directions. As x is squeezed y is stretched.

6 Geometrical Formulation of Vectors and Tensors (not examinable)

Throughout the module vectors and tensors have been developed through their transformation properties under special and general changes of coordinates. This follows the historical path where Special Relativity was developed from the Lorentz transformation considered in chapter 1. In this short chapter vectors and tensors are developed as geometrical objects. For a detailed development of General Relativity starting from differentiable manifolds see 'The Large Scale Structure of Space-time' by S. W. Hawking and G. F. R. Ellis, Cambridge University press (1973)[21].

In chapter 3 contravariant vector fields on a N-dimensional space were defined as N functions $v^a(x)$ with the transformation property

$$v^{a'} = \frac{\partial x^{a'}}{\partial x^b} v^b,$$

under a change of coordinates from N 'old' coordinates x^a $(a = 1, ..., N)$ to N 'new' coordinates $x^{a'}$. The v^a transform in the same way as the differentials dx^a. In elementary geometry a vector is introduced as an object with a magnitude and direction. Consider a vector field as

$$v(x) = v^a(x) e_a(x),$$

where the $e_a(x)$ are N basis vectors $(a = 1, ..., N)$ and the summation convention is used. Here the basis vectors are position dependent. Under a change of basis vectors the components v^a change but the vector, $v(x)$, is unchanged.

In order to define basis vectors, a general definition of vectors is required. To motivate this, consider the directional derivative in vector calculus. Let $\phi(\mathbf{r})$ be a scalar field (a mapping from $\mathbb{R}^3$ to $\mathbb{R}$). Consider the directional derivative of ϕ in the direction $\mathbf{V}$

$$D_{\mathbf{V}(\mathbf{r})}\phi = \lim_{\epsilon \to 0} \frac{\phi(\mathbf{r} + \epsilon \mathbf{V}(\mathbf{r})) - \phi(\mathbf{r})}{\epsilon}.$$

Here $\mathbf{V}$ depends on $\mathbf{r}$ (it is a vector field). The directional derivative can also be written in the form

$$D_{\mathbf{V}(\mathbf{r})}\phi = \mathbf{V} \cdot \nabla \phi = V^x \frac{\partial \phi}{\partial x} + V^y \frac{\partial \phi}{\partial y} + V^z \frac{\partial \phi}{\partial z}.$$

The directional derivative, $D_{\mathbf{V}}$, has the properties

$$D_{\mathbf{V}+\mathbf{W}} = D_{\mathbf{V}} + D_{\mathbf{W}},$$

[21] Available online from the library.

that is linearity, and the Leibniz property

$$D_{\mathbf{V}}(\phi_1\phi_2) = (D_{\mathbf{V}}\phi_1)\phi_2 + \phi_1(D_{\mathbf{V}}\phi_2).$$

Using $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ to denote unit vectors in the x, y, z directions $D_{\mathbf{i}}\phi = \partial\phi/\partial x$. In this way the vector $\mathbf{i}$ can be identified with the operator $\partial/\partial x$. This motivates the following definition: a vector field v is a linear mapping assigning a function $v\phi$ to every function ϕ. This linear mapping is assumed to have the Leibniz property

$$v(\phi_1\phi_2) = (v\phi_1)\phi_2 + \phi_1(v\phi_2).$$

For example

$$v = x\frac{\partial}{\partial x} + y^2\frac{\partial}{\partial z},$$

has the Leibniz property. It is also the directional derivative $D_{\mathbf{V}}$ in the direction $\mathbf{V} = x\mathbf{i}+y^2\mathbf{k}$.

For coordinates x^a $(a = 1, ..., N)$ it is natural to use the basis vectors

$$e_a = \frac{\partial}{\partial x^a}.$$

A general vector can be written in the form

$$v(x) = v^a(x)\frac{\partial}{\partial x^a}.$$

Under a change of coordinates $v(x)$ is unchanged

$$v^a(x) = v^a(x)\frac{\partial}{\partial x^a} = v^{a'}(x)\frac{\partial}{\partial x^{a'}}.$$

For example, on a sphere with coordinates θ, ϕ

$$v = \sin\theta\frac{\partial}{\partial\phi},$$

is a vector field with components $v^\theta = 0$ and $v^\phi = \sin\theta$.

In chapter 3 covariant vector fields $v_a(x)$ were defined through the transformation rule

$$v_{a'} = \frac{\partial x^b}{\partial x^{a'}}v_b.$$

These are components of one-forms

$$\omega = v_a e^a.$$

One-forms (linearly) map vector fields to scalar fields $\omega(v)$ is a scalar field (or function). The

e^a are dual basis vectors in that

$$e^a(e_b) = \delta^a_b.$$

For coordinates x^a $(a = 1, ..., N)$ a natural choice is $e_a = \partial/\partial x^a$. Here the dual basis vectors are 'differentials'

$$e^a = dx^a.$$

That is

$$dx^a \left(\frac{\partial}{\partial x^b} \right) = \delta^a_b,$$

defines the differentials. They are not 'small' or 'infinitesimal'.

The metric which maps pairs of vector fields to a scalar field

$$g(u, v) = g(v, u),$$

is bilinear

$$g(u + v, w) = g(u, w) + g(v, w).$$

The metric generalises the idea of the scalar product.

Components can be defined through

$$g_{ab} = g(e_a, e_b).$$

Using the $e_a = \partial/\partial x^a$

$$g = g_{ab}(x) dx^a \otimes dx^b.$$

Covariant derivatives are defined as a generalisation of the directional derivative and written ∇_v (the derivative in the direction v). For a scalar

$$\nabla_v \phi = v\phi,$$

since vector fields have been defined as a directional derivative of a scalar. How is a vector differentiated? To define $\nabla_v w$ the operation is defined for basis vectors:

$$\nabla_{e_c} e_b = \Gamma^a_{bc} e_a.$$

Here ∇ is called a connection. Using the shorthand notation

$$\nabla_c = \nabla_{e_c},$$

$$\nabla_c e_b = \Gamma^a_{bc} e_a.$$

We can differentiate a general vector field using the Leibniz rule

$$\nabla_c v = \nabla_c(v^b e_b) = (\nabla_c v^b)e_b + v^b(\nabla_c e_b)$$

$$= (\partial_c v^b)e_b + v^b \Gamma^a_{bc} e_a = \left(\partial_c v^a + \Gamma^a_{bc} v^b\right) e_a,$$

using the standard basis $e_a = \partial/\partial x^a$. This is in line with the definition $\nabla_c v^a = \partial_c v^a + \Gamma^a_{bc} v^b$ used in chapter 3. However, it is important to note that the second term comes from differentiating the basis vectors. Using the definitions of this chapter $\nabla_c v^a$ is just $\partial_c v^a$ as v^a is really a function or scalar.

Curvature is defined through

$$R(u, v)w = \nabla_u \nabla_v w - \nabla_v \nabla_u w - \nabla_{[u,v]} w.$$

Here the vector field $[u, v]$ is the commutator of the vector fields u and v. Using basis vectors

$$R(e_c, e_d)e_b = R^a_{bcd} e_a.$$

With the standard basis vectors $[e_c, e_d] = 0$.