GRAVITATION

J.PEARSON

JULY 20, 2009

Abstract

These are a set of notes I have made, based on lectures given by A.Pilaftsis at the University of Manchester Sept-Dec '08. Please e-mail me with any comments/corrections: jon@jpoffline.com. These notes may be found at www.jpoffline.com.

Contents

1 Recap of Special Relativity

Let us quickly recap the principles of special relativity that are assumed to be known.

The postulates of SR:

- All laws of nature are the same for all inertial observers;
- The speed of light, c, is the same for all inertial observers.

1.1 The Lorentz Transformations

Consider a frame Σ', within which an observer is stationary. The coordinates in that frame are the "primed ones", (ct', x', y', z'). Now, consider another frame, Σ, such that Σ' is moving at constant velocity $\beta \equiv v/c$ relative to a stationary observer in Σ. The coordinates in the "stationary frame" are unprimed (ct, x, y, z).

The two sets of coordinates are related via the transformations

$$ct' = \gamma(ct - \beta x), \quad x' = \gamma(c - \beta ct), \quad y' = y, \quad z' = z. \tag{1.1}$$

We have defined the quantities

$$\gamma \equiv \frac{1}{\sqrt{1-\beta^2}}, \quad \beta \equiv \frac{v}{c}.$$

From the transformations, we can compute "the invariance of the interval", thus

$$c^2t'^2 - x'^2 - y'^2 - z'^2 = ct^2 - x^2 - y^2 - z^2.$$

The physical consequences of this is that of Fitzgerald contraction (moving bodies shorten), time dilation (moving clocks run slow).

1.2 Covariant Formalism

The title "covariant formalism" is a little misleading: it should read "invariant formalism", but convention leaves it so.

Let us define the *contravariant* position 4-vector as

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z). \tag{1.2}$$

The metric of SR is flat, called the Minkowski metric, and written $\eta_{\mu\nu}$. The elements of the metric may be represented as

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

Notice that this metric is symmetric; $\eta_{\mu\nu} = \eta_{\nu\mu}$. Consider constructing an inverse matrix to this metric. That is, we require

$$\eta\eta^{-1} = \mathbf{1}_4,$$

where $\mathbf{1}_4$ is the 4-D identity matrix diag(1,1,1,1). Inspection will see that the inverse matrix has the same elements as the original. We denote the inverse of the metric as

$$(\eta^{-1})_{\mu\nu} \equiv \eta^{\mu\nu},$$

thus, we have that

$$\eta_{\mu\nu}\eta^{\nu\lambda} = \delta^{\lambda}_{\mu}.$$

Now, in Euclidean space, suppose we have a vector $\mathbf{x} = x^i \mathbf{e}_i$, where $\mathbf{e}_i$ is a basis vector and $i \in [1, n]$, where n is the dimension of the Euclidean space (usually 3). Then, the dot-product of the vector with itself can be written as

$$\mathbf{x} \cdot \mathbf{x} = x^i x^j \mathbf{e}_i \cdot \mathbf{e}_j,$$

and we "mix the basis vectors" via the Kronecker-delta, which is the metric of Euclidean space

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} \quad \Rightarrow \quad \mathbf{x} \cdot \mathbf{x} = x^i x^j \delta_{ij} = x^i x_i.$$

If we expand out this implied summation, we get the radius of a sphere in the n-dimensional Euclidean space

$$x^i x_i = x^2 + y^2 + z^2.$$

Now, we make the analogy to Minkowski space. We denote a contravariant vector as $\mathsf{x} = x^\mu \mathsf{e}_\mu$, so that the inner-product of the vector with itself is written

$$\mathsf{x} \cdot \mathsf{x} = x^\mu x^\nu \mathsf{e}_\mu \cdot \mathsf{e}_\nu,$$

and again we mix the basis vectors by the metric of the space; the metric of Minkowski space is $\eta_{\mu\nu}$. Thus,

$$\mathsf{e}_\mu \cdot \mathsf{e}_\nu = \eta_{\mu\nu},$$

and therefore

$$\mathsf{x} \cdot \mathsf{x} = x^\mu x^\nu \eta_{\mu\nu}.$$

If we say that

$$x_\mu = \eta_{\mu\nu} x^\nu, \tag{1.3}$$

then we see that

$$\mathsf{x} \cdot \mathsf{x} = x^\mu x_\mu.$$

From this, we are able to define the *covariant* position 4-vector as

$$x_\mu = \eta_{\mu\nu} x^\nu = (ct, -x, -y, -z).$$

And therefore, carrying out the summation, we find that the inner-product of the position 4-vector with itself is the radius of a 4-D sphere in Minkowski space;

$$x^\mu x_\mu = (ct)^2 - x^2 - y^2 - z^2.$$

Of course, we can write the inner-product of one 4-vector with another

$$\mathsf{x} \cdot \mathsf{y} = x^\mu y^\nu \eta_{\mu\nu} = x^\mu y_\nu.$$

Just as we used the metric to lower a contravariant vectors index, to become a covariant index, we may use the inverse metric to raise a covariant index to become a contravariant one

$$x^\mu = \eta^{\mu\nu} x_\nu. \tag{1.4}$$

Therefore, using these relations, we are able to see that

$$x_\nu y^\nu = x^\nu y_\nu.$$

1.2.1 Lorentz Boost

Consider again the 4-vector $\mathsf{x} = x^\mu \mathsf{e}_\mu$. Then, consider that the vector is the same in another frame, then we must have that

$$x^\mu \mathsf{e}_\mu = x'^\mu \mathsf{e}'_\mu.$$

The way we transform between frames is via a Lorentz boost;

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu, \tag{1.5}$$

where we use

$$(\Lambda^\mu{}_\nu) = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

If we note all of our definitions used thus far (for contravariant vectors, and their components), and the expressions forming the Lorentz transformations, (1.1), we see that

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

We say that $\Lambda^\mu{}_\nu$ (as defined above) constitutes a boost along the x-axis. It is infact a rotation about the $y - z$-plane.

Hence, we have a rule for transforming contravariant component, between frames: (1.5). Then, how does a covariant component transform?

Consider using the metric to change from a contravariant vector to a covariant one, in the primed frame,

$$x'_\mu = \eta_{\mu\kappa} x'^\kappa,$$

then we use (1.5) to transform the contravariant vector on the RHS

$$\eta_{\mu\kappa} x'^\kappa = \eta_{\mu\kappa} \Lambda^\kappa{}_\lambda x^\lambda,$$

then lower the index on the RHS

$$\eta_{\mu\kappa} \Lambda^\kappa{}_\lambda x^\lambda = \eta_{\mu\kappa} \Lambda^\kappa{}_\lambda \eta^{\lambda\nu} x_\nu.$$

Although not previously stated, we can imagine that the metric can lower/raise indices on anything, not just position vector-components. Thus, we see that $\eta_{\mu\kappa}\Lambda^\kappa{}_\lambda = \Lambda_{\mu\lambda}$. Hence, the above reads

$$\eta_{\mu\kappa} \Lambda^\kappa{}_\lambda \eta^{\lambda\nu} x_\nu = \Lambda_\mu{}^\nu x_\nu.$$

Now, let us define the inverse Lorentz transform as

$$(\Lambda^{-1})^\nu{}_\mu \equiv \Lambda_\mu{}^\nu.$$

Therefore, writing this stream of algebra down, from start to finish, we arrive at our result

$$\begin{aligned} x'_\mu &= \eta_{\mu\kappa} x'^\kappa \\ &= \eta_{\mu\kappa} \Lambda^\kappa{}_\lambda x^\lambda \\ &= \eta_{\mu\kappa} \Lambda^\kappa{}_\lambda \eta^{\lambda\nu} x_\nu \\ &= \Lambda_\mu{}^\nu x_\nu \\ &= (\Lambda^{-1})^\nu{}_\mu x_\nu. \end{aligned}$$

That is, to find the covariant components of a vector in the primed frame, we relate them to the unprimed frame via the inverse Lorentz transformation

$$x'_\mu = (\Lambda^{-1})^\nu{}_\mu x_\nu. \tag{1.6}$$

Let us then right our two Lorentz transformation rules; one for contravariant components & one for covariant

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu, \quad x'_\mu = (\Lambda^{-1})^\nu{}_\mu x_\nu. \tag{1.7}$$

Notice that the inverse Lorentz transformation matrix may be written as

$$((\Lambda^{-1})^\nu{}_\mu) = \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (\Lambda^{-1})^\nu{}_\mu = \frac{\partial x^\nu}{\partial x'^\mu}.$$

Notice that the product of $\Lambda^{\mu}{}_{\nu}$ and $(\Lambda^{-1})^{\nu}{}_{\mu}$ is the identity matrix, as they are inverses

$$\Lambda^{\nu}{}_{\lambda}(\Lambda^{-1})^{\mu}{}_{\nu} = \delta^{\mu}_{\lambda}.$$

We are now in a position to be able to prove the invariance of the interval, in Minkowski space, under Lorentz transformations. Consider the inner-product of two vectors in the primed frame,

$$\mathsf{x}' \cdot \mathsf{y}' = x'^{\mu} y'_{\mu},$$

we then transform each expression on the RHS, according to the relevant rule

$$x'^{\mu} y'_{\mu} = \Lambda^{\mu}{}_{\nu}(\Lambda^{-1})^{\lambda}{}_{\mu} x^{\nu} y_{\lambda},$$

then, noting the relation between the transformation & its inverse,

$$\Lambda^{\mu}{}_{\nu}(\Lambda^{-1})^{\lambda}{}_{\mu} x^{\nu} y_{\lambda} = \delta^{\lambda}_{\nu} x^{\nu} y_{\lambda},$$

which easily gives

$$\delta^{\lambda}_{\nu} x^{\nu} y_{\lambda} = x^{\nu} y_{\nu}.$$

And therefore, putting it all together

$$\begin{aligned} x'^{\mu} y'_{\mu} &= \Lambda^{\mu}{}_{\nu}(\Lambda^{-1})^{\lambda}{}_{\mu} x^{\nu} y_{\lambda} \\ &= \delta^{\lambda}_{\nu} x^{\nu} y_{\lambda} \\ &= x^{\nu} y_{\nu}. \end{aligned}$$

And thus, we have shown that the inner-product is invariant under Lorentz transformation (the invariance of the interval).

1.3 Standard Relations

Here we shall merely state the standard definitions of various 4-vectors.

The infinitesimal 4-position is defined as

$$dx^{\mu} = (cdt, \mathbf{x}), \quad \Rightarrow \quad dx_{\mu} = \eta_{\mu\nu} dx^{\nu} = (cdt, -\mathbf{x}).$$

The line element:

$$ds^2 = \eta_{\mu\nu} dx^{\mu} dx^{\nu} = c^2 dt^2 - (d\mathbf{x})^2.$$

Proper time:

$$d\tau = \frac{1}{c}\sqrt{dx^{\nu} dx_{\nu}} = \frac{dt}{\gamma}.$$

4-velocity:

$$u^{\mu} = \frac{dx^{\mu}}{d\tau} = (c\gamma, \gamma\mathbf{v}).$$

4-momentum:

$$p^\mu = m u^\mu = (E/c, \mathbf{p}).$$

Differential operator:

$$\partial_\mu \equiv \left(\frac{1}{c}\frac{\partial}{\partial t}, \nabla\right).$$

Charge conservation:

$$\partial_\mu J^\mu = 0, \quad J^\mu = (c\rho, \mathbf{J}).$$

Lorentz gauge:

$$\partial_\mu A^\mu = 0, \quad A^\mu = (\phi/c, \mathbf{A}).$$

1.4 The Equivalence Principles

Here we shall discuss some thought experiments which lead to the development of general relativity.

1.4.1 The Weak Equivalence Principle

Imagine an observer and "ball" inside a sealed lift. The observer is stationary relative to the ball, and are unable to see out of the lift. Suppose that the lift is suspended above a homogeneous gravitational field.

Then, suppose that the cable holding the lift up, is cut. The lift will accelerate downwards, $\boldsymbol{a} = \boldsymbol{g}$; where the acceleration due to gravity is just given by

$$\mathbf{g} = -\nabla\phi_g.$$

Now, experience tells us that both the observer and ball will remain at rest, relative to each other, inside the lift.

From Newton, we have the relation between the resultant force on a body (which will be the gravitational mass times the gravitational field), and the inertial mass with acceleration:

$$m_i \mathbf{a} = m_g \mathbf{g}.$$

Thus, as $\mathbf{a} = \mathbf{g}$, we therefore easily see that $m_i = m_g$. This leads to the statement of the weak equivalence principle:

"*Gravity couples in the same way to all mass & energy*".

1.4.2 The Strong Equivalence Principle

Consider the same setup as before: observer & ball at rest inside a sealed lift. This time, let the lift be in free space (i.e. no gravitational fields, anywhere).

Then, suppose that we accelerate the lift (using a rocket) such that $\mathbf{a} = \mathbf{g}$. We see that there is no difference in this situation as to one in which the lift is sat on the earths surface. Thus, the string equivalence principle:

"All laws of physics are the same in an accelerated frame, and in a uniform static gravitational field".

1.5 Gravitational Redshift

Consider a lift with a stationary observer in. Also in the lift, is a light blub, which emits light at frequency ν', according to the observer stationary inside the lift. Now, consider that there is another observer, stationary, on the surface of the earth (which we model as having a homogenous gravitational field). We have $\hat{\mathbf{x}}$ pointing upwards, from the surface of the earth. Then,

$$\mathbf{g} = -\frac{d\phi}{d\ell}\hat{\mathbf{x}}.$$

Let the length of the lift be $d\ell$, and the light bulb reside at the top of the lift. Then, a signal traveling at speed c takes time $dt = cd\ell$ to traverse the length of the lift.

Now, suppose that the lift is traveling at speed v, and then the observer on the earth will see some shifted frequency, ν. The Doppler shift is just

$$\frac{\nu'}{\nu} = \left(\frac{1 + v/c}{1 - v/c}\right)^{1/2} \approx 1 + \frac{v}{c},$$

after using the binomial expansion. From this, we see that

$$\frac{d\nu}{\nu} = \frac{v}{c}.$$

Using the relation that $v = du = gdt$, this simply gives that

$$\frac{d\nu}{\nu} = \frac{g}{c}dt,$$

which gives, using $cdt = d\ell$

$$\frac{d\nu}{\nu} = \frac{g}{c^2}d\ell.$$

Now, if we use the fact that $gd\ell = -d\phi$, then this is just

$$\frac{d\nu}{\nu} = -\frac{d\phi}{c^2}.$$

Therefore, we see that frequency shift is due to a changing gravitational potential. Thus, if a photon is moving out of a potential, then it will be red-shifted; and inward would be blue-shifted.

1.6 Einstein's Vision of General Relativity

Einstein's vision is that spacetime is a manifold, such that line elements are given by

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

where the metric is a function of coordinates. Within the metric (or, how the metric is constructed) is information on how spacetime is curved; and it is curved by any form of energy/momentum. According to the equivalence principle, one can always choose coordinates such that space is locally flat (Minkowski). Things in the spacetime travel along straight geodesics. Massive particles travel along time-like geodesics, which have $ds^2 > 0$, photons travel along null geodesics $ds^2 = 0$, and tachyons along $ds^2 < 0$.

2 Manifolds, Metrics & Tensors

2.1 Definitions

Let us state some rather (mathematically) loose definitions.

Manifold A manifold is a continuous set of points, which locally looks like an n-dimensional Minkowski space.

That is, given a manifold $\mathcal{M}$, if we "zoom in" on a little bit, that little bit will look flat. Suppose we zoom in on a bit which we label $u_i(\mathbf{p})$, where i just means that we chose one of many bits; and $\mathbf{p}$ is the point at the middle of the bit u_i. The coordinate system in u_1 (say) is Minkowski, $x^a(\mathbf{p})$. The whole collection of these little bits leads us to our next definition.

A manifold endowed with a metric is called a *Riemannian manifold*

Atlas An atlas is the complete set of coordinate systems $\{u_i\}$ in the manifold $\mathcal{M}$.

Curve A curve, in an n-manifold (where $\mathcal{M}$ merely has n coordinates), is a subset of points defined parametrically

$$x^a = x^a(\lambda), \quad a = 1, 2, \ldots, n, \quad \lambda \in \mathbb{R}.$$

For example, consider a 1-sphere (i.e. a circle), defined by the equation $x^2 + y^2 = 1$. We parameterise it thus

$$x^a = (x(\lambda), y(\lambda)) \quad \Rightarrow \quad x(\lambda) = \sin\lambda, y(\lambda) = \cos\lambda; \quad 0 \leq \lambda < 2\pi.$$

Surfaces A m-dim hypersurface in an n-manifold (whereby $m < n$), is defined as

$$x^a = x^a(\lambda_1, \ldots, \lambda_m); \quad \lambda_{1,\ldots,m} \in \mathbb{R}.$$

So that a curve is a 1D hypersurface. Or, alternatively, a surface is a generalisation of a curve.

For example, consider a 2-sphere (i.e. the surface of a ball), of constant radius r. It is defined by $x^2 + y^2 + z^2 = r^2 = \text{const}$. We parameterise the surface by (θ, ϕ), so that

$$x = r\sin\theta\cos\phi, \quad y = r\sin\theta\sin\phi, \quad z = r\cos\theta; \quad 0 \leq \theta < \pi, 0 \leq \phi < 2\pi.$$

2.2 Coordinate Transformations

Consider moving from one coordinate system to another

$$x^\mu \longmapsto x'^\mu = x'^\mu(x^\nu).$$

Such a transformation is defined by displacement vectors dx^μ and dx'^ν, such that

$$dx'^\mu = J^\mu{}_\nu dx^\nu, \tag{2.1}$$

whereby the inverse is just

$$dx^\mu = \left(J^{-1}\right)^\mu{}_\nu dx'^\nu.$$

By the chain rule, it is easy to see that the transformation matrix is just the Jacobian

$$J^\mu{}_\nu = \frac{\partial x'^\mu}{\partial x^\nu}. \tag{2.2}$$

The transformation & inverse satisfy

$$J^\mu{}_\nu \left(J^{-1}\right)^\nu{}_\sigma = \delta^\mu_\sigma. \tag{2.3}$$

This is easier to see if we represent the Jacobians in terms of differentials,

$$J^\mu{}_\nu \left(J^{-1}\right)^\nu{}_\sigma = \frac{\partial x'^\mu}{\partial x^\nu}\frac{\partial x^\nu}{\partial x'^\sigma} = \frac{\partial x'^\mu}{\partial x'^\sigma} = \delta^\mu_\sigma.$$

2.2.1 Example: Plane Polars

Consider that some point in the $\mathbb{R}^2$ plane may be defined by Cartesian coordinates (x, y) or plane polars, (r, θ). Then, we make the identifications

$$(x^1, x^2) = (x, y), \quad (x'^1, x'^2) = (r, \theta).$$

We also know that

$$x = r\cos\theta, \quad y = r\sin\theta; \quad r = \sqrt{x^2 + y^2}, \quad \theta = \tan^{-1} y/x.$$

Then, we can compute the elements of the Jacobian

$$\begin{aligned} J^i{}_j &= \frac{\partial x'^i}{\partial x^j} \\ &= \frac{\partial(r,\theta)}{\partial(x,y)} \\ &= \begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta & \sin\theta \\ -\frac{\sin\theta}{r} & \frac{\cos\theta}{r} \end{pmatrix}. \end{aligned}$$

And therefore,

$$\begin{aligned} dr &= \sum_j J^r{}_j dx^j \\ &= J^r{}_x dx + J^r{}_y dy \\ &= \cos\theta dx + \sin\theta dy. \end{aligned}$$

And similarly,

$$d\theta = -\frac{\sin\theta}{r}dx + \frac{\cos\theta}{r}dy.$$

2.3 Tangent Vector

Imagine that on a manifold $\mathcal{M}$, we have curves parameterised by u. On one curve, there is a point $\mathbf{p}(u)$. So, we have $x^\mu = x^\mu(u)$, then the tangent curve is defined to be

$$T^\mu = \left.\frac{dx^\mu}{du}\right|_{u=u_p}. \tag{2.4}$$

2.4 The Metric & Line Element

We have the line element

$$ds^2 = g_{\mu\nu}(x)dx^\mu dx^\nu. \tag{2.5}$$

Now, a common requirement, is the invariance of the line element (i.e. invariance of the interval). Thus, we require that

$$ds^2(x) = ds^2(x').$$

So, under transformation $x^\mu \mapsto x'^\nu(x^\mu)$, we want that

$$g_{\mu\nu}dx^\mu dx^\nu = g'_{\alpha\beta}dx'^\alpha dx'^\beta. \tag{2.6}$$

So, we proceed by writing down the known transformation of the RHS "primed" to "unprimed" displacement vectors,

$$g_{\mu\nu}dx^\mu dx^\nu = g'_{\alpha\beta}dx'^\alpha dx'^\beta = g'_{\alpha\beta}J^\alpha{}_\mu J^\beta{}_\nu dx^\mu dx^\nu.$$

But, this must always be consistent, so we see that we must have

$$g_{\mu\nu} = g'_{\alpha\beta}J^\alpha{}_\mu J^\beta{}_\nu. \tag{2.7}$$

We can derive a similar relation, by starting from (2.6), and instead of transforming the RHS, transform the LHS. So,

$$g'_{\alpha\beta}dx'^\alpha dx'^\beta = g_{\mu\nu}dx^\mu dx^\nu = g_{\mu\nu}\left(J^{-1}\right)^\nu{}_\beta\left(J^{-1}\right)^\mu{}_\alpha dx'^\alpha dx'^\beta,$$

which we require to always be true, leaving us with

$$g'_{\alpha\beta} = g_{\mu\nu}\left(J^{-1}\right)^\nu{}_\beta\left(J^{-1}\right)^\mu{}_\alpha. \tag{2.8}$$

The alternative way of writing the Jacobian leads us to be able to rewrite (trivially) expressions (2.7) and (2.8)

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

We call $g_{\mu\nu}$ the "metric", and $g^{\mu\nu}$ the "inverse metric"; where they must satisfy

$$g_{\mu\nu}g^{\nu\lambda} = \delta^\lambda_\mu. \tag{2.9}$$

2.4.1 Example: Polars

We know that the line element in plane polars is $ds^2 = dr^2 + r^2 d\theta^2$. Thus, we can read off the elements of the metric

$$(g_{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix},$$

and, by (2.9), we see that we require

$$(g^{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & 1/r^2 \end{pmatrix}.$$

In spherical polars, the line element is

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

and we can easily read off the metric

$$(g_{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix}, \quad (g^{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/r^2 & 0 \\ 0 & 0 & 1/r^2\sin^2\theta \end{pmatrix}.$$

Raising & Lowering We can use the metric to raise & lower indices. We shall not show this in use here; see the next subsection.

2.5 Vectors

We start this by discussing contravariant and covariant vectors.

2.5.1 Contravariant Vectors A^μ

These are sometimes just denoted "vectors".

These are defined to transform, under coordinate transformation $x^\mu \mapsto x'^\mu(x^\nu)$ as

$$A'^\mu = J^\mu{}_\nu A^\nu. \tag{2.10}$$

2.5.2 Covariant Vectors A_μ

These are sometimes called "covectors".

Let us say that we define a covector A_μ via

$$A_\mu = g_{\mu\nu} A^\nu.$$

Then, we may derive its transformation properties. Consider that

$$A'_\mu = g'_{\mu\nu} A'^\nu,$$

the RHS of which we know the transformation rules for

$$A'_\mu = g'_{\mu\nu} A'^\nu = \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu g_{\alpha\beta} J^\nu{}_\sigma A^\sigma.$$

We can rearrange the terms in this expression,

$$A'_\mu = g'_{\mu\nu} A'^\nu = \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu J^\nu{}_\sigma g_{\alpha\beta} A^\sigma,$$

so that we notice the appearance of a transformation-inverse multiplication, which results in a Kronecker-delta

$$A'_\mu = g'_{\mu\nu} A'^\nu = \left(J^{-1}\right)^\alpha{}_\mu \delta^\beta_\sigma g_{\alpha\beta} A^\sigma,$$

acting the Kronecker-delta results in (ignoring the middle equality now)

$$A'_\mu = \left(J^{-1}\right)^\alpha{}_\mu g_{\alpha\beta} A^\beta,$$

then lowering the index, via the metric,

$$A'_\mu = \left(J^{-1}\right)^\alpha{}_\mu A_\alpha. \tag{2.11}$$

And therefore, we have arrived at the relation we require.

2.5.3 The Scalar Product

The scalar product between two vectors is written

$$\mathsf{S} \cdot \mathsf{T} = S^\mu T^\nu g_{\mu\nu} = S_\nu T^\nu.$$

A fairly obvious thing we need to prove is the invariance of the dot-product. So,

$$S_\nu T^\nu = S'_\alpha T'^\beta \left(J^{-1}\right)^\alpha{}_\nu J^\nu{}_\beta = S'_\alpha T'^\beta \delta^\alpha_\beta = S'_\alpha T'^\alpha.$$

This is a very important proof. Infact, it also states that scalars are invariant under transformation.

Within the scalar product, we must briefly mention the modulus of a vector. We denote them as $||S||$, and define them

$$||S|| = \begin{cases} (S^\mu S_\mu)^{1/2} & \text{time}-\text{like}, ds^2 > 0 \\ (-S^\mu S_\mu)^{1/2} & \text{space}-\text{like}, ds^2 < 0. \end{cases}$$

2.5.4 Conformal Transformations

Following from the previous definition of the scalar product, we have the definition of the angle between two vectors;

$$\cos\theta = \frac{S^\mu T_\mu}{||S||\,||T||} = \frac{S^\mu T_\mu}{(S^\alpha S_\alpha)^{1/2}(T^\beta T_\beta)^{1/2}}. \tag{2.12}$$

A conformal transformation is defined as one whose angle between two vectors does not change. That is, under a conformal transformation, the angle between two vectors is unchanged.

Associated metrics are termed “conformal metrics”. How can we find such metrics? They are given by

$$\tilde{g}_{\mu\nu} = \Omega(x) g_{\mu\nu}, \quad \Omega(x) \neq 0. \tag{2.13}$$

We can see this by putting this new metric into the $\cos\theta$ expression,

$$\cos\tilde{\theta} = \frac{\tilde{g}_{\mu\nu} S^\mu T^\nu}{(\tilde{g}_{\alpha\gamma} S^\alpha S^\gamma)^{1/2}(\tilde{g}_{\beta\delta} T^\beta T^\delta)^{1/2}},$$

and by substituting $\tilde{g}_{\mu\nu} = \Omega(x) g_{\mu\nu}$, we see that the factors of Ω end up canceling, leaving the angle unchanged.

2.5.5 It is a Proper Vector?

Here, we ask if various quantities are “proper vectors”, or not.

Consider $C^\mu(x) = aA^\mu(x) + bB^\mu(x)$. It is clearly a proper vector, as each of its constituents transform as we expect - each is defined at the same coordinate point.

Consider $C^\mu = aA^\mu(x_1) + bB^\mu(x_2)$. This is not a proper vector, as the constituents are defined at different points, and different points transform differently.

2.6 Tensors

These are basically vectors, with more indices. We can also mix the indices, so that we have some up, some down.

For example, consider $F^{\mu\nu} \equiv A^\mu B^\nu$. We call it a second rank contravariant tensor, or a $\binom{2}{0}$-tensor. It clearly transforms as

$$F'^{\mu\nu} = A'^\mu B'^\nu = J^\mu{}_\alpha J^\nu{}_\beta A^\alpha B^\beta = J^\mu{}_\alpha J^\nu{}_\beta F^{\alpha\beta}.$$

Similarly, a second rank covariant tensor, or a $\binom{0}{2}$-tensor, transforms like

$$F'_{\mu\nu} = A'_\mu B'_\nu = \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu A_\alpha B_\beta = \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu F_{\alpha\beta}.$$

Finally, a mixed $\binom{1}{1}$-tensor transforms

$$F'^\mu{}_\nu = J^\mu{}_\alpha \left(J^{-1}\right)^\beta{}_\nu F^\alpha{}_\beta.$$

This obviously generalises to higher-rank tensors. One must include a Jacobian for each contravariant index, and one inverse Jacobian for each covariant index.

Getting equations & other expressions into tensorial form (i.e. into a form consistent with the above tensor transformations), is extremely useful. For example, given a tensor equation in one frame of reference, one therefore knows the form in all frames of reference. This becomes particularly useful when one finds a frame in which a particular equation becomes simple to analyse; then, one can simply transform out of that frame, and know that the analysis still holds.

Also, consider a tensor for whom all components are zero. Then, one cannot make a coordinate transformation that will be able to "reinstate" those (completely) zero components. That is, a tensor with zero components in one frame, has zero components in all frames. This is a very useful concept. If a quantity is not a tensor, then this does not hold true. That is, a non-tensor with zero components in one frame may have non-zero components in another.

2.6.1 Symmetric & Anti-symmetric Tensors

A symmetric $\binom{2}{0}$-tensor is one where

$$A^{\mu\nu} = A^{\nu\mu},$$

that is, the sign is unchanged under exchange of the indices. An anti-symmetric tensor is one for whom

$$B^{\mu\nu} = -B^{\nu\mu}.$$

Now then, using these relations (definitions, if you will), we can see some interesting formulae.

Suppose that $A^{\mu\nu}$ is a symmetric tensor. Then, $A^{\mu\nu} = A^{\nu\mu}$. Then, we see that

$$A^{\mu\nu} = \tfrac{1}{2}(A^{\mu\nu} + A^{\nu\mu}) = \tfrac{1}{2}(A^{\mu\nu} + A^{\mu\nu}) = \tfrac{1}{2}2A^{\mu\nu} = A^{\mu\nu}.$$

Similarly, suppose that $B^{\mu\nu}$ is an anti-symmetric tensor. Then,

$$B^{\mu\nu} = \tfrac{1}{2}(B^{\mu\nu} - B^{\nu\mu}) = \tfrac{1}{2}(B^{\mu\nu} + B^{\mu\nu}) = B^{\mu\nu}.$$

These obviously all hold for covariant tensors. Lets introduce some notation that will be pretty useful.

Suppose we have some tensor, defined as

$$T_{\mu\nu} \equiv \frac{1}{2}(B_{\mu\nu} - B_{\nu\mu}),$$

then, we write

$$B_{[\mu\nu]} \equiv \frac{1}{2}(B_{\mu\nu} - B_{\nu\mu}).$$

That is, we could say that $T_{\mu\nu}$ is formed by the anti-symmetric interchange of indices on $B_{\mu\nu}$. We use the "square brackets" to denote the anti-symmetric interchange. Similarly, suppose we have

$$C_{\mu\nu} \equiv \frac{1}{2}(A_{\mu\nu} + A_{\nu\mu}),$$

then, we define

$$A_{(\mu\nu)} \equiv \frac{1}{2}(A_{\mu\nu} + A_{\nu\mu}).$$

Thus, we say that $C_{\mu\nu}$ is formed by the symmetric interchange of indices. We used "round brackets" to denote the symmetric interchange.

Suppose we have some tensor, $Y_{\mu\nu}$. Then, we can write it as the sum of an anti-symmetric part, and a symmetric part. That is,

$$Y_{\mu\nu} = A_{[\mu\nu]} + A_{(\mu\nu)} = \frac{1}{2}(A_{\mu\nu} - A_{\nu\mu}) + \frac{1}{2}(A_{\mu\nu} + A_{\nu\mu}).$$

This is infact pretty obvious. If the tensor is symmetric, then $A_{[\mu\nu]} = 0$. And, if the tensor is anti-symmetric, then $A_{(\mu\nu)} = 0$.

The notation of a lower bracket to denote index interchange can be used in another way. Recall the electromagnetic field tensor,

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu,$$

then, we can write this as

$$F_{\mu\nu} = 2\partial_{[\mu} A_{\nu]}.$$

Also recall that two of Maxwells equations may be recovered from

$$\partial_\mu F_{\alpha\beta} + \partial_\alpha F_{\beta\mu} + \partial_\beta F_{\mu\alpha} = 0,$$

well, we can denote this (notice that this is a cyclic interchange of index) as

$$\partial_{(\mu} F_{\alpha\beta)} = 0.$$

In this final example, we were a little sloppy. There is infact a numerical factor associated with this; however, it gets very messy, and the factor cancels out anyway. However, one should be aware that there is a factor there.

3 Tensor Calculus

Here we shall lay some formal groundwork for dealing with objects in curved spacetime. We start by looking at differentiation, going on to geodesics.

3.1 Covariant Differentiation

Let us just state some notation. We have

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu}, \quad \partial^\mu \equiv \frac{\partial}{\partial x_\mu}.$$

Now, let us look at the coordinate transformation $x^\mu \mapsto x'^\mu(x^\nu)$. Then, we have that

$$dx'^\mu = J^\mu{}_\nu dx^\nu, \quad J^\mu{}_\nu = \frac{\partial x'^\mu}{\partial x^\nu} = \partial_\nu x'^\mu.$$

Now, let us consider differentiation of a scalar, and a coordination transformation (noting that scalars do not transform under a coordinate transformation); thus

$$\begin{aligned}\partial_\mu \phi \longmapsto \partial'_\mu \phi &= \frac{\partial}{\partial x'^\mu}\phi \\ &= \frac{\partial x^\nu}{\partial x'^\mu}\partial_\nu \phi \\ &= \left(J^{-1}\right)^\nu{}_\mu \partial_\nu \phi.\end{aligned}$$

Therefore, we see that the derivative of a scalar $\partial_\mu \phi$ transforms as a covariant vector

$$\partial'_\mu \phi = \left(J^{-1}\right)^\nu{}_\mu \partial_\nu \phi.$$

Now, let us try this with a vector (again, under a coordinate transformation)

$$\partial_\mu A^\nu \longmapsto \partial'_\mu A'^\nu;$$

where we want to derive how the RHS relates back to the LHS. Notice, if $\partial_\mu A^\nu$ is a $\binom{1}{1}$-tensor, then we know what it gives. However, let us derive it. So, using the known transformation rules for A^ν and ∂_μ,

$$\partial'_\mu A'^\nu = \left(J^{-1}\right)^\alpha{}_\mu \partial_\alpha J^\nu{}_\beta A^\beta.$$

Now, to continue, we must consider the partial derivative above. We must use the product rule on everything to the right of it. That is

$$\left(J^{-1}\right)^\alpha{}_\mu \partial_\alpha \left(J^\nu{}_\beta A^\beta\right) = \left(J^{-1}\right)^\alpha{}_\mu \left(\partial_\alpha J^\nu{}_\beta\right) A^\beta + \left(J^{-1}\right)^\alpha{}_\mu J^\nu{}_\beta \left(\partial_\alpha A^\beta\right).$$

This is not the transformation rule for a $\binom{1}{1}$-tensor, due to the presence of the first term on the RHS. We write the result, swapping the two terms on the RHS, to see this more clearly:

$$\partial'_\mu A'^\nu = \left(J^{-1}\right)^\alpha{}_\mu J^\nu{}_\beta \left(\partial_\alpha A^\beta\right) + \left(J^{-1}\right)^\alpha{}_\mu \left(\partial_\alpha J^\nu{}_\beta\right) A^\beta.$$

Therefore, we see that the partial derivative of a vector is not a tensor. The non-tensorial part is the added term on the far right. There is a rather more fundamental reasoning behind why the partial derivative of a vector is not tensorial. Recall that the partial derivative of a vector is defined as

$$\partial_\mu A^\nu(x^\alpha) = \lim_{\delta u\to 0} \frac{A^\nu(x^\alpha) - A^\nu(x^\alpha + \delta u)}{\delta u}.$$

So, the partial derivative is composed by finding the value of a vector at different points. As we have seen, the sum of two vectors evaluated at different points, is not a proper vector (this is due to the Jacobian being evaluated at different positions). Therefore, one should expect the partial derivative of a vector not to be tensorial; which is what we find.

Now, consider the vector

$$\mathbf{A}(x) = A^\nu(x)\mathbf{e}_\nu(x) = A'^\nu(x)\mathbf{e}'_\nu(x),$$

where we use the fact that a vector is the same in all frames. Now consider differentiating $\mathbf{A}$, noting that the components and basis vectors are all function of coordinate;

$$\partial_\nu \mathbf{A} = \partial_\nu \left(A^\mu \mathbf{e}_\mu\right) = (\partial_\nu A^\mu)\mathbf{e}_\mu + A^\mu(\partial_\nu \mathbf{e}_\mu).$$

Now, to continue, we shall write the final bracketed term as a sum over coefficients

$$\partial_\nu \mathbf{e}_\mu = \Gamma^\rho{}_{\nu\mu}\mathbf{e}_\rho.$$

The logic behind this will become clear. However, one may think of it in a similar way to quantum theory. Given a state, one can write it as a sum over coefficients times the basis. What we are doing here, is to say that $\partial_\nu \mathbf{e}_\mu$ is a "new object", and write that new object as a sum over the original basis $\mathbf{e}_\rho$, with coefficients $\Gamma^\rho{}_{\nu\mu}$. Notice that this then results in

$$\partial_\nu \mathbf{A} = (\partial_\nu A^\mu)\mathbf{e}_\mu + A^\mu \Gamma^\rho{}_{\nu\mu}\mathbf{e}_\rho.$$

In the final term, let us swap indices $\rho \to \mu$ and $\mu \to \beta$,

$$A^\mu \Gamma^\rho{}_{\mu\nu}\mathbf{e}_\rho \to A^\beta \Gamma^\mu{}_{\nu\beta}\mathbf{e}_\mu.$$

This therefore results in

$$\partial_\nu \mathbf{A} = (\partial_\nu A^\mu)\mathbf{e}_\mu + A^\beta \Gamma^\mu{}_{\nu\beta}\mathbf{e}_\mu,$$

which we factorise (and move the position of the final A^β) to

$$\partial_\nu \mathbf{A} = \mathbf{e}_\mu \left(\partial_\nu A^\mu + \Gamma^\mu{}_{\nu\beta}A^\beta\right).$$

Furthermore, we define the bracketed quantity as

$$\nabla_\nu A^\mu \equiv \partial_\nu A^\mu + \Gamma^\mu{}_{\nu\beta} A^\beta. \tag{3.1}$$

This defines the covariant derivative of a contravariant vector. We can use this rule for the covariant derivative of a contravariant vector to derive the rule for a covariant vector.

The covariant derivative of a contravariant vector is

$$\nabla_\alpha A^\mu = \partial_\alpha A^\mu + \Gamma^\mu{}_{\alpha\lambda} A^\lambda.$$

A covector is constructed from the contravariant vector via

$$A_\nu = g_{\nu\mu} A^\mu.$$

So,

$$\begin{aligned}
\nabla_\alpha A^\mu &= \nabla_\alpha \left(g^{\mu\nu} A_\nu\right) \\
&= g^{\mu\nu} \nabla_\alpha A_\nu + A_\nu \nabla_\alpha g^{\mu\nu} \\
&= \partial_\alpha A^\mu + \Gamma^\mu{}_{\alpha\lambda} A^\lambda \\
&= \partial_\alpha \left(g^{\mu\nu} A_\nu\right) + \Gamma^\mu{}_{\alpha\lambda} \left(g^{\lambda\beta} A_\beta\right) \\
&= A_\nu \partial_\alpha g^{\mu\nu} + g^{\mu\nu} \partial_\alpha A_\nu + \Gamma^\mu{}_{\alpha\lambda} g^{\lambda\beta} A_\beta.
\end{aligned}$$

If we equate the second and last lines,

$$g^{\mu\nu} \nabla_\alpha A_\nu + A_\nu \nabla_\alpha g^{\mu\nu} = A_\nu \partial_\alpha g^{\mu\nu} + g^{\mu\nu} \partial_\alpha A_\nu + \Gamma^\mu{}_{\alpha\lambda} g^{\lambda\beta} A_\beta.$$

Now, the index ν is a "dummy index", so we can swap $\beta \to \nu$ in the last term, to give

$$g^{\mu\nu} \nabla_\alpha A_\nu + A_\nu \nabla_\alpha g^{\mu\nu} = A_\nu \partial_\alpha g^{\mu\nu} + g^{\mu\nu} \partial_\alpha A_\nu + \Gamma^\mu{}_{\alpha\lambda} g^{\lambda\nu} A_\nu,$$

collecting terms,

$$g^{\mu\nu} \nabla_\alpha A_\nu = \left(\partial_\alpha g^{\mu\nu} + \Gamma^\mu{}_{\alpha\lambda} g^{\lambda\nu} - \nabla_\alpha g^{\mu\nu}\right) A_\nu + g^{\mu\nu} \partial_\alpha A_\nu.$$

We then expand out the covariant derivative of the metric (the third term in the bracket), to give

$$g^{\mu\nu} \nabla_\alpha A_\nu = \left(\partial_\alpha g^{\mu\nu} + \Gamma^\mu{}_{\alpha\lambda} g^{\lambda\nu} - \partial_\alpha g^{\mu\nu} - \Gamma^\mu{}_{\alpha\lambda} g^{\lambda\nu} - \Gamma^\nu{}_{\alpha\lambda} g^{\mu\lambda}\right) A_\nu + g^{\mu\nu} \partial_\alpha A_\nu.$$

Now, the first and third terms cancel each other out, as do the second and fourth. Leaving

$$g^{\mu\nu} \nabla_\alpha A_\nu = g^{\mu\nu} \partial_\alpha A_\nu - \Gamma^\nu{}_{\alpha\lambda} g^{\mu\lambda} A_\nu.$$

If we multiply through by $g_{\pi\mu}$, then we see that

$$g_{\pi\mu} g^{\mu\lambda} = \delta^\lambda_\pi, \quad g_{\pi\lambda} g^{\mu\nu} = \delta^\nu_\pi.$$

Hence, this gives

$$\delta^\nu_\pi \nabla_\alpha A_\nu = \delta^\nu_\pi \partial_\alpha A_\nu - \delta^\lambda_\pi \Gamma^\nu{}_{\alpha\lambda} A_\nu,$$

which is

$$\nabla_\alpha A_\pi = \partial_\alpha A_\pi - \Gamma^\nu{}_{\alpha\pi} A_\nu.$$

Putting into more "standard indices", we have our desired result. Hence, the covariant derivative of a covariant vector is

$$\nabla_\nu A_\mu \equiv \partial_\nu A_\mu - \Gamma^\beta{}_{\nu\mu} A_\beta. \tag{3.2}$$

Now, remember that a scalar is invariant; and that the derivative of a scalar is a tensor, we should have that $\nabla_\mu (A^\nu A_\nu) = \partial_\mu (A^\nu A_\nu)$. This can be checked. So,

$$\begin{aligned}
\nabla_\mu (A^\nu A_\nu) &= (\nabla_\mu A^\nu) A_\nu + A^\nu (\nabla_\mu A_\nu) \\
&= A_\nu \left(\partial_\mu A^\nu + \Gamma^\nu{}_{\mu\beta} A^\beta \right) + A^\nu \left(\partial_\mu A_\nu - \Gamma^\alpha{}_{\mu\nu} A_\alpha \right) \\
&= \partial_\mu \left(A^\nu A_\nu \right) + A_\nu \Gamma^\nu{}_{\mu\beta} A^\beta - A^\nu \Gamma^\alpha{}_{\mu\nu} A_\alpha \\
&= \partial_\mu \left(A^\nu A_\nu \right) + A_\nu A^\beta \Gamma^\nu{}_{\mu\beta} - A^\nu A_\alpha \Gamma^\alpha{}_{\mu\nu}.
\end{aligned}$$

Now, the last two expressions can be shown to cancel, by interchanging indices. Let us manipulate the final expression

$$A^\nu A_\alpha \Gamma^\alpha{}_{\mu\nu} \quad \alpha \to \nu \to \beta \quad \Rightarrow \quad A^\beta A_\nu \Gamma^\nu{}_{\mu\beta},$$

and so, if we put this expression back in, we see that

$$\begin{aligned}
\nabla_\mu (A^\nu A_\nu) &= \partial_\mu \left(A^\nu A_\nu \right) + A_\nu A^\beta \Gamma^\nu{}_{\mu\beta} - A^\beta A_\nu \Gamma^\nu{}_{\mu\beta} \\
&= \partial_\mu \left(A^\nu A_\nu \right).
\end{aligned}$$

Therefore, we see an expected result: the covariant derivative of a scalar is the same as the partial derivative.

We call the expansion coefficients $\Gamma^\lambda{}_{\nu\mu}$ the *affine connection.*

We are able to find the covariant derivative of tensors of arbitrary rank. A few are given below.

$$\begin{aligned}
\nabla_\alpha A^{\mu\nu} &= \partial_\alpha A^{\mu\nu} + \Gamma^\mu{}_{\alpha\lambda} A^{\lambda\nu} + \Gamma^\nu{}_{\alpha\lambda} A^{\mu\lambda}, \\
\nabla_\alpha A_{\mu\nu} &= \partial_\alpha A_{\mu\nu} - \Gamma^\lambda{}_{\alpha\mu} A_{\lambda\nu} - \Gamma^\lambda{}_{\alpha\nu} A_{\mu\lambda}, \\
\nabla_\alpha A^\mu{}_\nu &= \partial_\alpha A^\mu{}_\nu + \Gamma^\mu{}_{\alpha\lambda} A^\lambda{}_\nu - \Gamma^\lambda{}_{\alpha\nu} A^\mu{}_\lambda, \\
\nabla_\alpha A^{\mu\nu\sigma} &= \partial_\alpha A^{\mu\nu\sigma} + \Gamma^\mu{}_{\alpha\lambda} A^{\lambda\nu\sigma} + \Gamma^\nu{}_{\alpha\lambda} A^{\mu\lambda\sigma} + \Gamma^\sigma{}_{\alpha\lambda} A^{\mu\nu\lambda}.
\end{aligned}$$

Basically, for each contravariant component, there should be a positive connection term, and for each covariant a negative term.

3.1.1 Parallel Transport

The main idea in parallel transport is this:

Consider moving a vector from one place to another. Then, in general, that vector will change direction; thus, a change in the vector upon moving said vector. So, we can find the difference in a vector,

$$DA^\mu = A^\mu(x') - \bar{A}^\mu(x').$$

Considering how the basis changes as well, we end up with

$$DA^\mu = \delta x^\nu \left(\partial_\nu A^\mu + \Gamma^\mu{}_{\nu\lambda} A^\lambda\right)$$

The bracketed quantity is just the covariant derivative. Thus,

$$DA^\mu = \delta x^\nu \nabla_\nu A^\mu.$$

Now, the point is that this gives another insight as to what the covariant derivative is. When moving a vector around a manifold, one must consider how the basis vectors change from point to point, as well as the components. This information is within the affine connection.

For an example as to what parallel transport is, consider a circle in the plane. Consider that there is an arrow living on the circle, pointing in a given direction (say parallel to the y-axis). Then, consider moving the arrow around the circle. The arrow undergoes parallel transport if it always points in the same direction, nomatter what its position on the circle. Now, consider that the entire space is the circle-line. That is, we have a 1D manifold. For a vector living on the manifold, parallel transport means moving on tangents to the circle.

3.1.2 Absolute Derivative

We define the absolute derivative as

$$\frac{DA^\mu}{Du} = \frac{dx^\nu}{du} \nabla_\nu A^\mu,$$

where we have considered a curve, parameterised so that

$$A^\mu = A^\mu(x^\nu(u)).$$

3.1.3 Transformation of $\Gamma^\lambda{}_{\nu\mu}$

Let us consider the transformation property of the affine connection, $\Gamma^\lambda{}_{\nu\mu}$. Let us start with our previous definition, but in the primed-frame (we will then transform to the unprimed)

$$\Gamma'^\rho{}_{\mu\nu} \mathbf{e}'_\rho = \partial'_\mu \mathbf{e}'_\nu.$$

Then, we know how to transform the RHS,

$$\partial'_\mu \mathbf{e}'_\nu = \left(J^{-1}\right)^\alpha{}_\mu \partial_\alpha \left(J^{-1}\right)^\beta{}_\nu \mathbf{e}_\beta,$$

we then use the product rule on the RHS,

$$\left(J^{-1}\right)^\alpha{}_\mu \partial_\alpha \left(J^{-1}\right)^\beta{}_\nu \mathbf{e}_\beta = \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu \partial_\alpha \mathbf{e}_\beta + \left(J^{-1}\right)^\alpha{}_\mu \mathbf{e}_\beta \partial_\alpha \left(J^{-1}\right)^\beta{}_\nu .$$

Now, we also know that $\partial_\alpha \mathbf{e}_\beta = \Gamma^\delta{}_{\alpha\beta}\mathbf{e}_\delta$, so that

$$\left(J^{-1}\right)^\alpha{}_\mu \partial_\alpha \left(J^{-1}\right)^\beta{}_\nu \mathbf{e}_\beta = \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu \Gamma^\delta{}_{\alpha\beta}\mathbf{e}_\delta + \left(J^{-1}\right)^\alpha{}_\mu \mathbf{e}_\beta \partial_\alpha \left(J^{-1}\right)^\beta{}_\nu ,$$

remembering that the LHS is of course just

$$\Gamma'^\rho{}_{\mu\nu}\mathbf{e}'_\rho = \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu \Gamma^\delta{}_{\alpha\beta}\mathbf{e}_\delta + \left(J^{-1}\right)^\alpha{}_\mu \mathbf{e}_\beta \partial_\alpha \left(J^{-1}\right)^\beta{}_\nu .$$

If we then transform the basis vector on the LHS, we have

$$\Gamma'^\rho{}_{\mu\nu} \left(J^{-1}\right)^\lambda{}_\rho \mathbf{e}_\lambda = \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu \Gamma^\delta{}_{\alpha\beta}\mathbf{e}_\delta + \left(J^{-1}\right)^\alpha{}_\mu \mathbf{e}_\beta \partial_\alpha \left(J^{-1}\right)^\beta{}_\nu .$$

On the RHS, let us change the indices on the basis vectors, so that they are the same as those on the left. That is, $\delta \to \lambda$ and $\beta \to \lambda$;

$$\Gamma'^\rho{}_{\mu\nu} \left(J^{-1}\right)^\lambda{}_\rho \mathbf{e}_\lambda = \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu \Gamma^\lambda{}_{\alpha\beta}\mathbf{e}_\lambda + \left(J^{-1}\right)^\alpha{}_\mu \mathbf{e}_\lambda \partial_\alpha \left(J^{-1}\right)^\lambda{}_\nu ,$$

which allows us to then cancel off the basis vectors,

$$\Gamma'^\rho{}_{\mu\nu} \left(J^{-1}\right)^\lambda{}_\rho = \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu \Gamma^\lambda{}_{\alpha\beta} + \left(J^{-1}\right)^\alpha{}_\mu \partial_\alpha \left(J^{-1}\right)^\lambda{}_\nu .$$

If we then multiply this through by something which will kill-off the inverse Jacobian on the LHS, we will have got to our result. Notice that $J^\pi{}_\lambda$ will do this. So,

$$\begin{aligned}
\Gamma'^\rho{}_{\mu\nu} J^\pi{}_\lambda \left(J^{-1}\right)^\lambda{}_\rho &= J^\pi{}_\lambda \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu \Gamma^\lambda{}_{\alpha\beta} + J^\pi{}_\lambda \left(J^{-1}\right)^\alpha{}_\mu \partial_\alpha \left(J^{-1}\right)^\lambda{}_\nu \\
\Rightarrow \quad \Gamma'^\rho{}_{\mu\nu}\delta^\pi_\rho &= J^\pi{}_\lambda \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu \Gamma^\lambda{}_{\alpha\beta} + J^\pi{}_\lambda \left(J^{-1}\right)^\alpha{}_\mu \partial_\alpha \left(J^{-1}\right)^\lambda{}_\nu \\
\Rightarrow \quad \Gamma'^\pi{}_{\mu\nu} &= J^\pi{}_\lambda \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu \Gamma^\lambda{}_{\alpha\beta} + J^\pi{}_\lambda \left(J^{-1}\right)^\alpha{}_\mu \partial_\alpha \left(J^{-1}\right)^\lambda{}_\nu .
\end{aligned}$$

We therefore have our result: the transformation of the affine connection is

$$\Gamma'^\pi{}_{\mu\nu} = J^\pi{}_\lambda \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu \Gamma^\lambda{}_{\alpha\beta} + J^\pi{}_\lambda \left(J^{-1}\right)^\alpha{}_\mu \partial_\alpha \left(J^{-1}\right)^\lambda{}_\nu . \tag{3.3}$$

Now, although not a notation we have been using much, we can represent the Jacobians by differentials,

$$J^\mu{}_\nu = \frac{\partial x'^\mu}{\partial x^\nu}, \quad \left(J^{-1}\right)^\mu{}_\nu = \frac{\partial x^\mu}{\partial x'^\nu};$$

and, using this notation, the transformation of the affine connection looks like

$$\Gamma'^{\pi}{}_{\mu\nu} = \frac{\partial x'^{\pi}}{\partial x^{\lambda}}\frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial x^{\beta}}{\partial x'^{\nu}}\Gamma^{\lambda}{}_{\alpha\beta} + \frac{\partial x'^{\pi}}{\partial x^{\lambda}}\frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial}{\partial x^{\alpha}}\frac{\partial x^{\lambda}}{\partial x'^{\nu}}.$$

We can see that this immediately shows that the affine connection is not a tensor (due to the existence of the second term on the RHS). Now, if the affine connection were a tensor, then, if one were to find a coordinate system in which all the components were zero, then they must be zero in all coordinate systems (this is a general property of tensors). That the affine connection is not a $\binom{1}{2}$-tensor means that even if the connection has zero components in one frame, there exists frames in which the components are non-zero. Infact, one can show that there exists a frame in which the components are zero, at a point. We shall now show that.

3.1.4 Locally Inertial Frames

This will all seem a little pointless, until we reach the very end of our discussion.

Let us make the following coordinate transformation,

$$x'^{\mu} = \bar{x}^{\mu} + \frac{1}{2}\Gamma^{\mu}{}_{\alpha\beta}\bar{x}^{\alpha}\bar{x}^{\beta}, \quad \bar{x}^{\mu} \equiv x^{\mu} - x^{\mu}_{*},$$

where x^{μ}_{*} is a single point. Now, under this transformation, we can write down the Jacobian

$$\begin{aligned} J^{\mu}{}_{\nu} = \frac{\partial x'^{\mu}}{\partial x^{\nu}} &= \frac{\partial}{\partial x^{\nu}}\left(\bar{x}^{\mu} + \frac{1}{2}\Gamma^{\mu}{}_{\alpha\beta}\bar{x}^{\alpha}\bar{x}^{\beta}\right) \\ &= \delta^{\mu}_{\nu} + \frac{1}{2}\bar{x}^{\alpha}\bar{x}^{\beta}\partial_{\nu}\Gamma^{\mu}{}_{\alpha\beta} + \frac{1}{2}\Gamma^{\mu}{}_{\alpha\beta}\left(\delta^{\alpha}_{\nu}\bar{x}^{\beta} + \delta^{\beta}_{\nu}\bar{x}^{\alpha}\right) \\ &= \delta^{\mu}_{\nu} + \frac{1}{2}\bar{x}^{\alpha}\bar{x}^{\beta}\partial_{\nu}\Gamma^{\mu}{}_{\alpha\beta} + \Gamma^{\mu}{}_{\nu\beta}\bar{x}^{\beta}, \end{aligned}$$

thus, the Jacobian is

$$J^{\mu}{}_{\nu} = \delta^{\mu}_{\nu} + \frac{1}{2}\bar{x}^{\alpha}\bar{x}^{\beta}\partial_{\nu}\Gamma^{\mu}{}_{\alpha\beta} + \Gamma^{\mu}{}_{\nu\beta}\bar{x}^{\beta}. \tag{3.4}$$

Notice that this can be written,

$$J^{\mu}{}_{\nu} = \delta^{\mu}_{\nu} + \mathcal{O}(\bar{x}^{\beta}). \tag{3.5}$$

Infact, the inverse Jacobian is also this,

$$\left(J^{-1}\right)^{\mu}{}_{\nu} = \delta^{\mu}_{\nu} - \mathcal{O}(\bar{x}^{\beta}). \tag{3.6}$$

Now then, returning to (3.4), we see that we can differentiate it,

$$\begin{aligned} \partial_{\alpha}J^{\pi}{}_{\lambda} &= \partial_{\alpha}\left(\delta^{\pi}_{\lambda} + \frac{1}{2}\left(\Gamma^{\pi}{}_{\lambda\beta}\bar{x}^{\beta} + \Gamma^{\pi}{}_{\lambda\nu}\bar{x}^{\nu}\right)\right) + \mathcal{O}(\bar{x}^{\beta}) \\ &= \frac{1}{2}\left(\Gamma^{\pi}{}_{\lambda\beta}\delta^{\beta}_{\alpha} + \Gamma^{\pi}{}_{\nu\lambda}\delta^{\nu}_{\alpha}\right) + \mathcal{O}(\bar{x}^{\beta}) \\ &= \frac{1}{2}\left(\Gamma^{\pi}{}_{\lambda\alpha} + \Gamma^{\pi}{}_{\alpha\lambda}\right) + \mathcal{O}(\bar{x}^{\beta}) \\ &= \Gamma^{\pi}{}_{\lambda\alpha} + \mathcal{O}(\bar{x}^{\beta}). \end{aligned}$$

Thus,

$$\partial_\alpha J^\pi{}_\lambda = \Gamma^\pi{}_{\lambda\alpha} + \mathcal{O}(\bar{x}^\beta). \tag{3.7}$$

Now, we previously derived the transformation rule of the affine connection,

$$\Gamma'^\pi{}_{\mu\nu} = J^\pi{}_\lambda \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu \Gamma^\lambda{}_{\alpha\beta} + J^\pi{}_\lambda \left(J^{-1}\right)^\alpha{}_\mu \partial_\alpha \left(J^{-1}\right)^\lambda{}_\nu .$$

Let us look at the final term,

$$J^\pi{}_\lambda \left(J^{-1}\right)^\alpha{}_\mu \partial_\alpha \left(J^{-1}\right)^\lambda{}_\nu ,$$

we see that we can write it as

$$-\left(J^{-1}\right)^\lambda{}_\nu \left(J^{-1}\right)^\alpha{}_\mu \partial_\alpha J^\pi{}_\lambda.$$

To see how we can do this, consider that

$$\delta^\alpha_\beta = \frac{\partial x'^\alpha}{\partial x'^\beta} = \frac{\partial x'^\alpha}{\partial x^\pi}\frac{\partial x^\pi}{\partial x'^\beta}.$$

Also, $\partial_\nu \delta^\alpha_\beta = 0$. Then, that means that

$$\begin{aligned}
\partial_\nu \delta^\alpha_\beta &= \frac{\partial}{\partial x^\nu}\frac{\partial x'^\alpha}{\partial x^\pi}\frac{\partial x^\pi}{\partial x'^\beta} \\
&= \frac{\partial x'^\alpha}{\partial x^\pi}\frac{\partial^2 x^\pi}{\partial x^\nu \partial x'^\beta} + \frac{\partial x^\pi}{\partial x'^\beta}\frac{\partial^2 x'^\alpha}{\partial x^\nu \partial x^\pi} \\
&= 0.
\end{aligned}$$

That is,

$$\frac{\partial x'^\alpha}{\partial x^\pi}\frac{\partial^2 x^\pi}{\partial x^\nu \partial x'^\beta} = -\frac{\partial x^\pi}{\partial x'^\beta}\frac{\partial^2 x'^\alpha}{\partial x^\nu \partial x^\pi}.$$

Or, using the Jacobian notation,

$$J^\alpha{}_\pi \partial_\nu \left(J^{-1}\right)^\pi{}_\beta = -\left(J^{-1}\right)^\pi{}_\beta \partial_\nu J^\alpha{}_\pi.$$

Thus, we have shown that we can do the “swap” we did above.

Therefore, we write the transformation rule of the connection once again, with this rewriting of the last term,

$$\Gamma'^\pi{}_{\mu\nu} = J^\pi{}_\lambda \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu \Gamma^\lambda{}_{\alpha\beta} - \left(J^{-1}\right)^\lambda{}_\nu \left(J^{-1}\right)^\alpha{}_\mu \partial_\alpha J^\pi{}_\lambda.$$

Now, we have all of these expression. We use (3.5) and (3.6) for the Jacobian/inverse, and (3.7) for the derivative of the Jacobian;

$$\Gamma'^\pi{}_{\mu\nu} = \delta^\pi_\lambda \delta^\alpha_\mu \delta^\beta_\nu \Gamma^\lambda{}_{\alpha\beta} - \delta^\lambda_\nu \delta^\alpha_\mu \Gamma^\pi{}_{\alpha\lambda} + \mathcal{O}(\bar{x}^\rho).$$

Using the Kronecker-deltas results in

$$\Gamma'^{\pi}{}_{\mu\nu} = \Gamma^{\pi}{}_{\mu\nu} - \Gamma^{\pi}{}_{\mu\nu} + \mathcal{O}(\bar{x}^{\rho}) = \mathcal{O}(\bar{x}^{\rho}).$$

Therefore, the components of the transformed connection are all

$$\Gamma'^{\pi}{}_{\mu\nu} = \mathcal{O}(\bar{x}^{\rho}).$$

Now, if we let $\bar{x}^{\mu} \to 0$, which is equivalent (by our definition of $\bar{x}^{\mu}$) to saying $x^{\mu} = x_*^{\mu}$, then we see that

$$\Gamma'^{\pi}{}_{\mu\nu}(x^{\rho} = x_*^{\rho}) = 0.$$

Therefore, we have a transformation which renders all components of the affine connection zero. That is, we can transform to a frame in which the geometry is Euclidean (flat), at that single point. This is actually an incredibly useful & important result. Notice that if the Christofell symbols are zero, then the covariant derivative is just the partial derivative. This tends to hugely simplify calculations. In the later discussions on curvature, we shall see that in transforming to a locally inertial frame, where the connection components are zero, we can compute this a lot easier. And, as the things we are transforming are tensors, the results hold in any frame.

Some literature call such a set of coordinates, geodesic coordinates.

Alternative Derivation Here we shall present a rather more mathematically rigorous derivation of the existence of geodesic coordinates.

Let $x^{\mu} = a^{\mu}$ be coordinates at a point A in the frame Σ. Let us transform to a new frame, via the transformation

$$x^{\mu} = a^{\mu} + x'^{\mu} + \frac{1}{2}a^{\mu}{}_{\nu\lambda}x'^{\nu}x'^{\lambda},$$

where the coefficient $a^{\mu}{}_{\nu\lambda}$ is symmetric in its lower indices, and is constant (i.e. we define this as part of the transformation). Thus, at the point A, $x'^{\mu} = 0$. So, let us compute the differentials of the transformation;

$$\begin{aligned}\frac{\partial x^{\mu}}{\partial x'^{\nu}} &= \delta^{\mu}_{\nu} + \frac{1}{2}a^{\mu}{}_{\kappa\lambda}\frac{\partial}{\partial x'^{\nu}}\left(x'^{\kappa}x'^{\lambda}\right) \\ &= \delta^{\mu}_{\nu} + \frac{1}{2}a^{\mu}{}_{\kappa\lambda}\left(x'^{\kappa}\delta^{\lambda}_{\nu} + x'^{\lambda}\delta^{\kappa}_{\nu}\right) \\ &= \delta^{\mu}_{\nu} + a^{\mu}{}_{\nu\lambda}x'^{\lambda}.\end{aligned}$$

Hence,

$$\frac{\partial^2 x^{\mu}}{\partial x'^{\nu}\partial x'^{\lambda}} = a^{\mu}{}_{\nu\lambda}.$$

Hence, at the point A (i.e. where $x'^{\mu} = 0$), we see that

$$\frac{\partial x^{\mu}}{\partial x'^{\nu}} = \delta^{\mu}_{\nu}, \quad \frac{\partial^2 x^{\mu}}{\partial x'^{\nu}\partial x'^{\lambda}} = a^{\mu}{}_{\nu\lambda}. \tag{3.8}$$

Now, we can do a little work to get a relation between the coefficents $a^\mu{}_{\nu\lambda}$ and the metric. The metric transforms via

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

Then, differentiating it,

$$\frac{\partial g'_{\mu\nu}}{\partial x'^\lambda} = \frac{\partial^2 x^\alpha}{\partial x'^\lambda \partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu} g_{\alpha\beta} + \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial^2 x^\beta}{\partial x'^\lambda \partial x'^\nu} g_{\alpha\beta} + \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu}\frac{\partial g_{\alpha\beta}}{\partial x'^\lambda}.$$

Now, rewrite the last term using the chain rule;

$$\frac{\partial g_{\alpha\beta}}{\partial x'^\lambda} = \frac{\partial g_{\alpha\beta}}{\partial x^\sigma}\frac{\partial x^\sigma}{\partial x'^\lambda},$$

so that we have

$$\frac{\partial g'_{\mu\nu}}{\partial x'^\lambda} = \frac{\partial^2 x^\alpha}{\partial x'^\lambda \partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu} g_{\alpha\beta} + \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial^2 x^\beta}{\partial x'^\lambda \partial x'^\nu} g_{\alpha\beta} + \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu}\frac{\partial g_{\alpha\beta}}{\partial x^\sigma}\frac{\partial x^\sigma}{\partial x'^\lambda}.$$

So, at the point A, using (3.8), this reads

$$\begin{aligned}\frac{\partial g'_{\mu\nu}}{\partial x'^\lambda} &= a^\alpha{}_{\lambda\mu}\delta^\beta_\nu g_{\alpha\beta} + a^\beta{}_{\lambda\nu}\delta^\alpha_\mu g_{\alpha\beta} + \delta^\alpha_\mu\delta^\beta_\nu\delta^\sigma_\lambda\frac{\partial g_{\alpha\beta}}{\partial x^\sigma} \\ &= g_{\alpha\nu}a^\alpha{}_{\lambda\nu} + g_{\mu\beta}a^\beta{}_{\mu\beta} + \frac{\partial g_{\mu\nu}}{\partial x^\lambda}.\end{aligned} \tag{3.9}$$

Now let us choose that

$$\frac{\partial g'_{\mu\nu}}{\partial x'^\lambda} = 0. \tag{3.10}$$

which is equivalent to choosing the metric to be flat at that point A. Now, note that

$$g_{\alpha\nu}a^\alpha{}_{\lambda\mu} = a_{\nu\lambda\mu}.$$

Then, we see that (3.9) becomes

$$a_{\nu\lambda\mu} + a_{\mu\lambda\nu} + \frac{\partial g_{\mu\nu}}{\partial x^\lambda} = 0,$$

which trivially becomes

$$a_{\nu\lambda\mu} + a_{\mu\lambda\nu} = -\frac{\partial g_{\mu\nu}}{\partial x^\lambda}. \tag{3.11}$$

Now, if we permute the indices $\nu \to \mu \to \lambda \to \nu$, this becomes

$$a_{\mu\nu\lambda} + a_{\lambda\nu\mu} = -\frac{\partial g_{\lambda\mu}}{\partial x^\nu}, \tag{3.12}$$

permuting again,

$$a_{\lambda\mu\nu} + a_{\nu\mu\lambda} = -\frac{\partial g_{\nu\lambda}}{\partial x^\mu}. \tag{3.13}$$

Now, if we form (3.12) + (3.13)–(3.11), then we see

$$\begin{aligned} a_{\mu\nu\lambda} + a_{\lambda\nu\mu} + a_{\lambda\mu\nu} + a_{\nu\mu\lambda} - a_{\nu\lambda\mu} - a_{\mu\lambda\nu} \\ = -\frac{\partial g_{\nu\lambda}}{\partial x^\mu} - \frac{\partial g_{\lambda\mu}}{\partial x^\nu} + \frac{\partial g_{\mu\nu}}{\partial x^\lambda}. \end{aligned}$$

Now, as $a_{\mu\nu\lambda} = a_{\mu\lambda\nu}$, we see that the fourth and fifth terms cancel, as do the first and sixth, leaving

$$2a_{\lambda\nu\mu} = -\left(\frac{\partial g_{\nu\lambda}}{\partial x^\mu} + \frac{\partial g_{\lambda\mu}}{\partial x^\nu} - \frac{\partial g_{\mu\nu}}{\partial x^\lambda}\right),$$

that is,

$$a_{\lambda\nu\mu} = [\lambda\nu, \mu],$$

where

$$[\lambda\nu, \mu] \equiv -\frac{1}{2}\left(\frac{\partial g_{\nu\lambda}}{\partial x^\mu} + \frac{\partial g_{\lambda\mu}}{\partial x^\nu} - \frac{\partial g_{\mu\nu}}{\partial x^\lambda}\right).$$

We call the $[\lambda\nu, \mu]$ a *Christofell symbol of the first kind.* Now, by (3.10) we see that

$$a'_{\lambda\nu\mu} = 0$$

at the point A.

Therefore, we have derived that under the coordinate transformation $x^\mu = a^\mu + x'^\mu + \frac{1}{2}a^\mu{}_{\nu\lambda}x'^\nu x'^\lambda$, the Christofell symbols are zero at the point $x^\mu = a^\mu$; such coordinates are geodesic coordinates. In the derivation, we assumed that:

$a^\mu{}_{\nu\lambda}$	constant and symmetric in lower indices,
$\dfrac{\partial g'_{\mu\nu}}{\partial x'^\lambda} = 0$	constant metric at point we transform to.

3.1.5 Torsion

Let us define torsion to be

$$T^\rho{}_{\mu\nu} \equiv \frac{1}{2}\left(\Gamma^\rho{}_{\mu\nu} - \Gamma^\rho{}_{\nu\mu}\right) = \Gamma^\rho{}_{[\mu\nu]}. \tag{3.14}$$

We shall work with symmetric affine connections; so that the torsion goes to zero. A torsion free space merely allows us to interchange the lower indices on the connection components at will. This expression for torsion is a tensor; let us prove it.

So, the transformation of torsion can be written

$$T'^{\rho}{}_{\mu\nu} = J^{\rho}{}_{\alpha} \left(J^{-1}\right)^{\beta}{}_{\mu} \left(J^{-1}\right)^{\gamma}{}_{\nu} T^{\alpha}{}_{\beta\gamma} + \delta T^{\rho}{}_{\mu\nu},$$

where $\delta T^{\rho}{}_{\mu\nu}$ is a term that can be easily seen from the transformation rule of the connection
,

$$\delta T^{\rho}{}_{\mu\nu} = J^{\rho}{}_{\lambda} \left(J^{-1}\right)^{\alpha}{}_{\mu} \partial_{\alpha} \left(J^{-1}\right)^{\lambda}{}_{\nu} - J^{\rho}{}_{\lambda} \left(J^{-1}\right)^{\alpha}{}_{\nu} \partial_{\alpha} \left(J^{-1}\right)^{\lambda}{}_{\mu}.$$

Now, a "trick" that we have used before is to note that

$$\begin{aligned}
\delta^{\mu}_{\nu} &= J^{\mu}{}_{\alpha} \left(J^{-1}\right)^{\alpha}{}_{\nu} \\
\Rightarrow \partial_{\beta}\delta^{\mu}_{\nu} &= \partial_{\beta} \left(J^{\mu}{}_{\alpha} \left(J^{-1}\right)^{\alpha}{}_{\nu}\right) \\
&= J^{\mu}{}_{\alpha}\partial_{\beta} \left(J^{-1}\right)^{\alpha}{}_{\nu} + \left(J^{-1}\right)^{\alpha}{}_{\nu} \partial_{\beta} J^{\mu}{}_{\alpha} \\
&= 0 \\
\Rightarrow J^{\mu}{}_{\alpha}\partial_{\beta} \left(J^{-1}\right)^{\alpha}{}_{\nu} &= -\left(J^{-1}\right)^{\alpha}{}_{\nu} \partial_{\beta} J^{\mu}{}_{\alpha}.
\end{aligned}$$

We can then use this in the expression for $\delta T^{\rho}{}_{\mu\nu}$, to see that

$$\begin{aligned}
\delta T^{\rho}{}_{\mu\nu} &= -\left(J^{-1}\right)^{\alpha}{}_{\mu} \left(J^{-1}\right)^{\lambda}{}_{\nu} \partial_{\alpha} J^{\rho}{}_{\lambda} + \left(J^{-1}\right)^{\alpha}{}_{\nu} \left(J^{-1}\right)^{\lambda}{}_{\mu} \partial_{\alpha} J^{\rho}{}_{\lambda} \\
&= \partial_{\alpha} J^{\rho}{}_{\lambda} \left[\left(J^{-1}\right)^{\alpha}{}_{\nu} \left(J^{-1}\right)^{\lambda}{}_{\mu} - \left(J^{-1}\right)^{\alpha}{}_{\mu} \left(J^{-1}\right)^{\lambda}{}_{\nu}\right].
\end{aligned}$$

Now, let us define

$$A^{\alpha\lambda}_{\nu\mu} \equiv \left(J^{-1}\right)^{\alpha}{}_{\nu} \left(J^{-1}\right)^{\lambda}{}_{\mu} - \left(J^{-1}\right)^{\alpha}{}_{\mu} \left(J^{-1}\right)^{\lambda}{}_{\nu},$$

then we see that

$$A^{\alpha\lambda}_{\nu\mu} = -A^{\lambda\alpha}_{\nu\mu},$$

i.e. it is anti-symmetric under interchange of α and λ. Now, notice that

$$\begin{aligned}
\partial_{\alpha} J^{\rho}{}_{\lambda} &= \frac{\partial^2 x'^{\rho}}{\partial x^{\alpha} \partial x^{\lambda}} \\
&= \frac{\partial^2 x'^{\rho}}{\partial x^{\lambda} \partial x^{\alpha}} \\
&= \partial_{\lambda} J^{\rho}{}_{\alpha}.
\end{aligned}$$

That is, $\partial_{\alpha} J^{\rho}{}_{\lambda}$ is symmetric under interchange of α and λ. Therefore, as the product of something which is symmetric and anti-symmetric is zero, we see that

$$\delta T^{\rho}{}_{\mu\nu} = 0.$$

Hence,

$$T'^{\rho}{}_{\mu\nu} = J^{\rho}{}_{\alpha} \left(J^{-1}\right)^{\beta}{}_{\mu} \left(J^{-1}\right)^{\gamma}{}_{\nu} T^{\alpha}{}_{\beta\gamma},$$

which is the rule of transformation of a $\binom{1}{2}$-tensor.

3.2 Geodesics

A geodesic is the curve which gives an extremal of motion. That we use the word extremal, rather than minima (or, indeed, maxima), is very important.

Suppose we are “living in a manifold” (suppose we are confined to the surface of a sphere). Then suppose that we wish to compute the equation of the line (in that manifold) that joins two points, where the equation of the line is an extremum. That is, we can compute many equations of that line, but only one of them will be an extremum. Then, that curve is a *geodesic*.

The geodesic will depend upon the geometry of the manifold; the line has its motion confined to the manifold. As we shall see, the metric is used to give the geometrical dependence.

3.2.1 The Affine Geodesic

We call an *affine geodesic* the curve for which the tangent vector is parallel transported to itself. That is,

$$\frac{DT^\mu}{Du} = \lambda(u)T^\mu, \quad T^\mu \equiv \frac{dx^\mu}{du}.$$

That is, we find a curve, along which the tangent vector does not change direction. It may get longer (hence the factor of $\lambda(u)$), but it does not change direction.

We have our definition of the absolute derivative,

$$\frac{DA^\mu}{Du} = \frac{dx^\nu}{du}\nabla_\nu A^\mu = T^\nu \nabla_\nu A^\mu.$$

Therefore, the affine geodesic satisfies

$$T^\nu \nabla_\nu T^\mu = \lambda(u)T^\mu,$$

which is, using the definition of the covariant derivative

$$T^\nu \left(\partial_\nu T^\mu + \Gamma^\mu{}_{\nu\gamma}T^\gamma\right) = \lambda T^\mu.$$

Now, consider

$$\partial_\nu = \frac{\partial}{\partial x^\nu} = \frac{du}{dx^\nu}\frac{d}{du} = \frac{1}{T^\nu}\frac{d}{du},$$

then, we see that the affine geodesic can be written

$$T^\nu \left(\frac{1}{T^\nu}\frac{d}{du}T^\mu + \Gamma^\mu{}_{\nu\gamma}T^\gamma\right) = \lambda T^\mu.$$

Therefore, noting that $T^\mu = \frac{dx^\mu}{du}$, we see that

$$\frac{d}{du}\frac{dx^\mu}{du} + T^\nu T^\gamma \Gamma^\mu{}_{\nu\gamma} = \lambda T^\mu,$$

which is of course just

$$\frac{d^2x^\mu}{du^2} + \Gamma^\mu{}_{\nu\gamma}\frac{dx^\gamma}{du}\frac{dx^\mu}{du} = \lambda T^\mu. \tag{3.15}$$

As an example, consider a Cartesian system, whereby the affine connections are all zero. The resultant differential equation has a straight line as the solution. That is, the affine geodesic in Cartesian coordinates is a straight line.

We say that u is the affine parameter. If $\lambda = 0$, then we say that the geodesic is affinely parameterised. That is,

$$T^\nu \nabla_\nu T^\mu = 0, \quad T^\mu \equiv \frac{dx^\mu}{du},$$

along an affinely parameterised geodesic.

3.2.2 The Metric Geodesic

This geodesic is perhaps a little less hand-wavey.

Consider two points in some space. Consider that they are joined by a line. Then, the metric geodesic is the line which extremises that joining line. So, given a line element

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

we see that the corresponding action is

$$S = \int ds.$$

Now, considering that the line is parameterised by u, the affine parameter, then we see that the action is simply

$$S = \int ds = \int \frac{ds}{du}du = \int du\sqrt{g_{\mu\nu}\frac{dx^\mu}{du}\frac{dx^\nu}{du}}.$$

Then, by the variational principle, the Euler-Lagrange equation

$$\frac{d}{du}\frac{\partial L}{\partial \dot{x}^\mu} - \frac{\partial L}{\partial x^\mu} = 0, \quad \dot{x}^\mu \equiv \frac{dx^\mu}{du}$$

extremises the action (note, we use the word extremise, rather than maximise or minimise). We must state that $g_{\mu\nu}(x^\rho)$, only.

So, the Lagrangian is

$$L = \left(g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta\right)^{1/2}.$$

We are now left to compute the elements of the EL equation. So,

$$\begin{aligned}
\frac{\partial L}{\partial \dot{x}^\mu} &= \frac{1}{2L}\frac{\partial}{\partial \dot{x}^\mu}\left(g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta\right) \\
&= \frac{1}{2L}g_{\alpha\beta}\left(\frac{\partial \dot{x}^\alpha}{\partial \dot{x}^\mu}\dot{x}^\beta + \dot{x}^\alpha\frac{\partial \dot{x}^\beta}{\partial \dot{x}^\mu}\right) \\
&= \frac{1}{2L}g_{\alpha\beta}\left(\delta^\alpha_\mu\dot{x}^\beta + \delta^\beta_\mu\dot{x}^\alpha\right) \\
&= \frac{1}{2L}g_{\alpha\beta}\left(\delta^\alpha_\mu\dot{x}^\beta + \delta^\alpha_\mu\dot{x}^\beta\right) \\
&= \frac{1}{L}g_{\alpha\beta}\dot{x}^\beta\delta^\alpha_\mu \\
&= \frac{1}{L}g_{\mu\beta}\dot{x}^\beta.
\end{aligned}$$

And also,

$$\frac{\partial L}{\partial x^\mu} = \frac{1}{2L}\dot{x}^\alpha\dot{x}^\beta\partial_\mu g_{\alpha\beta}.$$

Finally,

$$\begin{aligned}
\frac{d}{du}\frac{\partial L}{\partial \dot{x}^\mu} &= \frac{d}{du}\left(\frac{1}{L}g_{\mu\beta}\dot{x}^\beta\right) \\
&= -\frac{\dot{L}}{L^2}g_{\mu\beta}\dot{x}^\beta + \frac{1}{L}\frac{d}{du}\left(g_{\mu\beta}\dot{x}^\beta\right),
\end{aligned}$$

the last expression we evaluate via

$$\begin{aligned}
\frac{d}{du}\left(g_{\mu\beta}\dot{x}^\beta\right) &= g_{\mu\beta}\frac{d\dot{x}^\beta}{du} + \dot{x}^\beta\frac{d}{du}g_{\mu\beta} \\
&= g_{\mu\beta}\ddot{x}^\beta + \dot{x}^\beta\frac{dx^\gamma}{du}\frac{dg_{\mu\beta}}{dx^\gamma} \\
&= g_{\mu\beta}\ddot{x}^\beta + \dot{x}^\beta\dot{x}^\gamma\partial_\gamma g_{\mu\beta}.
\end{aligned}$$

Therefore,

$$\frac{d}{du}\frac{\partial L}{\partial \dot{x}^\mu} = -\frac{\dot{L}}{L^2}g_{\mu\beta}\dot{x}^\beta + \frac{1}{L}\left(g_{\mu\beta}\ddot{x}^\beta + \dot{x}^\beta\dot{x}^\gamma\partial_\gamma g_{\mu\beta}\right);$$

and consequently, the EL equation reads

$$-\frac{\dot{L}}{L^2}g_{\mu\beta}\dot{x}^\beta + \frac{1}{L}\left(g_{\mu\beta}\ddot{x}^\beta + \dot{x}^\beta\dot{x}^\gamma\partial_\gamma g_{\mu\beta}\right) - \frac{1}{2L}\dot{x}^\alpha\dot{x}^\beta\partial_\mu g_{\alpha\beta} = 0.$$

Now, the job is to get this into a "nice form", without mention to L. We can move the first term over to the RHS, and multiply through by L, giving

$$g_{\mu\beta}\ddot{x}^\beta + \dot{x}^\beta\dot{x}^\gamma\partial_\gamma g_{\mu\beta} - \frac{1}{2}\dot{x}^\alpha\dot{x}^\beta\partial_\mu g_{\alpha\beta} = \frac{\dot{L}}{L}g_{\mu\beta}\dot{x}^\beta.$$

Let us now multiply this by something that will kill-off the metric multiplying $\ddot{x}^\beta$, the LHS. Multiplying by $g^{\rho\mu}$ will work,

$$g^{\rho\mu} g_{\mu\beta} \ddot{x}^\beta + g^{\rho\mu} \left(\dot{x}^\beta \dot{x}^\gamma \partial_\gamma g_{\mu\beta} - \frac{1}{2} \dot{x}^\alpha \dot{x}^\beta \partial_\mu g_{\alpha\beta} \right) = \frac{\dot{L}}{L} g^{\rho\mu} g_{\mu\beta} \dot{x}^\beta,$$

noting that $g^{\rho\mu} g_{\mu\beta} = \delta^\rho_\beta$, and using this relation, we see we now have

$$\ddot{x}^\rho + g^{\rho\mu} \left(\dot{x}^\beta \dot{x}^\gamma \partial_\gamma g_{\mu\beta} - \frac{1}{2} \dot{x}^\alpha \dot{x}^\beta \partial_\mu g_{\alpha\beta} \right) = \frac{\dot{L}}{L} \dot{x}^\rho.$$

To continue, we use the simple result that if $a = b$, then $a = \frac{1}{2}(a + b)$. So, we see that

$$\dot{x}^\beta \dot{x}^\gamma \partial_\gamma g_{\mu\beta} = \frac{1}{2} \left(\dot{x}^\beta \dot{x}^\gamma \partial_\gamma g_{\mu\beta} + \dot{x}^\gamma \dot{x}^\beta \partial_\beta g_{\mu\gamma} \right),$$

thus, using this, we see that we have the geodesic equation being

$$\ddot{x}^\rho + g^{\rho\mu} \frac{1}{2} \left(\dot{x}^\beta \dot{x}^\gamma \partial_\gamma g_{\mu\beta} + \dot{x}^\gamma \dot{x}^\beta \partial_\beta g_{\mu\gamma} - \dot{x}^\alpha \dot{x}^\beta \partial_\mu g_{\alpha\beta} \right) = \frac{\dot{L}}{L} \dot{x}^\rho.$$

We can pull out common factors of the bracketed term, by relabeling indices $\alpha \to \gamma$, thus

$$\ddot{x}^\rho + g^{\rho\mu} \frac{1}{2} \left(\partial_\gamma g_{\mu\beta} + \partial_\beta g_{\mu\gamma} - \partial_\mu g_{\gamma\beta} \right) \dot{x}^\beta \dot{x}^\gamma = \frac{\dot{L}}{L} \dot{x}^\rho.$$

Now, by way of convenient notation, we define everything multiplying the $\dot{x}^\beta \dot{x}^\gamma$ as

$$\left\{ {}_\gamma{}^\rho{}_\beta \right\} \equiv g^{\rho\mu} \frac{1}{2} \left(-\partial_\mu g_{\gamma\beta} + \partial_\gamma g_{\mu\beta} + \partial_\beta g_{\mu\gamma} \right), \tag{3.16}$$

a symbol we call the *Christofell symbol*. Thus, the geodesic equation looks like

$$\ddot{x}^\rho + \left\{ {}_\gamma{}^\rho{}_\beta \right\} \dot{x}^\beta \dot{x}^\gamma = \frac{\dot{L}}{L} \dot{x}^\rho.$$

Now, if $\dot{L} = 0$, then this reads

$$\ddot{x}^\rho + \left\{ {}_\alpha{}^\rho{}_\beta \right\} \dot{x}^\alpha \dot{x}^\beta = 0, \quad \frac{d^2 s}{du^2} = 0.$$

Notice that the Christofell symbol is symmetric in its lower indices,

$$\left\{ {}_\gamma{}^\mu{}_\beta \right\} = \left\{ {}_\beta{}^\mu{}_\gamma \right\},$$

which we can see by its definition, noting that the metric is symmetric.

Just to write the result again,

$$\ddot{x}^\rho + \left\{ {}_\alpha{}^\rho{}_\beta \right\} \dot{x}^\alpha \dot{x}^\beta = 0 \tag{3.17}$$

is the *affinely parameterised metric geodesic.*

So, to recap these geodesics. The curve which preserves the direction of the tangent vector on that curve, is called the *affine geodesic*. When deriving the geodesic, one uses the $\Gamma^{\lambda}{}_{\mu\nu}$ symbol, so we call it the affine connection; or just *connection*. The second type of geodesic was derived to be the curve which extremises the path length between two points. This was called the *metric geodesic*. In deriving the metric geodesic, one defines some quantities, the *Christofell symbols*.

3.2.3 Relation Between Affine Connection & Christofell Symbol

We shall start by asserting that for a torsion free connection, $T^{\alpha}{}_{\mu\nu} = 0$, and that for a metric with zero covariant derivative, $\nabla_{\alpha} g_{\nu\mu} = 0$, then the affine connection is the Christofell symbol. That is,

$$\nabla_{\alpha} g_{\mu\nu} = 0, \quad T^{\mu}{}_{\alpha\beta} = 0 \quad \Rightarrow \quad \Gamma^{\mu}{}_{\alpha\beta} = \{{}_{\alpha}{}^{\mu}{}_{\beta}\}.$$

The way to use "torsion free" is that the final two indices on the affine connection can be interchanged. We also use a symmetric metric throughout.

Let us prove it.

We start by writing the covariant derivative of the metric,

$$\nabla_{\alpha} g_{\mu\nu} = \partial_{\alpha} g_{\mu\nu} - \Gamma^{\lambda}{}_{\alpha\mu} g_{\lambda\nu} - \Gamma^{\lambda}{}_{\alpha\nu} g_{\mu\lambda}.$$

But, by our definition of the problem, this is zero. So,

$$\partial_{\alpha} g_{\mu\nu} = \Gamma^{\lambda}{}_{\alpha\mu} g_{\lambda\nu} + \Gamma^{\lambda}{}_{\alpha\nu} g_{\mu\lambda}.$$

Let us now cyclicly change the indices. First, we shall do $\alpha \to \mu \to \nu \to \alpha$. Giving

$$\partial_{\mu} g_{\nu\alpha} = \Gamma^{\lambda}{}_{\mu\nu} g_{\lambda\alpha} + \Gamma^{\lambda}{}_{\mu\alpha} g_{\nu\lambda}.$$

Let us do the interchange again, on this new equation. Giving

$$\partial_{\nu} g_{\alpha\mu} = \Gamma^{\lambda}{}_{\nu\alpha} g_{\lambda\mu} + \Gamma^{\lambda}{}_{\nu\mu} g_{\alpha\lambda}.$$

Let us add the first two, and subtract the last equation. Giving

$$\partial_{\alpha} g_{\mu\nu} + \partial_{\mu} g_{\nu\alpha} - \partial_{\nu} g_{\alpha\mu} = \Gamma^{\lambda}{}_{\alpha\mu} g_{\lambda\nu} + \Gamma^{\lambda}{}_{\alpha\nu} g_{\mu\lambda} + \Gamma^{\lambda}{}_{\mu\nu} g_{\lambda\alpha} + \Gamma^{\lambda}{}_{\mu\alpha} g_{\nu\lambda} - \Gamma^{\lambda}{}_{\nu\alpha} g_{\lambda\mu} - \Gamma^{\lambda}{}_{\nu\mu} g_{\alpha\lambda}.$$

Now, we notice that by our torsion free assert, we can cancel off some of the terms on the RHS. These are the second with the fifth, and third with sixth. This leaves

$$\partial_{\alpha} g_{\mu\nu} + \partial_{\mu} g_{\nu\alpha} - \partial_{\nu} g_{\alpha\mu} = \Gamma^{\lambda}{}_{\alpha\mu} g_{\lambda\nu} + \Gamma^{\lambda}{}_{\mu\alpha} g_{\nu\lambda},$$

from which we further use the torsion-free assert, to see that the two terms on the RHS are identical, leaving

$$\partial_{\alpha} g_{\mu\nu} + \partial_{\mu} g_{\nu\alpha} - \partial_{\nu} g_{\alpha\mu} = 2\Gamma^{\lambda}{}_{\alpha\mu} g_{\lambda\nu}.$$

Rearranging this trivially results in

$$\Gamma^{\lambda}{}_{\alpha\mu} g_{\lambda\nu} = \frac{1}{2}\left(\partial_\alpha g_{\mu\nu} + \partial_\mu g_{\nu\alpha} - \partial_\nu g_{\alpha\mu}\right).$$

Now, let us multiply the whole thing by $g^{\rho\nu}$,

$$\begin{aligned} \Gamma^{\lambda}{}_{\alpha\mu} g^{\rho\nu} g_{\lambda\nu} &= \frac{1}{2} g^{\rho\nu}\left(\partial_\alpha g_{\mu\nu} + \partial_\mu g_{\nu\alpha} - \partial_\nu g_{\alpha\mu}\right) \\ \Rightarrow \quad \Gamma^{\lambda}{}_{\alpha\mu} \delta^{\rho}_{\lambda} &= \frac{1}{2} g^{\rho\nu}\left(\partial_\alpha g_{\mu\nu} + \partial_\mu g_{\nu\alpha} - \partial_\nu g_{\alpha\mu}\right), \end{aligned}$$

which is just

$$\Gamma^{\rho}{}_{\alpha\mu} = \frac{1}{2} g^{\rho\nu}\left(\partial_\alpha g_{\mu\nu} + \partial_\mu g_{\nu\alpha} - \partial_\nu g_{\alpha\mu}\right).$$

Now, if we switch over the μ index to β (just relabelling),

$$\Gamma^{\rho}{}_{\alpha\beta} = \frac{1}{2} g^{\rho\nu}\left(\partial_\alpha g_{\beta\nu} + \partial_\beta g_{\nu\alpha} - \partial_\nu g_{\alpha\beta}\right).$$

Upon inspection of this with (3.16), we find that they are equal. Therefore,

$$\Gamma^{\rho}{}_{\alpha\beta} = \{{}_{\alpha}{}^{\rho}{}_{\beta}\}; \quad \nabla_\alpha g_{\mu\nu} = 0, \quad T^{\mu}{}_{\alpha\beta} = 0.$$

It is very important to note that this only holds for a torsion free connection, with metric having zero covariant derivative. Under these conditions, the affine connection is the Christofell symbol.

3.3 Isometries & Killing's Equation

Consider the coordinate transformation

$$g_{\mu\nu}(x) \longmapsto g_{\mu\nu}(x'),$$

so that the metric in the new frame has the same functional dependance as in the old frame. That is, the new metric depends on x' in the same way as the old metric depended on x. Then, we have

$$ds^2(x) = ds^2(x'),$$

and that

$$g'_{\mu\nu}(x') = g_{\mu\nu}(x). \tag{3.18}$$

Therefore, by the transformation rule of the metric,

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

and using (3.18) on the RHS, we see that

$$g_{\mu\nu}(x) = \frac{\partial x'^{\alpha}}{\partial x^{\mu}} \frac{\partial x'^{\beta}}{\partial x^{\nu}} g_{\alpha\beta}(x'). \tag{3.19}$$

So, a coordinate transformation leaving the metric in the same form (form invariant), is called an *isometry*.

The coordinate transformation we considered was $x^\mu \mapsto x'^\mu$. Let us consider a special case of this; namely

$$x^\mu \longmapsto x'^\mu = x^\mu + \epsilon \xi^\mu,$$

where ϵ is small, and ξ^μ a vector field. Now, the Jacobian,

$$J^\mu{}_\nu = \partial_\nu x'^\mu = \partial_\nu \left(x^\mu + \epsilon \xi^\mu \right),$$

which is clearly

$$J^\mu{}_\nu = \delta^\mu_\nu + \epsilon \partial_\nu \xi^\mu.$$

Now, by a Taylor expansion, we see that

$$g_{\alpha\beta}(x') = g_{\alpha\beta}(x^\mu + \epsilon\xi^\mu) = g_{\alpha\beta}(x^\mu) + \epsilon \xi^\mu \partial_\mu g_{\alpha\beta}(x^\mu) + \mathcal{O}(\epsilon^2).$$

So, we now have enough terms to be able to put them all into the metric isometry transformation equation (3.19). Thus,

$$\begin{aligned} g_{\mu\nu}(x) &= J^\alpha{}_\mu J^\beta{}_\nu g_{\alpha\beta}(x') \\ &= \left(\delta^\alpha_\mu + \epsilon \partial_\mu \xi^\alpha \right) \left(\delta^\beta_\nu + \epsilon \partial_\nu \xi^\beta \right) \left(g_{\alpha\beta}(x) + \epsilon \xi^\rho \partial_\rho g_{\alpha\beta}(x) \right). \end{aligned}$$

if we expand out the RHS, neglecting terms in $\mathcal{O}(\epsilon^2)$, one finds that

$$g_{\mu\nu}(x) = g_{\mu\nu}(x) + \epsilon \xi^\rho \partial_\rho g_{\mu\nu} + \epsilon g_{\mu\beta} \partial_\nu \xi^\beta + \epsilon g_{\alpha\nu} \partial_\mu \xi^\alpha,$$

rearranging,

$$g_{\mu\nu}(x) = g_{\mu\nu}(x) + \epsilon \left(g_{\mu\beta} \partial_\nu \xi^\beta + g_{\alpha\nu} \partial_\mu \xi^\alpha + \xi^\rho \partial_\rho g_{\mu\nu} \right),$$

which is obviously just

$$g_{\mu\beta} \partial_\nu \xi^\beta + g_{\alpha\nu} \partial_\mu \xi^\alpha + \xi^\rho \partial_\rho g_{\mu\nu} = 0. \tag{3.20}$$

Now, notice that

$$\xi_\alpha = g_{\alpha\beta} \xi^\beta,$$

and then its differential is

$$\begin{aligned} \partial_\nu \xi_\alpha &= \partial_\nu \left(g_{\alpha\beta} \xi^\beta \right) \\ &= \xi^\beta \partial_\nu g_{\alpha\beta} + g_{\alpha\beta} \partial_\nu \xi^\beta. \end{aligned}$$

Hence, we can rearrange this into the form

$$g_{\alpha\beta}\partial_\nu \xi^\beta = \partial_\nu \xi_\alpha - \xi^\beta \partial_\nu g_{\alpha\beta}.$$

So, if we put this into (3.20) for the first and second expressions (being very careful in changing indices), we get

$$\partial_\nu \xi_\mu - \xi^\beta \partial_\nu g_{\mu\beta} + \partial_\mu \xi_\nu - \xi^\beta \partial_\mu g_{\nu\beta} + \xi^\beta \partial_\beta g_{\mu\nu} = 0.$$

Collecting terms,

$$\partial_\nu \xi_\mu + \partial_\mu \xi_\nu + \xi^\beta \left(\partial_\beta g_{\mu\nu} - \partial_\nu g_{\mu\beta} - \partial_\mu g_{\nu\beta}\right) = 0$$

The bracketed quantity is just $-2g_{\rho\beta}\Gamma^\rho{}_{\mu\nu}$, so that

$$\partial_\nu \xi_\mu + \partial_\mu \xi_\nu - 2\xi^\beta g_{\rho\beta}\Gamma^\rho{}_{\mu\nu} = 0,$$

which is just,

$$\partial_\nu \xi_\mu + \partial_\mu \xi_\nu - 2\Gamma^\rho{}_{\mu\nu}\xi_\rho = 0.$$

This is just the covariant derivative (noting the symmetry of the Christofell symbols),

$$\nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 0. \tag{3.21}$$

This is known as *Killing's equation*. A vector ξ^ν satisfying Killing's equation is called a *Killing vector*.

Let us just recap what we have done. A metric is said to have an isometry if it can transform, retaining its functional dependence. Then, under a small coordinate transformation, with a vector field ξ^μ, the field satisfying Killing's equation will give an isometry.

Now, a theorem states that, for a tangent and Killing vector, T^μ, ξ^μ respectively, there is a conserved quantity $T^\mu \xi_\mu$ along an affinely parameterised geodesic. So, to prove it, we consider

$$\frac{D}{Du}(T^\mu \xi_\mu) = T^\nu \nabla_\nu (T^\mu \xi_\mu) = T^\nu \left(\xi_\mu \nabla_\nu T^\mu + T^\mu \nabla_\nu \xi_\mu\right).$$

Now, the first term is zero, as we are on an affinely parameterised geodesic. Now, notice that we can write the final term as

$$\nabla_\nu \xi_\mu = \frac{1}{2}(\nabla_\nu \xi_\mu + \nabla_\mu \xi_\nu),$$

and thus we have

$$\frac{D}{Du}(T^\mu \xi_\mu) = T^\nu T^\mu \frac{1}{2}(\nabla_\nu \xi_\mu + \nabla_\mu \xi_\nu).$$

We were able to interchange the indices (with the factor of one-half to cancel out the double counting), because the things multiplying it are symmetric under interchange of indices. Notice that the bracketed term is just Killing's equation, for some Killing vector ξ^ν. Therefore,

$$\frac{D}{Du}(T^\mu \xi_\mu) = 0,$$

thus, $T^\mu \xi_\mu$ is some conserved quantity along an affinely parameterised geodesic.

3.4 Summary

We shall soon see some examples of geodesics, and what a Killing vector corresponds to; but before then we shall bring together our definitions of the Christofell symbol, and introduce a little new notation (just to be inkeeping with the literature).

We have that the affine connection, $\Gamma^{\mu}{}_{\nu\lambda}$ is the same as the geodesic connection $\{{}_{\mu}{}^{\alpha}{}_{\nu}\}$, for manifolds for whom

$$\nabla_{\alpha} g_{\mu\nu} = 0, \quad T^{\alpha}{}_{\mu\nu} = 0 \quad \Rightarrow \quad \Gamma^{\alpha}{}_{\mu\nu} = \{{}_{\mu}{}^{\alpha}{}_{\nu}\}.$$

We also derived that the relation between the Christofell symbol (as we may as well call it), and the metric, is

$$\Gamma^{\rho}{}_{\alpha\mu} = \frac{1}{2} g^{\rho\nu} \left(-\partial_{\nu} g_{\alpha\mu} + \partial_{\alpha} g_{\mu\nu} + \partial_{\mu} g_{\nu\alpha} \right).$$

Infact, $\Gamma^{\rho}{}_{\alpha\mu}$ are generally denoted *Christofell symbols of the second kind.* We can infact see that

$$\Gamma^{\rho}{}_{\alpha\mu} = g^{\rho\nu} \Gamma_{\nu\alpha\mu},$$

where we call the $\Gamma_{\nu\alpha\mu}$ the *Christofell symbols of the first kind.* When we refer to the "Christofell symbols", we shall mean those of the second kind.

We derived that the affine geodesic is the same as the metric geodesic, for affinely parameterised geodesics (satisfying the above torsion & covariant derivative relations).

We also saw that the Christofell symbols are not tensors. The non-tensorial nature of the symbols allowed us to derive that there exists a point in a manifold, where all components of the symbol are zero. That is, there exists a point where the manifold is flat.

The Lagrangian squared,

$$L^2 = \left(\frac{ds}{du} \right)^2 = g_{\mu\nu} \dot{x}^{\mu} \dot{x}^{\nu},$$

is just the line element length. Its possible values are $0, \pm 1$. We classify 0 as *null geodesics*, $+1$ as *time-like* and -1 as *space-like.*

Also by way of being inkeeping with the literature, some books denote partial & covariant derivatives in a different way. Sometimes one may see

$$\partial_{\nu} A_{\mu} \equiv A_{\nu,\mu}, \quad \nabla_{\nu} A_{\mu} \equiv A_{\nu;\mu}.$$

That is, a "comma" representing partial derivatives, and a "semi-colon" for covariant derivatives.

3.5 Examples

Here we shall see specific examples of geodesics, Killing vectors & how to compute Christofell symbols.

3.5.1 Computing Christofell Symbols: Effective Lagrangian

Now, before we go onto the effective Lagrangian method of computing the Christofell symbols, we shall see how to do so, via brute force.

Brute force: plane polars In plane polars, we have the line element

$$ds^2 = dr^2 + r^2 d\phi^2,$$

and therefore, reading off the components of the metric & inverse

$$(g_{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}, \quad (g^{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & 1/r^2 \end{pmatrix}.$$

Then, using the notation that $ds^i = (dr, d\phi)$, we see that $g_{rr} = 1, g_{\phi\phi} = r^2, g^{rr} = 1, g^{\phi\phi} = r^{-2}$ are the only non-zero components. So, to compute the Christofell symbols (the brute force way), we must find

$$\Gamma^i{}_{jk} = \frac{1}{2}\sum_a g^{ia}\left(-\partial_a g_{jk} + \partial_j g_{ak} + \partial_k g_{ja}\right), \quad i, j, k, a = r, \phi.$$

We shall spell out, in detail, how to compute one of the components;

$$\begin{aligned} \Gamma^r{}_{\phi\phi} &= \frac{1}{2}\sum_{a=r,\phi} g^{ra}\left(-\partial_a g_{\phi\phi} + \partial_\phi g_{\phi a} + \partial_\phi g_{\phi a}\right) \\ &= \frac{1}{2}\left[g^{rr}\left(-\partial_r g_{\phi\phi} + \partial_\phi g_{\phi r} + \partial_\phi g_{\phi r}\right) + g^{r\phi}\left(-\partial_\phi g_{\phi\phi} + \partial_\phi g_{\phi\phi} + \partial_\phi g_{\phi\phi}\right)\right]. \end{aligned}$$

Now, one of the first things we note, is that the metric is diagonal: all off-diagonal components are zero. So, the above reduces to

$$\begin{aligned} \Gamma^r{}_{\phi\phi} &= -\frac{1}{2} g^{rr}\partial_r g_{\phi\phi} \\ &= -\frac{1}{2}.1.\frac{\partial}{\partial r}r^2 \\ &= -r. \end{aligned}$$

We have thus found one of the components of the Christofell symbol. We shall state the rest of them (as going through how to find each one is very tedious).

$$\begin{aligned} &\Gamma^r{}_{rr} = 0, \quad \Gamma^r{}_{r\phi} = \Gamma^r{}_{\phi r} = 0, \quad \Gamma^r{}_{\phi\phi} = -r, \\ &\Gamma^\phi{}_{\phi\phi} = 0, \quad \Gamma^\phi{}_{r\phi} = \Gamma^\phi{}_{\phi r} = r^{-1}, \quad \Gamma^\phi{}_{rr} = 0. \end{aligned}$$

Now, we shall show how to find them in a more intelligent manner.

Effective Lagrangian Method When we derived the metric geodesic, we had that the Lagrangian was

$$L = \sqrt{\frac{ds}{du}}.$$

Now, consider

$$L_{\text{eff}} \equiv L^2.$$

The Euler-Lagrange equation for L_{eff} is

$$\frac{d}{du}\left(\frac{\partial L_{\text{eff}}}{\partial \dot{x}^\mu}\right) - \frac{\partial L_{\text{eff}}}{\partial x^\mu} = 0,$$

from which it is clear to see that

$$2L\left[\frac{d}{du}\left(\frac{\partial L}{\partial \dot{x}^\mu}\right) - \frac{\partial L}{\partial x^\mu}\right] = 0.$$

Thus, the if L satisfies the Euler-Lagrange equation, then so does L^2. This makes life a lot simpler, as we can consider just $g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$, rather than its square-root.

So, for plane polars, where

$$L_{\text{eff}} = L^2 = \dot{r}^2 + r^2\dot{\phi}^2,$$

we have two Euler-Lagrange equations, one for each coordinate r, ϕ. They are

$$\ddot{r} - r\dot{\phi}^2 = 0, \quad 2\dot{r}\dot{\phi} + r\ddot{\phi} = 0.$$

Now, if we get these equations into the form $\ddot{x} + C\dot{x}^1\dot{x}^2 = 0$,

$$\ddot{r} - r\dot{\phi}^2 = 0, \quad \ddot{\phi} + \frac{\dot{r}\dot{\phi}}{r} + \frac{\dot{\phi}\dot{r}}{r} = 0.$$

So, we see that we can read off the Christofell symbols, by inspection. To see this a little clearer, the "general" metric geodesic, for r, is

$$\ddot{r} + \sum_{i,j=r,\phi} \Gamma^r{}_{ij}\dot{x}^i\dot{x}^j = 0;$$

then, we can see that the only Christofell components that is non-zero is that where $i = j = \phi$, and that value is $-r$. Thus, we read off that $\Gamma^r{}_{\phi\phi} = -r$, which is in accord to what we had by the brute force method. For ϕ, the general geodesic is

$$\ddot{\phi} + \sum_{i,j=r,\phi} \Gamma^\phi{}_{ij}\dot{x}^i\dot{x}^j = 0;$$

and we therefore see two non-zero Christofell symbols: when $i = r, j = \phi$ and $i = \phi, j = r$. The corresponding Christofell symbols components are thus $\Gamma^\phi{}_{r\phi} = \Gamma^\phi{}_{\phi r} = r^{-1}$. Again, in accord with the brute force components.

3.5.2 Computing the Geodesic

Now, we are able to find the geodesic: a parameterised curve that extremises the distance between two points, in the plane polar coordinate system.

When we computed the Euler-Lagrange equation for ϕ, we had a term (which we didnt state above, but is easy to see, upon computation)

$$\frac{d}{du}(2r^2\dot{\phi}) = 0 \quad \Rightarrow \quad r^2\dot{\phi} = B = \text{const.}$$

That is, $\dot{\phi} = B/r^2$. Now, the effective Lagrangian is just ds^2/du^2, which is just the line element, which can be one of 3 values (as previously stated),

$$L_{\text{eff}} = L^2 = \begin{cases} 0 \\ 1 \\ -1 \end{cases} \equiv A.$$

So, the effective Lagrangian is just

$$L_{\text{eff}} = \dot{r}^2 + r^2\dot{\phi}^2 = A.$$

Hence, using our expression for $\dot{\phi}$,

$$\dot{r}^2 + r^2\frac{B^2}{r^4} = A \quad \Rightarrow \quad \dot{r} = \sqrt{A - \frac{B^2}{r^2}}.$$

Now, if we notice that

$$\frac{\dot{\phi}}{\dot{r}} = \frac{d\phi}{dr} = \frac{B}{r^2}\left(A - \frac{B^2}{r^2}\right)^{-1/2},$$

then this integrates to

$$r\cos(\phi - \phi_c) = \frac{B}{\sqrt{A}}.$$

If we take $A = 1$, so that we are talking about time-like geodesics, then the equation becomes

$$r\cos(\phi - \phi_c) = r\cos\phi\cos\phi_c + r\sin\phi\sin\phi_c = B.$$

We now note that ϕ_c is a constant, $x = r\cos\phi, y = r\sin\phi$, giving

$$mx + ty = B \quad \Rightarrow \quad y = mx + c.$$

That is, the time-like geodesic is a straight line.

Let us now consider the null geodesic. We appeal back to the effective Lagrangian, which becomes

$$\dot{r}^2 + r^2\dot{\phi}^2 = 0.$$

This has solution $\dot{r} = 0$ and $\dot{\phi} = 0$. That is, both radius & angle are constants. That is, a single point. Thus, the null geodesic is a point (null size).

When we consider the space-like geodesic, we find that there is no solution: it does not exist in plane polars.

Example of Geodesic 2 Let us compute another geodesic, for another line element,

$$ds^2 = \frac{1}{t^2}(dt^2 - dx^2).$$

So, we see that the effective Lagrangian is

$$L_{\text{eff}} = \frac{\dot{t}^2 - \dot{x}^2}{t^2}, \quad \dot{t} \equiv \frac{dt}{du}, \dot{x} \equiv \frac{dx}{du}.$$

Then, the Euler-Lagrange equations for this effective Lagrangian are

$$\ddot{t} - \frac{\dot{t}^2}{t} - \frac{\dot{x}^2}{t} = 0, \quad \ddot{x} - \frac{2}{t}\dot{x}\dot{t} = 0.$$

So, we can read off the Christofell symbol components. The only non-zero components are

$$\Gamma^t{}_{xx} = \Gamma^t{}_{tt} = -\frac{1}{t}, \quad \Gamma^x{}_{xt} = \Gamma^x{}_{tx} = -\frac{1}{t}.$$

3.5.3 Physical Meaning of the Killing Vector

Again, let us go back to plane polars. The line element is

$$ds^2 = dr^2 + r^2 d\phi^2.$$

We would like to think of a vector that leaves the line element unchanged. A transformation on ϕ works:

$$\phi \longmapsto \phi' = \phi + \epsilon,$$

so that

$$(r, \phi) \longmapsto (r', \phi') = (r, \phi + \epsilon) = (r, \phi) + \epsilon(0, 1).$$

Therefore, our Killing vector is

$$\xi^i = (0, 1).$$

Now, we stated that $T^i \xi_i$ is a conserved quantity. Let us consider what it is. So,

$$\begin{aligned}
\xi^i \dot{x}_i &= g_{ij} \xi^i \dot{x}^j \\
&= g_{\phi\phi} \xi^\phi \dot{x}^\phi \\
&= r^2 .1. \dot{\phi} \\
&= r^2 \dot{\phi}.
\end{aligned}$$

This quantity is a constant (as it is conserved). We also notice that it is the expression for the angular momentum of the system. Therefore, the conserved quantity associated with the Killing vector is the angular momentum, in plane polars.

3.5.4 Nordstrom's Theory of Gravity

Let us compute the connection associated with $\hat{g}_{\mu\nu} = \Omega^2 g_{\mu\nu}$. Now, the connection associated with $g_{\mu\nu}$ is

$$\Gamma^{\rho}{}_{\alpha\beta} = \frac{1}{2} g^{\rho\nu} \left(\partial_\alpha g_{\beta\nu} + \partial_\beta g_{\nu\alpha} - \partial_\nu g_{\alpha\beta} \right).$$

Hence,

$$\begin{aligned} \partial_\alpha \left(\hat{g}_{\mu\nu} \right) &= \partial_\alpha \left(\Omega^2 g_{\mu\nu} \right) \\ &= g_{\mu\nu} 2\Omega \partial_\alpha \Omega + \Omega^2 \partial_\alpha g_{\mu\nu}. \end{aligned}$$

Therefore,

$$\begin{aligned} \hat{\Gamma}^{\rho}{}_{\alpha\beta} &= \frac{1}{2} \hat{g}^{\rho\nu} \left(\partial_\alpha \hat{g}_{\beta\nu} + \partial_\beta \hat{g}_{\nu\alpha} - \partial_\nu \hat{g}_{\alpha\beta} \right) \\ &= \frac{1}{2} \frac{1}{\Omega^2} g^{\rho\nu} \big(2\Omega g_{\beta\nu} \partial_\alpha \Omega + \Omega^2 \partial_\alpha g_{\beta\nu} + 2\Omega g_{\nu\alpha} \partial_\beta \Omega + \\ &\quad \Omega^2 \partial_\beta g_{\nu\alpha} - 2\Omega g_{\alpha\beta} \partial_\nu \Omega - \Omega^2 \partial_\nu g_{\alpha\beta} \big) \\ &= \frac{1}{2} g^{\rho\nu} \left(\partial_\alpha g_{\beta\nu} + \partial_\beta g_{\nu\alpha} - \partial_\nu g_{\alpha\beta} \right) + \frac{1}{\Omega} g^{\rho\nu} \left(g_{\beta\nu} \partial_\alpha \Omega + g_{\nu\alpha} \partial_\beta \Omega - g_{\alpha\beta} \partial_\nu \Omega \right) \\ &= \Gamma^{\rho}{}_{\alpha\beta} + \frac{1}{\Omega} \left(\delta^\rho_\beta \partial_\alpha \Omega + \delta^\rho_\alpha \partial_\beta \Omega - g^{\rho\nu} g_{\alpha\beta} \partial_\nu \Omega \right) \end{aligned}$$

Hence,

$$\hat{\Gamma}^{\rho}{}_{\alpha\beta} = \Gamma^{\rho}{}_{\alpha\beta} + \frac{1}{\Omega} \left(\delta^\rho_\beta \partial_\alpha \Omega + \delta^\rho_\alpha \partial_\beta \Omega - g^{\rho\nu} g_{\alpha\beta} \partial_\nu \Omega \right).$$

Let us suppose that we have

$$\hat{g}_{\mu\nu} = e^{2\phi} \eta_{\mu\nu},$$

so that

$$\Omega = e^{\phi}, \quad g_{\mu\nu} = \eta_{\mu\nu}.$$

Hence,

$$\partial_\alpha \Omega = e^{\phi} \partial_\alpha \phi, \quad \Gamma^{\mu}{}_{\alpha\beta} = 0.$$

Hence, using these,

$$\begin{aligned} \hat{\Gamma}^{\rho}{}_{\alpha\beta} &= e^{-\phi} \left(\delta^\rho_\beta e^{\phi} \partial_\alpha \phi + \delta^\rho_\alpha e^{\phi} \partial_\beta \phi - \eta^{\rho\nu} \eta_{\alpha\beta} e^{\phi} \partial_\nu \phi \right) \\ &= \delta^\rho_\beta \partial_\alpha \phi + \delta^\rho_\alpha \partial_\beta \phi - \eta^{\rho\nu} \eta_{\alpha\beta} \partial_\nu \phi. \end{aligned}$$

So, the geodesic equation, with this connection, reads

$$\ddot{x}^\rho + \left(\delta^\rho_\beta \partial_\alpha \phi + \delta^\rho_\alpha \partial_\beta \phi - \eta^{\rho\nu} \eta_{\alpha\beta} \partial_\nu \phi \right) \dot{x}^\alpha \dot{x}^\beta = 0.,$$

which reduces to

$$\ddot{x}^\rho + \dot{x}^\rho 2 \partial_\alpha \phi \dot{x}^\alpha - \dot{x}^2 \partial^\rho \phi = 0.$$

Now, null geodesics have $\dot{x}^2 = 0$, so that this geodesics null value is

$$\ddot{x}^\rho + \dot{x}^\rho 2\partial_\alpha \phi \dot{x}^\alpha = 0.$$

Similarly, timelike geodesics have $\dot{x}^2 = 1$, so that this geodesics timelike value is

$$\ddot{x}^\rho + 2\partial_\alpha \phi \dot{x}^\alpha - \partial^\rho \phi = 0.$$

Now, this example provides us with some practice with using tensors & computing geodesics. In addition to this, we have found the geodesics for a theory whereby the metric is given by $e^{2\phi}\eta_{\mu\nu}$. This theory was proposed by Nordstrom before Einstein.

4 Curvature

We have now got enough mathematical tools to be able to consider the curvature of a manifold.

To continue, consider the commutator of covariant derivatives, acting upon a scalar,

$$[\nabla_\mu, \nabla_\nu]\,\phi = \nabla_\mu \nabla_\nu \phi - \nabla_\nu \nabla_\mu \phi.$$

Now, as we previously showed, the covariant derivative of a scalar is just the normal partial derivative. Therefore,

$$[\nabla_\mu, \nabla_\nu]\,\phi = \nabla_\mu(\partial_\nu \phi) - \nabla_\nu(\partial_\mu \phi).$$

We can now expand out the covariant derivatives. So,

$$\nabla_\mu(\partial_\nu \phi) = \partial_\mu \partial_\nu \phi - \Gamma^\lambda{}_{\mu\nu}\partial_\lambda \phi.$$

Therefore, the commutator is

$$[\nabla_\mu, \nabla_\nu]\,\phi = \partial_\mu \partial_\nu \phi - \Gamma^\lambda{}_{\mu\nu}\partial_\lambda \phi - \partial_\nu \partial_\mu \phi + \Gamma^\lambda{}_{\nu\mu}\partial_\lambda \phi.$$

In a torsion free manifold, the two Christofell terms cancel out, as do the partial derivative terms (as they commute naturally). Therefore, we see that

$$[\nabla_\mu, \nabla_\nu]\,\phi = 0.$$

So, the commutator of partial derivatives, acting upon a scalar, is zero. This result isn't perhaps that surprising. So, let us consider the commutator acting upon a vector.

4.1 The Riemann Tensor

As previously stated, we shall compute the commutator of covariant derivatives, acting upon a vector. That is,

$$[\nabla_\mu, \nabla_\nu]\,A^\rho = \nabla_\mu \nabla_\nu A^\rho - \nabla_\nu \nabla_\mu A^\rho.$$

Now, before, we expanded out the inner covariant derivatives first (as they resulted in just partial derivatives). However, if we do that this time, we will end up having to compute the covariant derivative of the Christofell symbol, which we don't know how to do. Hence, we expand out the outer derivatives first. So,

$$\nabla_\mu \nabla_\nu A^\rho = \partial_\mu(\nabla_\nu A^\rho) + \Gamma^\rho{}_{\mu\lambda}\nabla_\nu A^\lambda - \Gamma^\lambda{}_{\mu\nu}\nabla_\lambda A^\rho,$$

thus, the commutator reads

$$\begin{aligned}[\nabla_\mu, \nabla_\nu]\,A^\rho = \quad & \partial_\mu(\nabla_\nu A^\rho) + \Gamma^\rho{}_{\mu\lambda}\nabla_\nu A^\lambda - \Gamma^\lambda{}_{\mu\nu}\nabla_\lambda A^\rho \\ & -\partial_\nu(\nabla_\mu A^\rho) - \Gamma^\rho{}_{\nu\lambda}\nabla_\mu A^\lambda + \Gamma^\lambda{}_{\nu\mu}\nabla_\lambda A^\rho.\end{aligned}$$

So, we see that the final terms on the RHS cancel (i.e. third & sixth),

$$\begin{aligned}[\nabla_\mu, \nabla_\nu] A^\rho = \quad & \partial_\mu(\nabla_\nu A^\rho) + \Gamma^\rho{}_{\mu\lambda}\nabla_\nu A^\lambda \\ & -\partial_\nu(\nabla_\mu A^\rho) - \Gamma^\rho{}_{\nu\lambda}\nabla_\mu A^\lambda.\end{aligned}$$

Now, expanding out the remaining covariant derivatives,

$$\begin{aligned}[\nabla_\mu, \nabla_\nu] A^\rho = \quad & \partial_\mu\left(\partial_\nu A^\rho + \Gamma^\rho{}_{\lambda\nu}A^\lambda\right) + \Gamma^\rho{}_{\mu\lambda}\left(\partial_\nu A^\lambda + \Gamma^\lambda{}_{\nu\beta}A^\beta\right) \\ & -\partial_\nu\left(\partial_\mu A^\rho + \Gamma^\rho{}_{\lambda\mu}A^\lambda\right) - \Gamma^\rho{}_{\nu\lambda}\left(\partial_\mu A^\lambda + \Gamma^\lambda{}_{\mu\beta}A^\beta\right).\end{aligned}$$

Now, as partial derivatives commute, the two terms on the far LHS cancel. So, cancelling & expanding out the brackets, we have

$$\begin{aligned}[\nabla_\mu, \nabla_\nu] A^\rho = \quad & (\partial_\mu\Gamma^\rho{}_{\lambda\nu})A^\lambda + \Gamma^\rho{}_{\lambda\nu}\partial_\mu A^\lambda + \Gamma^\rho{}_{\mu\lambda}\partial_\nu A^\lambda + \Gamma^\rho{}_{\mu\lambda}\Gamma^\lambda{}_{\nu\beta}A^\beta \\ & -(\partial_\nu\Gamma^\rho{}_{\lambda\mu})A^\lambda - \Gamma^\rho{}_{\lambda\mu}\partial_\nu A^\lambda - \Gamma^\rho{}_{\nu\lambda}\partial_\mu A^\lambda - \Gamma^\rho{}_{\nu\lambda}\Gamma^\lambda{}_{\mu\beta}A^\beta.\end{aligned}$$

Now, we see that the second term cancels with the seventh, and the third with the sixth (again, by assuming torsion free manifolds). Leaving us with

$$[\nabla_\mu, \nabla_\nu] A^\rho = (\partial_\mu\Gamma^\rho{}_{\lambda\nu})A^\lambda + \Gamma^\rho{}_{\mu\lambda}\Gamma^\lambda{}_{\nu\beta}A^\beta - (\partial_\nu\Gamma^\rho{}_{\lambda\mu})A^\lambda - \Gamma^\rho{}_{\nu\lambda}\Gamma^\lambda{}_{\mu\beta}A^\beta.$$

Now, in the second & fourth terms, let us interchange $\beta \leftrightarrow \lambda$, giving

$$[\nabla_\mu, \nabla_\nu] A^\rho = (\partial_\mu\Gamma^\rho{}_{\lambda\nu})A^\lambda + \Gamma^\rho{}_{\mu\beta}\Gamma^\beta{}_{\nu\lambda}A^\lambda - (\partial_\nu\Gamma^\rho{}_{\lambda\mu})A^\lambda - \Gamma^\rho{}_{\nu\beta}\Gamma^\beta{}_{\mu\lambda}A^\lambda,$$

so that we can take out a common factor of A^λ,

$$[\nabla_\mu, \nabla_\nu] A^\rho = \left(\partial_\mu\Gamma^\rho{}_{\lambda\nu} + \Gamma^\rho{}_{\mu\beta}\Gamma^\beta{}_{\nu\lambda} - \partial_\nu\Gamma^\rho{}_{\lambda\mu} - \Gamma^\rho{}_{\nu\beta}\Gamma^\beta{}_{\mu\lambda}\right) A^\lambda.$$

Now, we define the bracketed quantity to be the *Riemann tensor*,

$$R^\rho{}_{\lambda\mu\nu} \equiv \partial_\mu\Gamma^\rho{}_{\lambda\nu} + \Gamma^\rho{}_{\mu\beta}\Gamma^\beta{}_{\nu\lambda} - \partial_\nu\Gamma^\rho{}_{\lambda\mu} - \Gamma^\rho{}_{\nu\beta}\Gamma^\beta{}_{\mu\lambda}, \tag{4.1}$$

so that the commutator reads

$$[\nabla_\mu, \nabla_\nu] A^\rho = R^\rho{}_{\lambda\mu\nu}A^\lambda. \tag{4.2}$$

The Riemann tensor is a $\binom{1}{3}$-tensor. It is clear that $R^\rho{}_{\lambda\mu\nu}$ is a tensor, as the LHS of the above is a tensor, the RHS must also be (as A^ρ is a tensor). This obviously not a rigorous proof of the tensorial nature of the Riemann "tensor", so we shall prove it.

We have

$$(\nabla_\mu\nabla_\nu - \nabla_\nu\nabla_\mu) A^\rho = R^\rho{}_{\lambda\mu\nu}A^\lambda,$$

and therefore that

$$\left(\nabla'_\mu\nabla'_\nu - \nabla'_\nu\nabla'_\mu\right) A'^\rho = R'^\rho{}_{\lambda\mu\nu}A'^\lambda.$$

Now, as covariant derivatives are tensors, we know that

$$\nabla'_\mu \nabla'_\nu A'^\rho = \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu J^\rho{}_\gamma \nabla_\alpha \nabla_\beta A^\gamma.$$

Hence,

$$\left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu J^\rho{}_\gamma \left(\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha\right) A^\gamma = R'^\rho{}_{\lambda\mu\nu} J^\lambda{}_\pi A^\pi.$$

Now, on the LHS, we see that $\left(\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha\right) A^\gamma = R^\gamma{}_{\sigma\alpha\beta} A^\sigma$. Therefore,

$$\left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu J^\rho{}_\gamma R^\gamma{}_{\sigma\alpha\beta} A^\sigma = R'^\rho{}_{\lambda\mu\nu} J^\lambda{}_\pi A^\pi.$$

Now, multiplying through by something that will 'kill off' the Jacobian on the RHS, $\left(J^{-1}\right)^\delta{}_\lambda$ for example,

$$\left(J^{-1}\right)^\delta{}_\lambda \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu J^\rho{}_\gamma R^\gamma{}_{\sigma\alpha\beta} A^\sigma = R'^\rho{}_{\lambda\mu\nu} \delta^\delta_\pi A^\pi = R'^\rho{}_{\lambda\mu\nu} A^\delta.$$

Now, as this must be valid for all A^μ, we must set $\delta = \sigma$. Therefore, doing so & canceling off the A^μ,

$$\left(J^{-1}\right)^\sigma{}_\lambda \left(J^{-1}\right)^\alpha{}_\mu \left(J^{-1}\right)^\beta{}_\nu J^\rho{}_\gamma R^\gamma{}_{\sigma\alpha\beta} = R'^\rho{}_{\lambda\mu\nu},$$

which is the transformation rule of a $\binom{1}{3}$-tensor. Therefore, we have proven that the Riemann tensor is infact a tensor.

Just to be in-keeping with some literature, the Riemann tensor is also called the Riemann-Christofell tensor, or the curvature tensor.

4.1.1 Symmetries of the Riemann Tensor

Now, in one of our previous discussions, we introduced the local inertial frame (LIF), whereby at a point $x^\mu = x^\mu_*$, the metric is flat, and the Christofell symbols are all zero;

$$g_{\mu\nu}(x_*) = \eta_{\mu\nu}, \quad \partial_\rho g_{\mu\nu}(x_*) = 0, \quad \Gamma^\rho{}_{\mu\nu}(x_*) = 0.$$

In a LIF, the Riemann tensor looks quite simple. So, we see that the Riemann tensor, in a LIF, is just

$$R^\rho{}_{\lambda\mu\nu} = \partial_\mu \Gamma^\rho{}_{\lambda\nu} - \partial_\nu \Gamma^\rho{}_{\lambda\mu}.$$

Putting in expressions for the Christofell symbols, and noting that the first derivative of the metric is zero;

$$R^\rho{}_{\lambda\mu\nu} = \frac{1}{2} g^{\rho\pi} \left(\partial_\mu \partial_\lambda g_{\nu\pi} + \partial_\mu \partial_\nu g_{\pi\lambda} - \partial_\mu \partial_\pi g_{\lambda\nu} - \partial_\nu \partial_\lambda g_{\mu\pi} - \partial_\nu \partial_\mu g_{\pi\lambda} + \partial_\nu \partial_\pi g_{\lambda\mu}\right),$$

the second & fifth terms cancel each other (as partial derivatives commute), leaving

$$R^\rho{}_{\lambda\mu\nu} = \frac{1}{2} g^{\rho\pi} \left(\partial_\mu \partial_\lambda g_{\nu\pi} - \partial_\mu \partial_\pi g_{\lambda\nu} - \partial_\nu \partial_\lambda g_{\mu\pi} + \partial_\nu \partial_\pi g_{\lambda\mu}\right).$$

Now, to get rid of the metric multiplying the bracket, we form

$$\begin{aligned} R_{\alpha\lambda\mu\nu} &= g_{\alpha\rho}R^{\rho}{}_{\lambda\mu\nu} \\ &= \frac{1}{2}g_{\alpha\rho}g^{\rho\pi}\left(\partial_\mu\partial_\lambda g_{\nu\pi} - \partial_\mu\partial_\pi g_{\lambda\nu} - \partial_\nu\partial_\lambda g_{\mu\pi} + \partial_\nu\partial_\pi g_{\lambda\mu}\right) \\ &= \frac{1}{2}\delta^{\pi}_{\alpha}\left(\partial_\mu\partial_\lambda g_{\nu\pi} - \partial_\mu\partial_\pi g_{\lambda\nu} - \partial_\nu\partial_\lambda g_{\mu\pi} + \partial_\nu\partial_\pi g_{\lambda\mu}\right) \\ &= \frac{1}{2}\left(\partial_\mu\partial_\lambda g_{\nu\alpha} - \partial_\mu\partial_\alpha g_{\lambda\nu} - \partial_\nu\partial_\lambda g_{\mu\alpha} + \partial_\nu\partial_\alpha g_{\lambda\mu}\right). \end{aligned}$$

Of course, it must be clear that this is only valid an a LIF. Now, although the above expression is only valid in a LIF, the resulting symmetries are valid everywhere (as the Riemann tensor is a tensor). We see that

$$R_{\alpha\lambda\mu\nu} = -R_{\lambda\alpha\mu\nu} = -R_{\alpha\lambda\nu\mu} = R_{\mu\nu\alpha\lambda} = R_{\lambda\alpha\nu\mu}. \tag{4.3}$$

And further that

$$R_{\alpha\lambda\mu\nu} + R_{\alpha\mu\nu\lambda} + R_{\alpha\nu\lambda\mu} = 0. \tag{4.4}$$

This can also be denoted

$$R_{\alpha(\lambda\mu\nu)} = 0,$$

where the notation is understood to mean cyclic interchange, and sum, over bracketed indices.

Theorem We state (without proof), that if all components of the Riemann tensor are zero, then the space is flat. That is

$$R^{\lambda}{}_{\mu\nu\delta} = 0 \quad \Rightarrow \quad \text{flat space}.$$

4.1.2 The Round Trip

Now, although we shall not go into the details here (we have already presented a full mathematical treatment, however, of the Riemann tensor), one can show that the Riemann tensor comes about from a round-trip around a rectangle.

Consider a rectangle, with horizontal sides of length Δx^μ and vertical sides length δx^μ. Then, if one makes the parallel-transported round trip $A \to B \to C \to D \to A$, and if one computes the coordinate shift (merely due to displacement) at each vertex, then one finds that

$$A^{\rho}_{1} = \left(1 + \delta x^\mu \Delta x^\nu \left[\nabla_\mu, \nabla_\nu\right]\right) A^{\rho}_{0},$$

where A^{ρ}_{1} is the component of A, after visiting that point after making the round trip (i.e. one starts at A^{ρ}_{0}). Then,

$$A^{\rho}_{1} - A^{\rho}_{0} = \Delta A^{\rho} = \delta x^\mu \Delta x^\nu \left[\nabla_\mu, \nabla_\nu\right] A^{\rho}_{0}.$$

Now, we see that $[\nabla_\mu, \nabla_\nu] A_0^\rho = R^\rho{}_{\alpha\mu\nu} A^\alpha$, and so,

$$\Delta A^\rho = \delta x^\mu \Delta x^\nu R^\rho{}_{\alpha\mu\nu} A^\alpha$$

Now, notice that

$$\begin{aligned} R^\rho{}_{\alpha\mu\nu} \delta x^\mu \Delta x^\nu &= \frac{1}{2}\left(R^\rho{}_{\alpha\mu\nu} \delta x^\mu \Delta x^\nu + R^\rho{}_{\alpha\nu\mu} \delta x^\nu \Delta x^\mu\right) \\ &= \frac{1}{2}\left(R^\rho{}_{\alpha\mu\nu} \delta x^\mu \Delta x^\nu - R^\rho{}_{\alpha\mu\nu} \delta x^\nu \Delta x^\mu\right) \\ &= \frac{1}{2} R^\rho{}_{\alpha\mu\nu} \Delta S^{\mu\nu}, \end{aligned}$$

where we have used the anti-symmetry identity of the Riemann tensors last two indices. Also, we have defined $\Delta S^{\mu\nu} \equiv \delta x^\mu \Delta x^\nu - \delta x^\nu \Delta x^\mu$. Therefore, we can write the round-trip expression as

$$\Delta A^\rho = \frac{1}{2} \Delta S^{\mu\nu} R^\rho{}_{\alpha\mu\nu} A^\alpha.$$

Therefore, we have a semi-geometrical interpretation of the Riemann tensor. It is able to tell us the difference in the orientation of a vector, after making a round trip about a rectangle, in a manifold.

4.2 The Ricci Identity

We call the commutator we defined above, the *Ricci identity*. That is, the Ricci identity is

$$[\nabla_\mu, \nabla_\nu] A^\rho = R^\rho{}_{\lambda\mu\nu} A^\lambda,$$

where $R^\rho{}_{\lambda\mu\nu}$ is the Riemann tensor, in a torsion-less manifold.

4.3 The Ricci Tensor & Scalar

If we contract the Riemann tensor on its first & third indices,

$$R^\rho{}_{\lambda\rho\nu} = g^{\rho\alpha} R_{\alpha\lambda\rho\nu},$$

we have a quantity we define

$$R_{\lambda\nu} \equiv R^\rho{}_{\lambda\rho\nu}.$$

If we further contract $R_{\lambda\nu}$,

$$R \equiv g^{\lambda\nu} R_{\lambda\nu}.$$

Thus, we have what we define the *Ricci tensor*, $R_{\mu\nu}$ and *Ricci scalar*, R. By the symmetries of the Riemann tensor above, we can easily see that the Ricci tensor is symmetric.

Now, when we stated that the condition for flat space was that all components of the Riemann tensor were zero; if the Ricci scalar is zero, then the space is not necessarily flat. One can see this, as upon contraction, some non-zero components may cancel each other out in summation.

4.3.1 Example: Plane Polars

Consider the line element
$$ds^2 = d\theta^2 + \sin^2\theta d\phi^2,$$
and suppose that we are given that
$$R^{\theta}{}_{\phi\theta\phi} = \sin^2\theta$$
is the only non-zero component of the Riemann tensor (obviously we can find the other non-zero components by symmetry relations); then, we can compute the Ricci scalar R.

The non-zero components of the metric are easily read off the line element;
$$g_{\theta\theta} = g^{\theta\theta} = 1, \quad g_{\phi\phi} = \sin^2\theta, \quad g^{\phi\phi} = \frac{1}{\sin^2\theta}.$$
Now,
$$R_{\theta\phi\theta\phi} = g_{\theta\theta} R^{\theta}{}_{\phi\theta\phi} = \sin^2\theta.$$
Now, by symmetry of the Riemann tensor,
$$R_{\theta\phi\theta\phi} = -R_{\phi\theta\theta\phi} = -R_{\theta\phi\phi\theta} = R_{\phi\theta\phi\theta}.$$
Now, the Ricci tensor is found by contraction,
$$R_{ij} = g^{nm} R_{nimj}.$$
We are slightly fortunate in that the metric is diagonal. So,
$$\begin{aligned} R_{\theta\theta} &= g^{ij} R_{i\theta j\theta} \\ &= g^{\theta\theta} R_{\theta\theta\theta\theta} + g^{\phi\phi} R_{\phi\theta\phi\theta} \\ &= 1.0 + \frac{1}{\sin^2\theta}.\sin^2\theta \\ &= 1. \end{aligned}$$
Also,
$$\begin{aligned} R_{\theta\phi} &= g^{\theta\theta} R_{\theta\theta\theta\phi} + g^{\phi\phi} R_{\phi\theta\phi\phi} \\ &= 0. \end{aligned}$$
And,
$$\begin{aligned} R_{\phi\phi} &= g^{\theta\theta} R_{\theta\phi\theta\phi} + g^{\phi\phi} R_{\phi\phi\phi\phi} \\ &= 1.\sin^2\theta + 0 \\ &= \sin^2\theta. \end{aligned}$$

Therefore, the Ricci scalar,

$$\begin{aligned} R &= g^{ij}R_{ij} \\ &= g^{\theta\theta}R_{\theta\theta} + g^{\phi\phi}R_{\phi\phi} \\ &= 1 + \frac{1}{\sin^2\theta}\sin^2\theta \\ &= 2. \end{aligned}$$

Therefore, the Ricci scalar is 2 for the plane polar metric.

Now, if we were to repeat this, for the line element

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

we would find that $R = 0$.

4.4 The Bianchi Identity

Consider the Riemann tensor, in a LIF,

$$R^\rho{}_{\lambda\mu\nu} = \partial_\mu \Gamma^\rho{}_{\lambda\nu} - \partial_\nu \Gamma^\rho{}_{\lambda\mu}.$$

Then, let us differentiate it,

$$\nabla_\pi R^\rho{}_{\lambda\mu\nu} = \nabla_\pi \partial_\mu \Gamma^\rho{}_{\lambda\nu} - \nabla_\pi \partial_\nu \Gamma^\rho{}_{\lambda\mu}.$$

Now, even though we dont know how to evaluate these expressions, we can still cycle indices to see what happens. So, making the change

$$\pi \longrightarrow \mu \longrightarrow \nu \longrightarrow \pi,$$

then

$$\nabla_\mu R^\rho{}_{\lambda\nu\pi} = \nabla_\mu \partial_\nu \Gamma^\rho{}_{\lambda\pi} - \nabla_\mu \partial_\pi \Gamma^\rho{}_{\lambda\nu},$$

and again,

$$\nabla_\nu R^\rho{}_{\lambda\pi\mu} = \nabla_\nu \partial_\pi \Gamma^\rho{}_{\lambda\mu} - \nabla_\nu \partial_\mu \Gamma^\rho{}_{\lambda\pi}.$$

Now, if we add these 3 expressions,

$$\begin{aligned} \nabla_\pi R^\rho{}_{\lambda\mu\nu} + \nabla_\mu R^\rho{}_{\lambda\nu\pi} + \nabla_\nu R^\rho{}_{\lambda\pi\mu} = \quad & \nabla_\pi \partial_\mu \Gamma^\rho{}_{\lambda\nu} - \nabla_\pi \partial_\nu \Gamma^\rho{}_{\lambda\mu} \\ & + \nabla_\mu \partial_\nu \Gamma^\rho{}_{\lambda\pi} - \nabla_\mu \partial_\pi \Gamma^\rho{}_{\lambda\nu}, \\ & + \nabla_\nu \partial_\pi \Gamma^\rho{}_{\lambda\mu} - \nabla_\nu \partial_\mu \Gamma^\rho{}_{\lambda\pi}. \end{aligned}$$

Now, in a LIF, the Christofell symbols are zero. Therefore, the covariant derivative is the same as the “usual” partial derivative. So, rather than changing the above symbols, we let covariant and partial derivative swap indices. Then, we can see that the entire RHS cancels itself out, leaving

$$\nabla_\pi R^\rho{}_{\lambda\mu\nu} + \nabla_\mu R^\rho{}_{\lambda\nu\pi} + \nabla_\nu R^\rho{}_{\lambda\pi\mu} = 0.$$

Now, if we drop the ρ-index (using a metric, but index relabeling is trivial),

$$\nabla_\pi R_{\rho\lambda\mu\nu} + \nabla_\mu R_{\rho\lambda\nu\pi} + \nabla_\nu R_{\rho\lambda\pi\mu} = 0.$$

And, using the symmetry property that $R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta}$, then

$$\nabla_\pi R_{\mu\nu\rho\lambda} + \nabla_\mu R_{\nu\pi\rho\lambda} + \nabla_\nu R_{\pi\mu\rho\lambda} = 0,$$

which we see is just a cyclic interchange of the first three indices of the whole expression. That is,

$$\nabla_{(\pi} R_{\mu\nu)\rho\lambda} = 0.$$

Hence, we have arrived at our result. The *Bianchi identity* is that

$$\nabla_\pi R_{\mu\nu\rho\lambda} + \nabla_\nu R_{\pi\mu\rho\lambda} + \nabla_\mu R_{\nu\pi\rho\lambda} = 0. \tag{4.5}$$

The Bianchi identity is infact the equivalent of the rectangular round-trip expression we derived above. The Bianchi identity will come about if one considers the difference in orientation of a vector being parallelly-transported around a cuboid.

Although we derived the Bianchi identity with the Riemann tensor in a LIF, the expression is completely valid in all frames. This is because the Riemann tensor is a tensor; and a tensor equation has the same form in all frames. Thus, one begins to see the power of getting expressions into tensorial form, and of the local inertial frame.

4.5 The Einstein Tensor

Now, let us take the Bianchi identity,

$$\nabla_\pi R_{\mu\nu\rho\lambda} + \nabla_\nu R_{\pi\mu\rho\lambda} + \nabla_\mu R_{\nu\pi\rho\lambda} = 0.$$

Now, let us figure out how to contract this expression, so that we have Ricci tensors, rather than Riemann tensors. Now, if we multiply the expression by $g^{\mu\lambda}$, then we will have achieved our goal (one can see that the indices of this metric are those on the first and last parts of the first Riemann tensor). However, let us do this methodically. So, the first expression will read,

$$g^{\mu\lambda}\nabla_\pi R_{\mu\nu\rho\lambda} = -g^{\mu\lambda}\nabla_\pi R_{\mu\nu\lambda\rho} = -\nabla_\pi R_{\nu\rho},$$

after using the anti-symmetry of the last two indices of the Riemann tensor. The second term can be rewritten, using the symmetry identity of the interchange first two & last two indices of the Riemann tensor;

$$g^{\mu\lambda}\nabla_\nu R_{\pi\mu\rho\lambda} = g^{\mu\lambda}\nabla_\nu R_{\mu\pi\lambda\rho} = \nabla_\nu R_{\pi\rho}.$$

Lastly, the final term of the contracted Bianchi identity is just

$$g^{\mu\lambda}\nabla_\mu R_{\nu\pi\rho\lambda} = \nabla^\lambda R_{\nu\pi\rho\lambda}.$$

Therefore, putting these all together, our contracted Bianchi identity looks like

$$\nabla_\nu R_{\pi\rho} - \nabla_\pi R_{\nu\rho} + \nabla^\lambda R_{\nu\pi\rho\lambda} = 0.$$

Now, multiplying this whole expression by $g^{\nu\rho}$ will contract the last Riemann tensor into a Ricci tensor; as well as the middle Ricci tensor into a Ricci scalar. Thus,

$$\begin{aligned} g^{\nu\rho}\nabla_\nu R_{\pi\rho} - g^{\nu\rho}\nabla_\pi R_{\nu\rho} + g^{\nu\rho}\nabla^\lambda R_{\nu\pi\rho\lambda} &= 0, \\ \Rightarrow \nabla^\rho R_{\pi\rho} - \nabla_\pi R + \nabla^\lambda R_{\pi\lambda} &= 0. \end{aligned}$$

Now, the first and last expressions are identical, as we can interchange the indices at will. Therefore, we have

$$2\nabla^\rho R_{\pi\rho} - \nabla_\pi R = 0.$$

Then, notice that we can write

$$2\nabla^\rho R_{\pi\rho} - g_{\pi\rho}\nabla^\rho R = 0,$$

and therefore that

$$\nabla^\rho \left(R_{\pi\rho} - \tfrac{1}{2} g_{\pi\rho} R\right) = 0.$$

Therefore, we can define the *Einstein tensor*,

$$G_{\mu\nu} \equiv R_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} R, \tag{4.6}$$

whereby

$$\nabla^\mu G_{\mu\nu} = 0; \tag{4.7}$$

after noting that the metric, Ricci & therefore the Einstein tensor are symmetric. This is called the *contracted Bianchi identity.*

4.6 Geodesic Deviation

Suppose we take, on flat space, two affinely parameterised geodesics $x^\mu(\tau), y^\mu(\tau)$, that are on a collision course. That is, the distance between the two lines,

$$\delta^\mu(\tau) \equiv x^\mu(\tau) - y^\mu(\tau),$$

decreases. On flat space, the distance will decrease linearly. That is,

$$\frac{d\delta^\mu}{d\tau} = \text{const} \quad \Rightarrow \quad \frac{d^2\delta^\mu}{d\tau^2} = 0.$$

Now consider a curved space. Let the paths be tangents. Then, the distance between the two wont decrease linearly. Instead, they will accelerate; thus

$$\frac{D^2\delta^\mu}{D\tau^2} = R^\mu{}_{\alpha\beta\rho}T^\alpha T^\beta \delta^\rho. \tag{4.8}$$

To imagine this in a physical situation, consider two balls falling towards the centre of the earth. Now, the balls will obviously move towards each other, as their motion is radial. However, there will be deviation from radial, and that deviation will be due to the curvature of space. That is, one will observe the balls accelerate towards each other (rather than the expected linear motion towards each other).

Derivation We can derive the geodesic deviation equation, by considering the 2-dim manifold swept out by two affinely parameterised geodesics, next to each other. The manifold may be parameterised by $x^\mu = x^\mu(\tau, \sigma)$ (i.e. two coordinates on this surface, rather than the usual one, on a curve). The tangent vectors are

$$T^\mu = \frac{dx^\mu}{d\tau}, \quad \delta^\mu = \frac{dx^\mu}{d\sigma}.$$

Now, let us show a useful relation. Consider

$$\begin{aligned} T^\mu \nabla_\mu \delta^\nu &= T^\mu \left(\partial_\mu \delta^\nu + \Gamma^\nu{}_{\mu\lambda}\delta^\lambda\right) \\ &= \frac{dx^\mu}{d\tau}\frac{\partial}{\partial x^\mu}\frac{dx^\nu}{d\sigma} + \Gamma^\nu{}_{\mu\lambda}T^\mu\delta^\lambda. \end{aligned}$$

Now, the first term can be rewritten

$$\begin{aligned} \frac{dx^\mu}{d\tau}\frac{\partial}{\partial x^\mu}\frac{dx^\nu}{d\sigma} &= \frac{d^2x^\nu}{d\tau d\sigma} \\ &= \frac{d^2x^\nu}{d\sigma d\tau} \\ &= \frac{dx^\mu}{d\sigma}\frac{\partial}{\partial x^\mu}\frac{dx^\nu}{d\tau}. \end{aligned}$$

Hence, we use this to see that

$$\begin{aligned} T^\mu \nabla_\mu \delta^\nu &= \frac{dx^\mu}{d\sigma}\frac{\partial}{\partial x^\mu}\frac{dx^\nu}{d\tau} + \Gamma^\nu{}_{\mu\lambda}T^\mu\delta^\lambda \\ &= \delta^\mu \partial_\mu T^\nu + \Gamma^\nu{}_{\lambda\mu}T^\lambda\delta^\mu \\ &= \delta^\mu \nabla_\mu T^\nu, \end{aligned}$$

where we have merely used the symmetry of the Christofell symbol. Hence, we have the relation

$$T^\mu \nabla_\mu \delta^\nu = \delta^\mu \nabla_\mu T^\nu. \tag{4.9}$$

Now, let us state the operator

$$\frac{D^2}{D\tau^2} = T^\alpha \nabla_\alpha T^\beta \nabla_\beta,$$

and compute its action upon δ^μ,

$$\frac{D^2 \delta^\mu}{D\tau^2} = T^\alpha \nabla_\alpha \left(T^\beta \nabla_\beta \delta^\mu \right).$$

Now, we use our relation (4.9), to see that

$$\frac{D^2 \delta^\mu}{D\tau^2} = T^\alpha \nabla_\alpha \left(\delta^\beta \nabla_\beta T^\mu \right).$$

Let us now expand this out,

$$\frac{D^2 \delta^\mu}{D\tau^2} = T^\alpha \nabla_\alpha \delta^\beta \nabla_\beta T^\mu + T^\alpha \delta^\beta \nabla_\alpha \nabla_\beta T^\mu.$$

Now, we can rewrite the two-covariant derivatives term on the far RHS, using the Ricci identity,

$$\left[\nabla_\mu, \nabla_\nu \right] A^\rho = R^\rho{}_{\lambda\mu\nu} A^\lambda,$$

so that we have

$$\begin{aligned}
\frac{D^2 \delta^\mu}{D\tau^2} &= T^\alpha \nabla_\alpha \delta^\beta \nabla_\beta T^\mu + T^\alpha \delta^\beta R^\mu{}_{\lambda\alpha\beta} T^\lambda + T^\alpha \delta^\beta \nabla_\beta \nabla_\alpha T^\mu \\
&= \delta^\alpha \nabla_\alpha T^\beta \nabla_\beta T^\mu + T^\beta \delta^\alpha \nabla_\alpha \nabla_\beta T^\mu + R^\mu{}_{\lambda\alpha\beta} T^\alpha T^\lambda \delta^\beta \\
&= \delta^\alpha \left(\nabla_\alpha T^\beta \nabla_\beta T^\mu + T^\beta \nabla_\alpha \nabla_\beta T^\mu \right) + R^\mu{}_{\lambda\alpha\beta} T^\alpha T^\lambda \delta^\beta.
\end{aligned}$$

In the first step we used our relation (4.9) on the first term, and changed dummy indices on the third term, then we merely factorised the expression. Now, notice that the bracketed term can be written

$$\nabla_\alpha T^\beta \nabla_\beta T^\mu + T^\beta \nabla_\alpha \nabla_\beta T^\mu = \nabla_\alpha \left(T^\beta \nabla_\beta T^\mu \right),$$

but the bracketed part on the RHS is zero on an affinely parameterised goedesic. Hence,

$$\frac{D^2 \delta^\mu}{D\tau^2} = R^\mu{}_{\lambda\alpha\beta} T^\alpha T^\lambda \delta^\beta,$$

or, trivially relabelling indices, we arrive at our equation for geodesic deviation

$$\frac{D^2 \delta^\mu}{D\tau^2} = R^\mu{}_{\alpha\beta\rho} T^\alpha T^\beta \delta^\rho. \tag{4.10}$$

5 Einstein's Equation

We almost have enough tools to be able to write Einstein's equation.

We have seen that freely-falling particles follow geodesics. In curved spacetime, the geodesics will probably be curves. So then, what makes the spacetime curved?

If we consider electromagnetic theory, there is a source for the electric field: the electron. For a field, there is a source. Therefore, we need a source term that will curve spacetime. We shall now discuss a term that is the "gravitation source term".

5.1 The Energy Momentum Tensor $T^{\mu\nu}$

We shall start by stating that there exists a tensor $T^{\mu\nu}$, which is symmetric. That is, $T^{\mu\nu} = T^{\nu\mu}$. Furthermore, we shall state that the components of this tensor contain all possible forms of energy and momentum (it will be this tensor which is the source-term). Let us state how to compute a given component of the tensor.

A given element $T^{\mu\nu}$ is the flux of p^μ that goes through the hypersurface $x^\nu = \text{const}$.

The structure of the tensor is clearly

$$(T^{\mu\nu}) = \begin{pmatrix} T^{tt} & T^{ti} \\ T^{it} & T^{ij} \end{pmatrix}.$$

Also, before we start to compute the components of the tensor, we must state that the full 4-volume is just $\Delta t \Delta x \Delta y \Delta z$.

5.1.1 Components of $T^{\mu\nu}$

Lets consider $T^{00} = T^{tt}$. Then, by our definition, that element is the flux of p^0 through the surface $x^0 = \text{const}$. Now, $p^0 = E$ and $x^0 = t$. Therefore, we see that T^{tt} is the flux of energy E through a 3-volume $\Delta x \Delta y \Delta z$ (it is the hypersurface that holds $x^0 = t$ constant). Therefore,

$$T^{tt} = \frac{E}{\Delta x \Delta y \Delta z} \equiv \varepsilon,$$

that is, the energy per unit volume, the energy density ε.

Consider the component $T^{01} = T^{tx}$. Then, we see that it is the flux of $p^0 = E$ through the hypersurface $x^1 = x = \text{const}$. That is,

$$T^{01} = \frac{E}{\Delta t \Delta y \Delta z},$$

which has the interpretation of being the energy flux through the $y - z$ plane, in unit time. This is easily extrapolated to the any term T^{0i}: the energy flux through a surface, in unit time.

Now consider the purely-spatial components, T^{ij}. For example,

$$T^{ix} = \frac{\Delta p^i}{\Delta t \Delta y \Delta z}.$$

Now, notice that we can rewrite this,

$$T^{ix} = \frac{\Delta p^i/\Delta t}{A_{yz}}, \quad A_{yz} \equiv \Delta y \Delta z;$$

where we have fairly obviously defined an area-element. A change in momentum per unit time is just a force. Thus,

$$T^{ix} = \frac{F^i}{A_{yz}},$$

which is a force per unit area: a pressure. Consider the specific component,

$$T^{yx} = \frac{\Delta p^y/\Delta t}{\Delta y \Delta z},$$

then, using the relation $p^i = v^i E$, we see that

$$T^{yx} = \frac{\Delta v^y E/\Delta t}{\Delta y \Delta z}.$$

So, as $v^i = x^i/t$, we see that this is just

$$T^{yx} = \frac{\Delta y/\Delta t E/\Delta t}{\Delta y \Delta z}.$$

Now then, as the Δy's cancel, we can just replace them with Δx's, thus

$$\begin{aligned} T^{yx} &= \frac{\Delta x/\Delta t E/\Delta t}{\Delta x \Delta z} \\ &= \frac{\Delta v^x E/\Delta t}{\Delta x \Delta z} \\ &= \frac{\Delta p^x/\Delta t}{\Delta x \Delta z} \\ &= T^{xy}. \end{aligned}$$

Therefore, with this little exercise, we see that the spatial components of the energy-momentum tensor are infact symmetric. One may be able to see that the off-diagonal components of the spatial part, those T^{ii}, correspond to the force perpendicular to a surface. Those off-digonal elements are the force parallel to a surface (shear). Therefore, the spatial components, T^{ij} are components of the stress-tensor.

The final part to the tensor, are those components T^{it}. Thus, we see that they are the flow of p^i through the hypersurface $t = \text{const}$. That is, the momentum flow in a given 3-volume, at a constant time. That is, how much momentum there exists in a unit volume, at a single time. This is clearly the momentum density.

$$T^{it} = \frac{\Delta p^i}{\Delta x \Delta y \Delta z} \equiv \pi^i.$$

To see that $T^{it} = T^{ti}$, consider the above expression; writing $p^i = v^i E = x^i E/t$, then,

$$T^{it} = \frac{\Delta x^i/\Delta t E}{\Delta x \Delta y \Delta z} = \frac{\Delta x^i E}{\Delta t \Delta x \Delta y \Delta z}.$$

Then, as i is changed through $i = x, y, z$, different components on the denominator will be cancelled out, leaving only those in the corresponding T^{ti}.

Therefore, we have seen that the energy-momentum tensor $T^{\mu\nu}$ is symmetric, by considering its components. The colloquial construction of the tensor is thus

$$(T^{\mu\nu}) = \begin{pmatrix} \text{energy density} & \text{energy flux} \\ \text{momentum density} & \text{stress tensor} \end{pmatrix}.$$

We shall write that $T^{it} = \pi^i$, $T^{tt} = \varepsilon$.

5.1.2 Conservation Equations

The standard conservation equation is that

$$\nabla_\nu T^{\mu\nu} = 0. \tag{5.1}$$

Let us consider this in a LIF. Then, it simply reads $\partial_\nu T^{\mu\nu} = 0$.

Now, take the time-component, $\mu = 0 = t$. Then, the conservation equation reads

$$\frac{\partial}{\partial t} T^{tt} + \frac{\partial}{\partial x^i} T^{ti} = 0,$$

which is just

$$\frac{\partial \varepsilon}{\partial t} + \frac{\partial \pi^i}{\partial x^i} = 0.$$

This equation can be written

$$\frac{\partial \varepsilon}{\partial t} + \nabla \cdot \boldsymbol{\pi} = 0,$$

and is the familiar continuity equation, for energy. Note that this is only valid in a LIF.

Let us take the spatial components, $\mu = i$ of the conservation equation. Thus,

$$\frac{\partial}{\partial t}T^{it} + \frac{\partial}{\partial x^j}T^{ij} = 0.$$

Now, if we write the force density, in a given direction

$$\phi^i \equiv -\frac{\partial T^{ij}}{\partial x^j},$$

then we see that the conservation equation reads

$$\frac{\partial \pi^i}{\partial t} - \phi^i = 0,$$

which is just the statement that the rate of change of momentum density is the force density. This is the familiar statement of Newton's second law. That is, the above equation is just

$$\frac{\partial \boldsymbol{\pi}}{\partial t} = \boldsymbol{\phi}.$$

Therefore, we see that the energy-momentum tensor $T^{\mu\nu}$ contains all sources of energy and momentum, and satisfies basic conservation relations.

5.1.3 Perfect Fluids

A perfect fluid is defined to be one for whom there is no viscosity or heat conduction. This "restriction" makes the energy-momentum tensor look a lot simpler.

That a fluid has no heat conduction means that there is no transfer of energy, across surfaces. Viscous forces are those which are parallel to a surface (shear). Thus, the absence of such forces, implies that all forces on surfaces are perpendicular to those surfaces.

Therefore, if we consider our previous "derivation" of the components of $T^{\mu\nu}$, we see that it must be diagonal. This is because

- No heat conduction implies no energy flux. Therefore, $T^{ti} = T^{it} = 0$.
- No viscosity means that all parallel forces are zero. This only leaves diagonal components to the stress tensor. All components left-over are just pressures (as discussed previously), P.

We shall change notation slightly, so that ρ is the energy density (which is clearly the case, via $\varepsilon = \rho c^2$, with $c = 1$). Therefore, we see that for a perfect fluid, at rest,

$$T^{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P \end{pmatrix} = \mathrm{diag}(\rho, P, P, P).$$

The Perfect Fluid Tensor The general expression for the energy-momentum tensor, for a perfect fluid in its LIF is

$$T^{\mu\nu} = (\rho + P)u^\mu u^\nu - P g^{\mu\nu}, \tag{5.2}$$

where $u^\mu = \gamma(1, \mathbf{u})$ is the 4-velocity, ρ the energy density and P the pressure of the fluid. If we take this tensor, with the fluid at rest in its LIF, in flat space, then, some of the components are

$$\begin{aligned} T^{00} &= (\rho + P) - P = \rho, \\ T^{12} &= 0, \\ T^{ij} &= P. \end{aligned}$$

Infact, all off-diagonal components are zero. Then, we see that we have recovered our previous expression for a perfect fluid at rest.

We can easily recover some standard fluid mechanics results from the perfect fluid tensor. Suppose we have a non-relativistic pressure-less fluid, $P = 0$, then, the energy-conservation equation is just

$$\partial_\mu(\rho u^\mu u^0) = 0,$$

which easily becomes

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho u^i}{\partial x^i} = 0,$$

which is just

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0.$$

5.2 Einstein's Equation

We now have a source term. The sources of energy and momentum can be written "into" the energy-momentum tensor, $T^{\mu\nu}$; a tensor which satisfies the conservation equation.

Now, from the previous sections contracted Bianchi identity,

$$\nabla^\mu G_{\mu\nu} = 0, \quad G_{\mu\nu} \equiv R_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} R,$$

we have an expression which takes care of the geometry of the spacetime. Recall that the Ricci tensor/scalar are composed of differentials (of various orders) of the metric, where the metric gives meaning to distances within a manifold. Then, if we can equate this expression to an expression which gives information as to what is doing the curving, then we have our general theory of relativity. We must use an expression which also has zero covariant derivative.

The obvious choice is the energy-momentum tensor. Therefore, we write

$$G_{\mu\nu} = \kappa T_{\mu\nu}.$$

Therefore, up to a constant κ, we have a LHS which describes the geometry of a manifold, and a RHS which describes the distribution of all forms of energy and momentum in the manifold. Therefore, we say that the distribution of energy-momentum in a manifold causes the manifold to become curved.

Therefore, *Einstein's equation* is

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R = \kappa T_{\mu\nu}. \tag{5.3}$$

Notice that both sides have vanishing covariant derivative. We shall be able to find the constant κ when we consider the Newtonian limit of the theory.

We can write this in an alternative form. Consider multiplying the whole expression by $g^{\mu\nu}$,

$$g^{\mu\nu}(R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R) = \kappa g^{\mu\nu}T_{\mu\nu},$$

then, writing the trace of the energy-momentum tensor $g^{\mu\nu}T_{\mu\nu} \equiv T$, and noting that we see that the Ricci tensor becomes the Ricci scalar upon contraction; thus, we see that

$$R - \tfrac{1}{2}g^{\mu}{}_{\mu}R = \kappa T.$$

Now, the metric multiplied by its inverse is just the Kronecker-delta. Thus, $g^{\mu}{}_{\mu} = \delta^{\mu}_{\mu} = 4$. Therefore,

$$R - \tfrac{1}{2}4R = \kappa T,$$

hence,

$$R = -\kappa T.$$

Thus, we have written the Ricci scalar in terms of the trace of the energy-momentum tensor. Hence, we can write the Einstein equation as

$$R_{\mu\nu} + \tfrac{1}{2}g_{\mu\nu}\kappa T = \kappa T_{\mu\nu},$$

which is just

$$R_{\mu\nu} = \kappa(T_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}T). \tag{5.4}$$

This an entirely equivalent form of Einstein's equation.

5.2.1 The Cosmological Constant

Now, if we require the covariant derivative of an expression to be zero, we may add on an "extra term", a constant, which will not change the value of the covariant derivative. The covariant derivative of the metric is zero, so we may add on any number of metrics and retain zero covariant derivative. Therefore,

$$G_{\mu\nu} = \kappa T_{\mu\nu} + \Lambda g_{\mu\nu}$$

is still consistent with zero covariant derivative. So, why is this a problem?

Consider the expression, from electrodynamic theory, in a LIF,

$$\partial^\mu F_{\mu\nu} = J_\nu.$$

Then, consider taking the differential of the expression,

$$\partial^\nu \partial^\mu F_{\mu\nu} = \partial^\nu J_\nu = 0;$$

where the equality with zero comes from the "usual" conservation equation. Now, consider that we try to add on an extra term,

$$\partial^\mu (F_{\mu\nu} + \Lambda \eta_{\mu\nu}).$$

Then, these two expressions are not the same. That is, we do not have the freedom to modify the field tensor by adding on an arbitrary quantity of metrics. The reason we are not able to do this, is because the field tensor is anti-symmetric, and the metric is symmetric.

Therefore, the reason we are able to add the constant metric term into Einstein's equation, is because both the Einstein tensor $G_{\mu\nu}$ and energy-momentum tensor $T_{\mu\nu}$ are symmetric (as is the metric); as well as the metric having zero covariant derivative.

The cosmological constant Λ has been measured to exist within the universe, having a very small numerical value. We shall usually define the cosmological constant within the energy-momentum tensor, so that we will essentially ignore it. However, it is to be understood that the term is within the energy-momentum tensor.

5.3 The Newtonian Limit

Let us discuss the correspondences of the theory of gravity on curved spacetime, with Newtonian gravity.

The equation of motion of a free particle, in Newton's theory, is just given by Newton's second law of motion,

$$\frac{d^2 x^i}{dt^2} = -\delta^{ij} \frac{\partial \Phi}{\partial x^j},$$

where Φ is the gravitational potential a particle feels. The corresponding equation of motion for a free particle, in curved spacetime, is the geodesic

$$\frac{d^2 x^\mu}{d\tau^2} = -\Gamma^\mu{}_{\alpha\beta} \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau},$$

where the Christofell symbol $\Gamma^\mu{}_{\alpha\beta}$ contains information about the geometry of the spacetime.

The field equation, which describes how "stuff" generates the gravitational field, for the Newtonian theory is

$$\nabla^2 \Phi = 4\pi G \rho_m.$$

That is, Poisson's equation. This equation tells us that for some mass density ρ_m, there is an associated gravitational potential Φ. Combined with the equation of motion, we see that a mass density gives rise to a gravitational potential, which affects how a free particle moves.

The field equation in general relativity, is Einstein's equation,

$$R_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} R = \kappa T_{\mu\nu}.$$

Some distribution of energy and momentum, defined within the energy-momentum tensor, gives rise to a different geometry. This geometrical information is then carried around by the Ricci tensor, within the metric. The metric then gives the Christofell symbols, which change the equation of motion - the geodesic.

The basic correspondence is

$$g_{\mu\nu} \longleftrightarrow \Phi \qquad T_{\mu\nu} \longleftrightarrow \rho_m.$$

Notice that we have only been referring to "free-particles". A free-particle is one which does not have any external influences on its motion. For example, this could mean a stone being dropped, in vacuum, from a building. The stones motion is only affected by the gravitational potential from the earth. Notice then, that the motion of a freely-falling particle in a curved spacetime is entirely due to the spacetime through which is moves. That is, its trajectory will be curved because of the geometry of the spacetime.

To modify these equations for a particle which is acted upon by an external force, $\mathbf{F}_{ext}$, one must merely add this to each component of the equation of motion.

5.3.1 Newtonian Gravity from Einstein's Gravity

Let us consider the geodesic equation,for a free particle,

$$\frac{d^2x^\mu}{d\tau} + \Gamma^\mu{}_{\alpha\beta} \frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} = 0.$$

Now, let us consider the non-relativistic limit of this geodesic.

Firstly, for non-relativistic motion, $\tau = t$. Second, $dx^i/dt \ll 1$. Then, we can write the geodesic equation as

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu{}_{00} \left(\frac{dt}{d\tau}\right)^2 + \mathcal{O}\left(\frac{dx^i}{d\tau}\right)^2 = 0,$$

which is just

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu{}_{00} = 0.$$

Now then, to continue, we make an assumption about the metric. We say that the metric is Minkowskian, with a small perturbation,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad h_{\mu\nu} \ll 1.$$

We shall only work to first order in the perturbation. That is, we shall neglect any terms $\mathcal{O}(h^2)$. We further say that the perturbation is static. That is, $h_{\mu\nu}(x^i)$ only; which immediately tells us that

$$\partial_0 h_{\mu\nu} = \partial_t h_{\mu\nu} = 0.$$

Now, the general expression for the Christofell symbol is

$$\Gamma^\rho{}_{\alpha\beta} = \frac{1}{2} g^{\rho\nu}\left(\partial_\alpha g_{\beta\nu} + \partial_\beta g_{\nu\alpha} - \partial_\nu g_{\alpha\beta}\right).$$

Then, the components that we are interested in are just

$$\Gamma^\mu{}_{00} = \frac{1}{2}\sum_\nu g^{\mu\nu}(\partial_0 g_{0\nu} + \partial_0 g_{\nu 0} - \partial_\nu g_{00}).$$

Now, as the time-differential of the metric is zero, all but the last term is zero. We shall also drop the implied summation;

$$\Gamma^\mu{}_{00} = -\frac{1}{2} g^{\mu\nu}\partial_\nu g_{00}.$$

We shall now drop the greek index on the RHS, and only use roman. This is because the time-differential of the metric is zero. Thus,

$$\Gamma^\mu{}_{00} = -\frac{1}{2} g^{\mu i}\partial_i g_{00}.$$

Inserting our expression for the metric,

$$\begin{aligned}\Gamma^\mu{}_{00} &= -\frac{1}{2}(\eta^{\mu i} - h^{\mu i})\partial_i(\eta_{00} + h_{00}) \\ &= -\frac{1}{2}\left(\eta^{\mu i}\partial_i h_{00} - h^{\mu i}\partial_i h_{00}\right).\end{aligned}$$

Now, the expression on the far right is $\mathcal{O}(h^2)$ thus, we ignore it. Therefore,

$$\Gamma^\mu{}_{00} = -\frac{1}{2}\eta^{\mu i}\partial_i h_{00}.$$

Finally, recall that the Minkowski metric is diagonal. Therefore, we only have contribution for $\mu = i$. Therefore, as $\eta^{ii} = -1$, to first order static-perturbation

$$\Gamma^i{}_{00} = \frac{1}{2}\partial_i h_{00}, \quad \Gamma^0{}_{00} = 0.$$

Therefore, the geodesic equation is

$$\frac{d^2 t}{d\tau^2} = 0, \quad \frac{d^2 x^i}{d\tau^2} + \frac{1}{2}\partial_i h_{00} = 0.$$

Now, we use the first expression to tell us that $dt = A d\tau$. We then set $A = 1$, to see that $t = \tau$. Therefore, the second expression is just

$$\frac{d^2 x^i}{dt^2} + \frac{1}{2}\partial_i h_{00} = 0.$$

Writing this as a vector equation, this is just

$$\frac{d^2\mathbf{x}}{dt^2} + \frac{1}{2}\nabla h_{00} = 0,$$

trivially rewriting results in

$$\frac{d^2\mathbf{x}}{dt^2} = -\frac{1}{2}\nabla h_{00}.$$

Now then, recall the Newtonian equation,

$$\frac{d^2\mathbf{x}}{dt^2} = -\nabla\Phi(x).$$

Then, we can read off the correspondence,

$$h_{00} = 2\Phi.$$

Finally, as the metric is just $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, then

$$g_{00} = 1 + 2\Phi.$$

Therefore, we see that the time-component of a static perturbation to the Minkowski metric is the gravitational potential.

Recall the Riemann tensor, in a LIF,

$$R^\rho{}_{\lambda\mu\nu} = \partial_\mu \Gamma^\rho{}_{\lambda\nu} - \partial_\nu \Gamma^\rho{}_{\lambda\mu},$$

and thus the Ricci tensor,

$$R_{\lambda\nu} = R^\rho{}_{\lambda\rho\nu} = \partial_\rho \Gamma^\rho{}_{\lambda\nu} - \partial_\nu \Gamma^\rho{}_{\lambda\rho}.$$

Now, let us compute the component R_{00}. Then,

$$R_{00} = \partial_\rho \Gamma^\rho{}_{00} - \partial_0 \Gamma^\rho{}_{0\rho},$$

noting that

$$\Gamma^i{}_{00} = \frac{1}{2}\partial_i h_{00}, \quad \Gamma^0{}_{00} = 0,$$

we then see that

$$\begin{aligned} R_{00} &= \partial_i \Gamma^i{}_{00} \\ &= \frac{1}{2}\partial_i \partial_i h_{00} \\ &= \frac{1}{2}\nabla^2 h_{00}. \end{aligned}$$

Further recall that we just derived that $h_{00} = 2\Phi$, then

$$R_{00} = \nabla^2\Phi.$$

Now then, we are now in a position to compute the constant κ in Einstein's field equation. Let us use the alternative form of the field equation, and take the "00" components;

$$R_{00} = \kappa(T_{00} - \tfrac{1}{2}g_{00}T).$$

Now, the trace T is just

$$T \equiv g^{\mu\nu}T_{\mu\nu} = T^{\mu}_{\ \mu}.$$

Therefore,

$$\begin{aligned} g_{00}T &= (\eta_{00} + h_{00})(\eta^{00} - h^{00})T_{00} \\ &= (\eta_{00}\eta^{00} - \eta_{00}h^{00} + \eta^{00}h_{00} - h_{00}h^{00})T_{00} \\ &= T_{00} + \mathcal{O}(h^2). \end{aligned}$$

Let us suppose that the field is generated by a static, non-relativistic body, mass density ρ_m. Then, $T_{00} = \rho_m$. Therefore, the field equation becomes

$$R_{00} = \kappa(\rho_m - \tfrac{1}{2}\rho_m) = \kappa\tfrac{1}{2}\rho_m.$$

Now, we have the Poisson equation, $\nabla^2\Phi = 4\pi G\rho_m$, and also that $R_{00} = \nabla^2\Phi$. Therefore, equating the two,

$$\nabla^2\Phi = 4\pi G\rho_m = \tfrac{1}{2}\kappa\rho_m,$$

we see that

$$\kappa = 8\pi G.$$

Therefore, the full field equation is

$$R_{\mu\nu} = 8\pi G(T_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}T). \tag{5.5}$$

5.4 Linearised Gravity

Let us take our perturbed metric,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu},$$

where $h_{\mu\nu} << 1$. Now then, notice that

$$\begin{aligned} g^{\mu\nu}g_{\nu\lambda} &= (\eta^{\mu\nu} - h^{\mu\nu})(\eta_{\nu\lambda} + h_{\nu\lambda}) \\ &= \eta^{\mu\nu}\eta_{\nu\lambda} + \eta^{\mu\nu}h_{\nu\lambda} - h^{\mu\nu}\eta_{\nu\lambda} - h^{\mu\nu}h_{\nu\lambda} \\ &= \delta^{\mu}_{\lambda} + \mathcal{O}(h^2). \end{aligned}$$

Now, consider a coordinate transformation,

$$x^{\mu} \mapsto x'^{\mu} = x^{\mu} + \epsilon^{\mu}, \quad x^{\mu} = x'^{\mu} - \epsilon^{\mu}.$$

Then, the Jacobians are clearly

$$J^{\mu}{}_{\nu} = \delta^{\mu}_{\nu} + \partial_{\nu}\epsilon^{\mu}, \quad \left(J^{-1}\right)^{\mu}{}_{\nu} = \delta^{\mu}_{\nu} - \partial_{\nu}\epsilon^{\mu}.$$

The $\epsilon^{\mu} << 1$. So, we work to first order in ϵ^{μ} only. Now then, lets consider the transformation of the metric,

$$g'_{\mu\nu} = \left(J^{-1}\right)^{\alpha}{}_{\mu} \left(J^{-1}\right)^{\beta}{}_{\nu} g_{\alpha\beta}.$$

Then, using our Jacobians for the coordinate transformation, this becomes

$$\begin{aligned} g'_{\mu\nu} &= (\delta^{\alpha}_{\mu} - \partial_{\mu}\epsilon^{\alpha})(\delta^{\beta}_{\nu} - \partial_{\nu}\epsilon^{\beta})g_{\alpha\beta} \\ &= (\delta^{\alpha}_{\mu}\delta^{\beta}_{\nu} - \delta^{\alpha}_{\mu}\partial_{\nu}\epsilon^{\beta} - \partial_{\mu}\epsilon^{\alpha}\delta^{\beta}_{\nu} + \partial_{\mu}\epsilon^{\alpha}\partial_{\nu}\epsilon^{\beta})g_{\alpha\beta} \\ &= g_{\mu\nu} - \partial_{\nu}\epsilon^{\beta}g_{\mu\beta} - \partial_{\mu}\epsilon^{\alpha}g_{\alpha\nu} + \mathcal{O}(\epsilon^2). \end{aligned}$$

Now then, notice that by the product rule,

$$\partial_{\nu}\epsilon_{\mu} = \partial_{\nu}(g_{\mu\beta}\epsilon^{\beta}) = \epsilon^{\beta}\partial_{\nu}g_{\mu\beta} + g_{\mu\beta}\partial_{\nu}\epsilon^{\beta},$$

and therefore that

$$g_{\mu\beta}\partial_{\nu}\epsilon^{\beta} = \partial_{\nu}\epsilon_{\mu} - \epsilon^{\beta}\partial_{\nu}g_{\mu\beta}.$$

Hence, using this, we see that the transformation of the metric looks like

$$g'_{\mu\nu} = g_{\mu\nu} - \partial_{\nu}\epsilon_{\mu} - \partial_{\mu}\epsilon_{\nu} + \epsilon^{\beta}\partial_{\nu}g_{\mu\beta} + \epsilon^{\alpha}\partial_{\mu}g_{\alpha\nu}.$$

Now, the last two terms on the right are both $\mathcal{O}(\epsilon^2)$. This is because the metric is of $\mathcal{O}(\epsilon)$, and therefore ϵ times the derivative of the metric is $\mathcal{O}(\epsilon^2)$. Therefore,

$$g'_{\mu\nu} = g_{\mu\nu} - \partial_{\nu}\epsilon_{\mu} - \partial_{\mu}\epsilon_{\nu} + \mathcal{O}(\epsilon^2).$$

Now, using the fact that $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, and $\eta_{\mu\nu} = \eta'_{\mu\nu}$, then the above is just

$$\eta_{\mu\nu} + h'_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} - \partial_{\nu}\epsilon_{\mu} - \partial_{\mu}\epsilon_{\nu},$$

which simply becomes

$$h'_{\mu\nu} = h_{\mu\nu} - \partial_{\nu}\epsilon_{\mu} - \partial_{\mu}\epsilon_{\nu}. \tag{5.6}$$

5.4.1 Linearising Einstein's Equation

Now, recall that Einstein's equation was composed of the Ricci tensor and the energy-momentum tensor. Now, the Ricci tensor was composed of derivatives of the Christofell symbol, which in turn contained derivatives of the metric. Now, we can recompute the Einstein equation under the coordinate transformation defined above as

$$x^{\mu} \mapsto x'^{\mu} = x^{\mu} + \epsilon^{\mu} \quad \Rightarrow \quad h'_{\mu\nu} = h_{\mu\nu} - \partial_{\nu}\epsilon_{\mu} - \partial_{\mu}\epsilon_{\nu}.$$

So, consider

$$\partial_\nu g_{\alpha\beta} = \partial_\nu(\eta_{\alpha\beta} + h_{\alpha\beta}) = \partial_\nu h_{\alpha\beta}.$$

Therefore, the Christofell symbol, defined as

$$\Gamma^\rho{}_{\alpha\beta} = \frac{1}{2} g^{\rho\nu}(\partial_\alpha g_{\beta\nu} + \partial_\beta g_{\nu\alpha} - \partial_\nu g_{\alpha\beta}),$$

becomes

$$\Gamma^\rho{}_{\alpha\beta} = \frac{1}{2} \eta^{\rho\nu}(\partial_\alpha h_{\beta\nu} + \partial_\beta h_{\nu\alpha} - \partial_\nu h_{\alpha\beta}).$$

Now, the Ricci tensor has components such as the product of two Christofell symbols. It is clear that these will be $\mathcal{O}(\epsilon^2)$, and therefore negligible. Hence, the Ricci tensor would look like

$$R_{\mu\nu} = \partial_\rho \Gamma^\rho{}_{\mu\nu} - \partial_\nu \Gamma^\rho{}_{\mu\rho}.$$

Then, plugging in our Christofell symbols,

$$R_{\mu\nu} = \frac{1}{2} \eta^{\rho\sigma}(\partial_\rho\partial_\mu h_{\nu\sigma} + \partial_\rho\partial_\nu h_{\sigma\mu} - \partial_\rho\partial_\sigma h_{\mu\nu} - \partial_\nu\partial_\mu h_{\rho\sigma} - \partial_\nu\partial_\rho h_{\sigma\mu} + \partial_\nu\partial_\sigma h_{\mu\rho}).$$

This becomes, after noting that partial derivatives commute, the Minkowski metric commutes with partial derivatives and that the second and fifth terms cancel,

$$2R_{\mu\nu} = \partial^\sigma\partial_\mu h_{\nu\sigma} - \partial^\rho\partial_\rho h_{\mu\nu} - \partial_\nu\partial_\mu h^\rho{}_\rho + \partial^\rho\partial_\nu h_{\mu\rho}.$$

Now, $h^\rho{}_\rho \equiv h$, and changing the σ index on the first expression to a ρ,

$$R_{\mu\nu} = \frac{1}{2}\left(\partial^\rho\partial_\mu h_{\nu\rho} + \partial^\rho\partial_\nu h_{\mu\rho} - \partial^\rho\partial_\rho h_{\mu\nu} - \partial_\nu\partial_\mu h\right).$$

Then, the Ricci scalar is

$$\begin{aligned} R &= g^{\mu\nu} R_{\mu\nu} \\ &= \eta^{\mu\nu} R_{\mu\nu} \\ &= \frac{1}{2}\left(\partial^\rho\partial^\nu h_{\nu\rho} + \partial^\rho\partial^\mu h_{\mu\rho} - \partial^\rho\partial_\rho h - \partial_\nu\partial^\nu h\right) \\ &= \partial^\rho\partial^\nu h_{\nu\rho} - \partial_\nu\partial^\nu h. \end{aligned}$$

Now then, the Einstein tensor is defined as

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

Therefore, using our linearised Ricci tensor and scalar,

$$\begin{aligned} G_{\mu\nu} = \quad & \frac{1}{2}\left(\partial^\rho\partial_\mu h_{\nu\rho} + \partial^\rho\partial_\nu h_{\mu\rho} - \partial^\rho\partial_\rho h_{\mu\nu} - \partial_\nu\partial_\mu h \right. \\ & \left. -\eta_{\mu\nu}\partial^\sigma\partial^\pi h_{\pi\sigma} + \eta_{\mu\nu}\partial^\rho\partial_\rho h\right). \end{aligned} \tag{5.7}$$

Now, let us define

$$\bar{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h, \tag{5.8}$$

and that the *Lorentz gauge* is

$$\partial_\mu \bar{h}^{\mu\nu} = \partial^\mu \bar{h}_{\mu\nu} = 0. \tag{5.9}$$

That is,

$$\partial^\mu h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\partial^\mu h = 0,$$

which is just the statement that

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

Hence, using this in (5.7) (and swapping the $\partial^\mu\partial_\nu \leftrightarrow \partial_\nu\partial^\mu$ at will), we see that

$$G_{\mu\nu} = \frac{1}{2}\left(\frac{1}{2}\partial_\mu\partial_\nu h + \frac{1}{2}\partial_\nu\partial_\mu h - \partial^\rho\partial_\rho h_{\mu\nu} - \partial_\nu\partial_\mu h - \frac{1}{2}\eta_{\mu\nu}\partial^\sigma\partial_\sigma h + \eta_{\mu\nu}\partial^\sigma\partial_\sigma h\right).$$

Now, the first and second terms are identical, but their sum cancels with the fourth term. Hence,

$$G_{\mu\nu} = \frac{1}{2}\left(-\partial^\rho\partial_\rho h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\partial^\sigma\partial_\sigma h + \eta_{\mu\nu}\partial^\sigma\partial_\sigma h\right).$$

The second and third terms add, to give

$$G_{\mu\nu} = \frac{1}{2}\left(-\partial^\rho\partial_\rho h_{\mu\nu} + \frac{1}{2}\eta_{\mu\nu}\partial^\sigma\partial_\sigma h\right).$$

Now, if we use a little bit of notation,

$$\Box \equiv \partial^\mu\partial_\mu,$$

then

$$G_{\mu\nu} = -\frac{1}{2}\left(\Box h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\Box h\right),$$

or

$$G_{\mu\nu} = -\frac{1}{2}\Box\left(h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h\right).$$

Hence, using our substitution (5.8) again,

$$G_{\mu\nu} = -\frac{1}{2}\Box\bar{h}_{\mu\nu}.$$

Then, if we write down Einstein's equation,

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \quad \Rightarrow \quad \Box\bar{h}_{\mu\nu} = -16\pi G T_{\mu\nu}.$$

Therefore, we have a wave equation in the metric perturbation, with the energy-momentum tensor as the source. This is the equation for gravitational radiation.

5.4.2 Gravitational Radiation

Under the Lorentz gauge (to be inkeeping with the literature, this is sometimes also referred to as the Einstein gauge, or harmonic gauge),

$$\partial_\mu \bar{h}^{\mu\nu} = 0, \quad \bar{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} h,$$

Einstein's equation becomes

$$\Box \bar{h}_{\mu\nu} = -16\pi G T_{\mu\nu}, \quad \Box \equiv \partial_\mu \partial^\mu.$$

That is, a wave equation.

$$\Box \bar{h}_{\mu\nu} = -16\pi G T_{\mu\nu}. \tag{5.10}$$

We can write down the solution to this directly, if one recalls the solution to the equivalent equation from electrodynamic theory.

In electrodynamics, under the Lorentz gauge $\partial_\mu A^\mu = 0$, we could derive the wave equation

$$\Box A^\nu = \mu_0 J^\nu,$$

which has solution

$$A^i = \frac{\mu_0}{4\pi} \int d^3x' \frac{J^i_{\text{ret}}}{|\mathbf{x} - \mathbf{x}'|}.$$

Hence, we can basically read off our solution by analogy,

$$\bar{h}^{ij} = 4G \int d^3x' \frac{T^{ij}_{\text{ret}}}{|\mathbf{x} - \mathbf{x}'|}. \tag{5.11}$$

One should recall that these are retarded integrals. The minus sign has "gone" because we have raised indices.

Therefore, we have derived that upon linearising Einstein's equation, and using the Lorentz gauge, we have derived that there is a wave equation in the metric perturbation. The source to the wave is the distribution of energy-momentum.

6 The Schwarzschild Solution

We can write Einstein's equation, in a vacuum, as

$$R_{\mu\nu} = 0. \tag{6.1}$$

That is, in a vacuum, where $T_{\mu\nu} = 0$, the "alternative form" of Einstein's equation reduces to the above.

Now, we can look for *spherically symmetric* solutions to this. That is, we are looking for a line element which possesses spherical symmetry. The most general such line element is

$$ds^2 = e^{\nu(r,t)}dt^2 - e^{\lambda(r,t)}dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2).$$

The reason we make this supposedly general line element diagonal, is that we can transform out of a frame in which there are diagonal elements.

In the line element we chose to use exponentials, as they are generally easy to work with (differentiating them is easy). Hence, the aim is to now find those functions $\nu(r,t), \lambda(r,t)$.

Now, although we shall not derive them, the only non-zero components of the Ricci tensor are

$$\begin{aligned}
R_{tt} &= \frac{1}{2}e^{-\lambda}\left(\nu'' + \frac{1}{2}\nu'(\nu' - \lambda') + \frac{2\nu'}{r}\right) \\
&\quad + e^{-\nu}\left(\dot{\lambda}(\dot{\nu} - \dot{\lambda}) - \frac{1}{2}\ddot{\lambda}\right), \\
R_{tr} &= \frac{\dot{\lambda}}{2r}, \\
R_{rr} &= \frac{1}{2}e^{-\nu}\left(\ddot{\lambda} - \frac{1}{2}\dot{\lambda}(\dot{\nu} - \dot{\lambda})\right) \\
&\quad - \frac{1}{2}e^{-\lambda}\left(\nu'' + \frac{1}{2}\nu'(\nu' - \lambda') - \frac{2\lambda'}{r}\right), \\
R_{\theta\theta} &= 1 - e^{-\lambda}\left(1 + \frac{1}{2}r(\nu' - \lambda')\right), \\
R_{\phi\phi} &= \sin^2\theta R_{\theta\theta}.
\end{aligned}$$

We have used that an over-dot represents derivative with respect to time t, and a prime with respect to r.

Hence, due to the reduction of Einstein's equation to the form $R_{\mu\nu} = 0$, each of these equation are equal to 0.

The easiest to start with, is the R_{tr} term. So,

$$\frac{\dot{\lambda}}{2r} = 0,$$

which immediately allows us to state that $\lambda(r)$ only. That is, λ does not have any dependance upon t. Thus, using $\dot{\lambda} = 0$ allows R_{tt} and R_{rr} to look very similar. Infact, as $R_{rr} = R_{tt} = 0$, then $R_{rr} + R_{tt} = 0$. This then easily shows that

$$R_{tt} + R_{rr} = \frac{1}{2}e^{-\lambda}\left(\frac{2\nu'}{r} + \frac{2\lambda'}{r}\right) = 0,$$

that is, assuming $r \neq 0$,

$$\nu' + \lambda' = 0.$$

Integrating this easily shows that

$$\nu + \lambda = f(t).$$

Now, we can set $f(t)$ to zero, by a time coordinate transformation. Then, $\nu = -\lambda$. Therefore,

$$\nu(r) = -\lambda(r).$$

Hence, using this in $R_{\theta\theta}$, we see that

$$R_{\theta\theta} = 1 - e^{\nu}(1 + r\nu') = 0,$$

that is,

$$e^{\nu} + r\nu' e^{\nu} = 1.$$

Now, we can rewrite this as

$$(re^{\nu})' = e^{\nu} + r\nu' e^{\nu} = 1,$$

that is, as

$$(re^{\nu})' = 1.$$

Integrating easily reveals that

$$e^{\nu} = 1 + \frac{C}{r},$$

where C is some constant. We can find the value of C, by considering the Newtonian limit of the metric. That is, recall that we derived

$$g_{00} = 1 + 2\Phi,$$

where we know that

$$\Phi = -\frac{GM}{r}.$$

Now, $e^{\nu} = g_{00}$ by inspection (it is the coefficient of the dt^2 term). Hence,

$$1 - \frac{2GM}{r} = 1 + \frac{C}{r} \quad \Rightarrow \quad C = -2GM.$$

Let us recall that this M is the mass of the body generating the potential Φ. That is, it will be the mass of the planet/star that is curving the spacetime. Therefore,

$$e^{\nu} = 1 - \frac{2GM}{r}, \quad e^{\lambda} = \left(1 - \frac{2GM}{r}\right)^{-1}.$$

And finally, we have our metric,

$$ds^2 = \left(1 - \frac{2GM}{r}\right) dt^2 - \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{6.2}$$

That is, we have the vacuum solution of Einstein's equation, due to a body of mass M; where $r > 0$. This metric is called the *Schwarzschild metric*.

Properties of the Schwarzschild Metric The metric, by construction, is spherically symmetric. Also, the metric is static; it clearly is not a function of time. That the metric is static, then means that changing the time coordinate by a constant amount leaves the metric unchanged. That is, the metric is invariant under constant translations and reflections. Also notice that the metric has Killing vectors $(1, 0, 0, 0)$ and $(0, 0, 0, 1)$ (i.e. on t and ϕ); these correspond to conservation of energy and angular momentum.

Notice that as $r \to \infty$, the metric goes over to Minkowski. That is, we say that the metric is asymptotically flat.

Also notice, at $r = 2GM$, the g_{tt} and g_{rr} components flip sign. We call this the *Schwarzschild radius*, or the *event horizon*. We denote the event horizon as

$$r_{\mathrm{s}} \equiv 2GM. \tag{6.3}$$

6.0.3 Gravitational Redshift

Consider some radial slices in the metric, so that $ds^2 = g_{tt}dt^2$. Also consider that

$$d\tau = \frac{ds}{c},$$

hence,

$$d\tau = \sqrt{g_{tt}}\frac{dt}{c}.$$

Now, it is fairly obvious that a frequency is inversely proportional to the proper time. That is,

$$\nu \propto \frac{1}{\Delta\tau}.$$

Now, if we take two events which are at the same t, then

$$\frac{\nu_1}{\nu_2} = \sqrt{\frac{g_{tt}(2)}{g_{tt}(1)}}.$$

If we use a weak gravitational field, then we can use the previously derived relation $g_{tt} = 1 + 2\Phi$. Hence, this gives

$$\frac{\nu_1}{\nu_2} = 1 + \Phi(2) - \Phi(1).$$

That is, the shift in frequency is a function of the distance from the gravitating body.

6.1 Dynamics in the Schwarzschild Spacetime

Recall that the effective Lagrangian is

$$L_{\text{eff}} = \left(\frac{ds}{d\tau}\right)^2.$$

Therefore, the effective Lagrangian is

$$L_{\text{eff}} = \left(1 - \frac{r_{\text{s}}}{r}\right)\dot{t}^2 - \left(1 - \frac{r_{\text{s}}}{r}\right)^{-1}\dot{r}^2 - r^2(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2), \tag{6.4}$$

where an over-dot denotes derivative with respect to the affine parameter τ, and $r_{\text{s}} = 2GM$. So, let us consider the first integrals of this effective Lagrangian.

The Euler-Lagrange equations, for this effective Lagrangian are

$$\frac{d}{d\tau}\frac{\partial L_{\text{eff}}}{\partial \dot{x}^\mu} - \frac{\partial L_{\text{eff}}}{\partial x^\mu} = 0.$$

Then, consider that

$$\frac{\partial L_{\text{eff}}}{\partial t} = 0, \quad \frac{\partial L_{\text{eff}}}{\partial \dot{t}} = 2\left(1 - \frac{r_{\text{s}}}{r}\right)\dot{t},$$

then, the t-first integral is that

$$2\left(1 - \frac{r_{\text{s}}}{r}\right)\dot{t} = \text{const} \equiv 2\varepsilon.$$

Similarly,

$$\frac{\partial L_{\text{eff}}}{\partial \phi} = 0, \quad \frac{\partial L_{\text{eff}}}{\partial \dot{\phi}} = 2r^2\sin^2\theta\dot{\phi},$$

with its first integral being

$$2r^2\sin^2\theta\dot{\phi} = \text{const} \equiv 2\ell.$$

These constants, ε, ℓ, are related to the conserved energy and angular momentum, per unit mass. Recall that these were predicted to be conserved, by the associated Killing vectors.

Finally, the effective Lagrangian is just the line element, and that can take on one of 3 values;

$$L_{\text{eff}} = K = \begin{cases} 0 & \text{null}, \\ +1 & \text{time-like}, \\ -1 & \text{space-like}. \end{cases}$$

Hence, using the derived relations for ℓ, ε, we can easily see that

$$\dot{t}^2 = \varepsilon^2\left(1 - \frac{r_{\text{s}}}{r}\right)^{-2}, \quad \dot{\phi}^2 = \frac{\ell^2}{r^4\sin^4\theta}.$$

And thus, using that the effective Lagrangian is just a constant K, we can easily put the effective Lagrangian into the form

$$K = \left(1 - \frac{r_{\mathrm{s}}}{r}\right)^{-1}\left(\varepsilon^2 - \dot{r}^2\right) - r^2\left(\dot{\theta}^2 + \frac{\ell^2}{r^4\sin^2\theta}\right).$$

Now, as the system has spherical symmetry, we may as well take a value of θ that makes the above expression look simpler. Taking $\theta = \pi/2$ (note that then $\dot{\theta} = 0$), we see that

$$K = \left(1 - \frac{r_{\mathrm{s}}}{r}\right)^{-1}\left(\varepsilon^2 - \dot{r}^2\right) - \frac{\ell^2}{r^2},$$

which is trivially just

$$K = \left(1 - \frac{r_{\mathrm{s}}}{r}\right)^{-1}\left[\varepsilon^2 - \left(\frac{dr}{d\tau}\right)^2\right] - \frac{\ell^2}{r^2}.$$

Before we carry on with this expression, let us compute the Christofell symbols and geodesics.

6.1.1 Geodesics & Christofell Symbols

Let us compute the geodesics and Christofell symbols for the effective Lagrangian (6.4) in this Schwarzschild spacetime.

We can compute the geodesic for the θ-component of the effective Lagrangian. We have that

$$\frac{\partial L_{\mathrm{eff}}}{\partial\dot{\theta}} = -2r^2\dot{\theta}, \quad \frac{\partial L_{\mathrm{eff}}}{\partial\theta} = -2r^2\sin\theta\cos\theta\dot{\phi}^2,$$

hence,

$$\frac{d}{d\tau}\frac{\partial L_{\mathrm{eff}}}{\partial\dot{\theta}} = -4r\dot{r}\dot{\theta} - 2r^2\ddot{\theta}.$$

Therefore, the Euler-Lagrange equation for the θ-component, is

$$-4r\dot{r}\dot{\theta} - 2r^2\ddot{\theta} + 2r^2\sin\theta\cos\theta\dot{\phi}^2.$$

Putting this into a more usable form,

$$\ddot{\theta} + \frac{2}{r}\dot{r}\dot{\theta} - \sin\theta\cos\theta\dot{\phi}^2 = 0.$$

Thus, we have the geodesic for θ. Now, we can read off the Christofell symbols. The non-zero components are

$$\Gamma^{\theta}{}_{r\theta} = \Gamma^{\theta}{}_{\theta r} = \frac{1}{r}, \quad \Gamma^{\theta}{}_{\phi\phi} = -\sin\theta\cos\theta.$$

We can compute the geodesic for r. So,

$$\begin{aligned}
\frac{\partial L_{\mathrm{eff}}}{\partial\dot{r}} &= -2\left(1 - \frac{r_{\mathrm{s}}}{r}\right)^{-1}\dot{r}, \\
\frac{\partial L_{\mathrm{eff}}}{\partial r} &= \dot{t}^2\frac{r_{\mathrm{s}}}{r^2} - \frac{r_{\mathrm{s}}}{r^2}\left(1 - \frac{r_{\mathrm{s}}}{r}\right)^{-2}\dot{r}^2 - 2r(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2),
\end{aligned}$$

and

$$\frac{d}{d\tau}\frac{\partial L_{\text{eff}}}{\partial \dot{r}} = -2\ddot{r}\left(1-\frac{r_{\text{s}}}{r}\right)^{-1} + 2\dot{r}^2\frac{r_{\text{s}}}{r^2}\left(1-\frac{r_{\text{s}}}{r}\right)^{-2}.$$

Therefore, the geodesic is

$$-2\ddot{r}\left(1-\frac{r_{\text{s}}}{r}\right)^{-1} + 2\dot{r}^2\frac{r_{\text{s}}}{r^2}\left(1-\frac{r_{\text{s}}}{r}\right)^{-2} - \dot{t}^2\frac{r_{\text{s}}}{r^2} - \frac{r_{\text{s}}}{r^2}\left(1-\frac{r_{\text{s}}}{r}\right)^{-2}\dot{r}^2$$
$$+2r(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2) = 0.$$

This simplifies down to

$$\ddot{r} - \frac{\dot{r}^2 r_{\text{s}}}{2r^2}\left(1-\frac{r_{\text{s}}}{r}\right)^{-1} + \frac{\dot{t}^2 r_{\text{s}}}{2r^2}\left(1-\frac{r_{\text{s}}}{r}\right) + r\left(1-\frac{r_{\text{s}}}{r}\right)(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2) = 0.$$

This is the r-geodesic. From this, we can read off the non-zero Christofell symbols. They are

$$\Gamma^r{}_{rr=} - \frac{r_{\text{s}}}{2r^2}\left(1-\frac{r_{\text{s}}}{r}\right)^{-1}, \quad \Gamma^r{}_{tt=}\frac{r_{\text{s}}}{2r^2}\left(1-\frac{r_{\text{s}}}{r}\right),$$
$$\Gamma^r{}_{\theta\theta} = r\left(1-\frac{r_{\text{s}}}{r}\right), \quad \Gamma^r{}_{\phi\phi} = r\sin^2\theta\left(1-\frac{r_{\text{s}}}{r}\right).$$

Then, let us compile these four geodesics (i.e. including the two not explicitly computed here). The geodesics for the Schwarzschild spacetime are:

$$\ddot{t} + \frac{r_{\text{s}}}{r^2}\left(1-\frac{r_{\text{s}}}{r}\right)^{-1}\dot{t}\dot{r} = 0,$$
$$\ddot{r} - \frac{r_{\text{s}}}{2r^2}\left(1-\frac{r_{\text{s}}}{r}\right)^{-1}\dot{r}^2 + \frac{r_{\text{s}}}{2r^2}\left(1-\frac{r_{\text{s}}}{r}\right)\dot{t}^2 + r\left(1-\frac{r_{\text{s}}}{r}\right)(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2) = 0,$$
$$\ddot{\theta} + \frac{2}{r}\dot{r}\dot{\theta} - \sin\theta\cos\theta\dot{\phi}^2 = 0,$$
$$\ddot{\phi} + \frac{2}{r}\dot{r}\dot{\phi} + 2\cot\theta\dot{\theta}\dot{\phi} = 0.$$

These complicated non-linear differential equations can be solved to find the trajectories of particles in the spacetime. The non-zero Christofell symbols are easily read off, and can be seen to be

$$\Gamma^t{}_{rt} = \frac{r_{\text{s}}}{2r^2}\left(1-\frac{r_{\text{s}}}{r}\right)^{-1}, \quad \Gamma^r{}_{rr=} - \frac{r_{\text{s}}}{2r^2}\left(1-\frac{r_{\text{s}}}{r}\right)^{-1},$$
$$\Gamma^r{}_{tt=}\frac{r_{\text{s}}}{2r^2}\left(1-\frac{r_{\text{s}}}{r}\right),$$
$$\Gamma^r{}_{\theta\theta} = r\left(1-\frac{r_{\text{s}}}{r}\right), \quad \Gamma^r{}_{\phi\phi} = r\sin^2\theta\left(1-\frac{r_{\text{s}}}{r}\right),$$
$$\Gamma^\theta{}_{r\theta} = \Gamma^\theta{}_{\theta r} = \frac{1}{r}, \quad \Gamma^\theta{}_{\phi\phi} = -\sin\theta\cos\theta,$$
$$\Gamma^\phi{}_{r\phi} = \frac{1}{r}, \quad \Gamma^\phi{}_{\theta\phi} = \cot\theta.$$

6.1.2 Orbits

Let us return to the expression we derived, for $\theta = \pi/2$,

$$K = \left(1 - \frac{r_\mathrm{s}}{r}\right)^{-1} \left[\varepsilon^2 - \left(\frac{dr}{d\tau}\right)^2\right] - \frac{\ell^2}{r^2}.$$

We can rearrange it into the form

$$\dot{r}^2 = \varepsilon^2 - K - \left[\frac{\ell^2}{r^2}\left(1 - \frac{r_\mathrm{s}}{r}\right) - \frac{K r_\mathrm{s}}{r}\right],$$

and indeed into the form

$$\frac{1}{2}\dot{r}^2 = \frac{\varepsilon^2 - K}{2} - \left[\frac{\ell^2}{2r^2}\left(1 - \frac{r_\mathrm{s}}{r}\right) - \frac{K r_\mathrm{s}}{2r}\right].$$

Now, we put it into this form, as we see that the LHS is a "velocity term", the middle term is just the "energy", and the far-RHS we call the effective potential V_eff:

$$E = \frac{1}{2}\dot{r}^2 + V_\mathrm{eff}(r),$$

where

$$V_\mathrm{eff} \equiv \frac{\ell^2}{2r^2}\left(1 - \frac{r_\mathrm{s}}{r}\right) - \frac{K r_\mathrm{s}}{2r}. \tag{6.5}$$

Now, one familiar with the Newtonian derivation of this formula, will realise that this expression is not quite the same as its Newtonian counterpart. The GR "correction" is the r_s/r, creating a $1/r^3$ term.

Just to recap what the symbols are in this effective potential. ℓ is the angular momentum of the "moving thing", $r_\mathrm{s} \equiv 2GM$, where M is the mass of the "big body" that the "moving thing" is moving in. That is, the big body is curving spacetime, and some moving object is having its motion deflected, by the curved spacetime, which is due to the big body. The amount of deflection is just a function of the distance from the big body to the smaller one. We shall call the "smaller body" the *test mass*, and the "big body" the *gravitating mass*.

Suppose that $\ell = 0$. Then,

$$V_\mathrm{eff} = -\frac{K r_\mathrm{s}}{2r} = -K\frac{2GM}{2r} = -K\frac{GM}{r}.$$

This is the Newtonian result. That is, for a test mass with no angular momentum, the effective potential is just what we would expect.

Recall that we derived

$$\varepsilon = \left(1 - \frac{r_\mathrm{s}}{r}\right)\frac{dt}{d\tau},$$

then, we can clearly see that

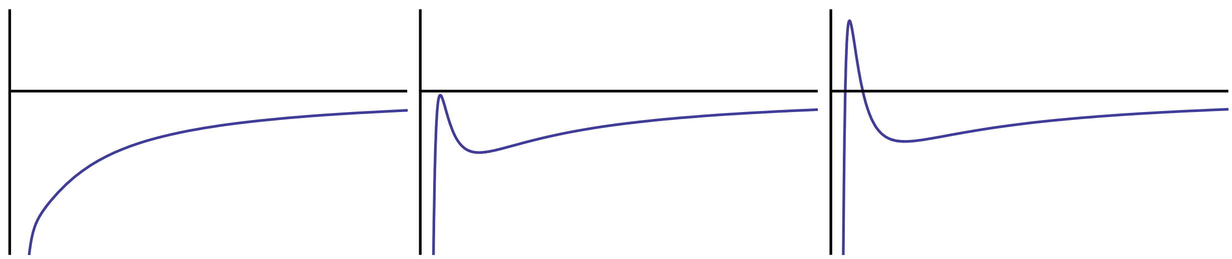
$$\frac{dt}{d\tau} = \varepsilon \left(1 - \frac{r_s}{r}\right)^{-1}.$$

That is, the proper time of a test mass is a function of the distance from the gravitating mass, and of the total energy.

Let us now give some results relating to orbits in the spacetime. Circular orbits have

$$\frac{dV_{\text{eff}}}{dr} = 0,$$

and stable circular orbits are those for whom the second differential of the potential is positive.



(a) $\ell < \sqrt{3}r_s$ (b) $\sqrt{3}r_s < \ell < 2r_s$ (c) $\ell > 2r_s$

Figure 6.1: The effective potential, as a function of distance from the gravitating body, for particle orbits.

Particle Orbits $K = 1$ If we vary the angular momentum, ℓ, with respect to the event horizon, r_s, then various shapes of effective potential are found. With reference to Figure (6.1), we see the 3 ranges of ℓ.

- $\ell < \sqrt{3}r_s$. Here, we see that any particle with energy $E > 0$, escapes, whilst any particle with $E < 0$ crushes back into the origin. No stable orbits exist.

- $\sqrt{3}r_s < \ell < 2r_s$. For this range, there are two positions in which orbits can exist, but only one of then is capable of sustaining stable orbits. If $E > 0$, then any particle will escape. If we define V_{max} as the position of the maximum of V_{eff}, and V_{min} as the minimum, then we can see that for any $0 > E > V_{\text{max}}$, a particle will crush into the origin. Also, for a particle with $E = V_{\text{min}}$, then there is a stable circular orbit. $E = V_{\text{max}}$ is an unstable circular orbit. Any particle trapped in the "well" will have some sort of elliptical orbit.

- $\ell > 2r_s$. Here, if $E = V_{\text{min}}$, the particle will have a stable circular orbit, and elliptical for perturbations about that minimum. If a particle has $E < V_{\text{max}}$, and lives to the left of the maximum, then it will crush into the origin. Now, if a particle has $E < V_{\text{max}}$, and approaches the system from the right of the maximum, then the particle will be repelled back to ∞. However, above a certain value, the particle will hit the origin. This is not present in Newtonian gravity, where particles are always repelled.

Photon Orbits $K = 0$ In this case, we have that

$$V_{\text{eff}} \equiv \frac{\ell^2}{2r^2}\left(1 - \frac{r_{\text{s}}}{r}\right).$$

Upon plotting the effective potential, we see that for a given $E < V_{\text{max}}$, it depends on where

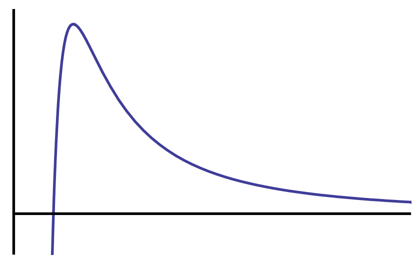


Figure 6.2: The effective potential, as a function of distance from the gravitating body, all for photon orbits.

the photon is, relative to the peak. That is, if the photon is within the peak, the photon will crush to the origin. If the photon is outside, then the photon will repel to infinity.

6.1.3 Summary

Let us just summarise the results obtained, as they will be useful in subsequent discussions.

We derived that, on a $\theta = \pi/2$ trajectory,

$$\frac{1}{2}\dot{r}^2 = E - V_{\text{eff}},$$

where the "energy" is given by

$$E = \frac{\varepsilon^2 - K}{2},$$

and the effective potential by

$$V_{\text{eff}} = \frac{\ell^2}{2r^2}\left(1 - \frac{r_{\text{s}}}{r}\right) - \frac{K r_{\text{s}}}{2r}.$$

The angular momentum of the test mass was computed to be

$$\ell = r^2\dot{\phi},$$

and the energy density

$$\varepsilon = \left(1 - \frac{r_s}{r}\right) \dot{t}.$$

Light-like trajectories are those for whom $K = 0$. Particle-like are those for whom $K = 1$. The event horizon is related to the mass of the gravitating body $r_s = 2GM$, and is idealised so that all mass is concentrated at a single point. Over-dots represent derivative with respect to the affine parameter. Notice that we can then easily write that

$$\begin{aligned} \frac{d\phi}{dr} &= \frac{\dot{\phi}}{\dot{r}} \\ &= \pm\frac{1}{r^2}\frac{\ell}{\sqrt{2}(E - V_{\text{eff}})^{1/2}}. \end{aligned} \tag{6.6}$$

6.2 Light Deflection

We can compute the angle that light is deflected by, due to the curved spacetime of a star.

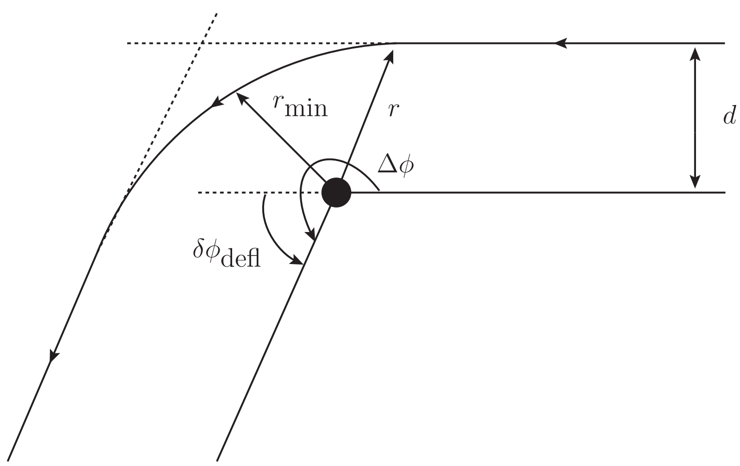


Figure 6.3: Light deflection due to a gravitating mass. Notice how various angles are defined. d is the impact parameter of the photon, with respect to the radius of the gravitating mass.

The effective potential, for photons with $K = 0$, reads

$$V_{\text{eff}} = \frac{\ell^2}{2r^2}\left(1 - \frac{r_s}{r}\right). \tag{6.7}$$

Consider the combination

$$\frac{\ell}{\varepsilon} = \frac{r^2\dot{\phi}}{\left(1 - \frac{r_s}{r}\right)\dot{t}},$$

then, considering that $r \gg r_\mathrm{s}$, then

$$\frac{\ell}{\varepsilon} \approx r^2 \frac{d\phi}{dt} + \mathcal{O}\left(\frac{r_\mathrm{s}}{r}\right) \quad \Rightarrow \quad \frac{\ell}{\varepsilon} = r^2 \frac{d\phi}{dt}. \tag{6.8}$$

Now, for small angles, we have that

$$\phi = \frac{d}{r}.$$

Then,

$$\frac{d\phi}{dt} = -\frac{d}{r^2}\frac{dr}{dt}.$$

Now,

$$\frac{dr}{dt} = -1,$$

where the unity comes from $c = 1$, and the minus-sign because distances are shrinking. Hence,

$$\frac{d\phi}{dt} = \frac{d}{r^2},$$

which we use in (6.8) to see that

$$\frac{\ell}{\varepsilon} = d.$$

Therefore, for photons,

$$d = \frac{\ell}{\varepsilon} = \frac{\ell}{\sqrt{2E}}. \tag{6.9}$$

Now, with reference to Figure (6.3), we see that the deflection angle is given by

$$\delta\phi_\mathrm{defl} = \Delta\phi - \pi.$$

The total angle change is just the integral

$$\Delta\phi = \int d\phi,$$

or, as we have an expression for $d\phi/dr$,

$$\Delta\phi = \int dr \frac{d\phi}{dr}.$$

Hence, using (6.6), we have that

$$\Delta\phi = 2\int_{r_\mathrm{min}}^{r_\mathrm{max}} dr \frac{1}{r^2}\frac{\ell}{\sqrt{2}(E - V_\mathrm{eff})^{1/2}}, \tag{6.10}$$

using the light-like effective potential (6.7),

$$\Delta\phi = 2\int_{r_\mathrm{min}}^{r_\mathrm{max}} dr \frac{1}{r^2}\frac{\ell}{\sqrt{2}\left[E - \frac{\ell^2}{2r^2}\left(1 - \frac{r_\mathrm{s}}{r}\right)\right]^{1/2}}.$$

We take $r_{\max} \to \infty$, and note that the factor of 2 out-front is due to the photon coming from infinity, the going back to infinity. If we put the factor of ℓ inside the square-root in the denominator, as well as the $\sqrt{2}$, then

$$\Delta\phi = 2\int_{r_{\min}}^{\infty} \frac{dr}{r^2}\left[\frac{2E}{\ell^2} - \frac{1}{r^2}\left(1 - \frac{r_{\mathrm{s}}}{r}\right)\right]^{-1/2}.$$

Now, noting that via (6.9), we rewrite

$$\frac{2E}{\ell^2} = \frac{1}{d^2},$$

and also change variables to

$$w \equiv \frac{d}{r} \quad \Rightarrow \quad dw = -\frac{d}{r^2}dr.$$

Hence, using this change of variables, and rewrite,

$$\Delta\phi = 2\int_{w_{\max}}^{0} -\frac{dw}{d}\left[\frac{1}{d^2} - \frac{w^2}{d^2}\left(1 - \frac{r_{\mathrm{s}}w}{d}\right)\right]^{-1/2},$$

the minus sign obviously flipping the integration limits to

$$\Delta\phi = 2\int_{0}^{w_{\max}} \frac{dw}{d}\left[\frac{1}{d^2} - \frac{w^2}{d^2}\left(1 - \frac{r_{\mathrm{s}}w}{d}\right)\right]^{-1/2}.$$

The factor of $\frac{1}{d}$ can be taken inside the square-root, giving

$$\Delta\phi = 2\int_{0}^{w_{\max}} dw\left[1 - w^2\left(1 - \frac{r_{\mathrm{s}}w}{d}\right)\right]^{-1/2}.$$

Now, if we refer to (6.10), we see that there is a singularity at $E = V_{\mathrm{eff}}$. It is an integral of the form

$$\int_0^1 \frac{dx}{\sqrt{x+\epsilon}} \approx \int_0^1 \frac{dx}{\sqrt{x}} - \frac{\epsilon}{2}\int_0^1 \frac{dx}{x^{3/2}},$$

whereby upon integration, the first term does not give a singularity, but the second does (at zero). Thus, we say that the integral has an *essential singularity*.

Let us continue. If we take out a factor, from the square root, then

$$\Delta\phi = 2\int_{0}^{w_{\max}} dw\left(1 - \frac{r_{\mathrm{s}}}{d}w\right)^{-1/2}\left[\left(1 - \frac{r_{\mathrm{s}}}{d}w\right)^{-1} - w^2\right]^{-1/2}.$$

Now, we can expand the two terms,

$$\left(1 - \frac{r_{\mathrm{s}}}{d}w\right)^{-1/2} = 1 + \frac{r_{\mathrm{s}}}{2d}w + \mathcal{O}\left(\frac{r_{\mathrm{s}}}{d}\right)^2, \quad \left(1 - \frac{r_{\mathrm{s}}}{d}w\right)^{-1} = 1 + \frac{r_{\mathrm{s}}}{d}w + \mathcal{O}\left(\frac{r_{\mathrm{s}}}{d}\right)^2,$$

so that

$$\Delta\phi = 2\int_0^{w_{\rm max}} dw\left(1+\frac{r_{\rm s}}{2d}w\right)\left(1+\frac{r_{\rm s}}{d}w-w^2\right)^{-1/2}+\mathcal{O}\left(\frac{r_{\rm s}}{d}\right)^2.$$

We can obviously now multiply out the bracket,

$$\Delta\phi = 2\int_0^{w_{\rm max}} dw\left(1+\frac{r_{\rm s}}{d}w-w^2\right)^{-1/2}+\frac{r_{\rm s}}{d}\int_0^{w_{\rm max}} dw w(1-w^2)^{-1/2}.$$

Now, we can see that there is a pole in the second integral, at $w_{\rm max}=1$. If we look up the values of the integrals, we find

$$\begin{aligned}\int_0^{w_{\rm max}} dw\left(1+\frac{r_{\rm s}}{d}w-w^2\right)^{-1/2} &= \frac{\pi}{2}+\frac{r_{\rm s}}{2d},\\ \int_0^{w_{\rm max}} dw w(1-w^2)^{-1/2} &= 1.\end{aligned}$$

Therefore, we see that

$$\Delta\phi = \pi+\frac{2r_{\rm s}}{d},$$

and hence, the deflection angle,

$$\delta\phi_{\rm defl} = \frac{2r_{\rm s}}{d}. \tag{6.11}$$

Therefore, we have derived the deflection angle of a photons trajectory, with impact parameter d with respect to a gravitating body of event horizon $r_{\rm s}$.

To get a handle on how big this angle is, consider the Sun. $r_{\rm s}\approx 3$km, and suppose the photon just grazes the suns surface. Then,

$$\delta\phi_\odot = \frac{2r_{\rm s}}{d} = \frac{2.3\text{km}}{7\times 10^{-5}\text{km}} \approx 10^{-5}\text{rad}.$$

This angle is equivalent to the observed height of a 1m high object, viewed from 10km away. That is, the effect is very small. However, this angle can be measured (best in solar eclipses), and has been confirmed to be closer to the actual value than the Newtonian prediction (which is a factor of 4 smaller).

This is one of the tests of general relativity.

6.3 Perihelion Precession

Here we consider the motion of a planet, about a star. Supposing that the orbit of the planet is elliptical, and that the "size" of the orbit is unchanged over may periods, does the "position" of the orbit change? That is, after each revolution, let us consider that $r_{\rm min}$ is the same, but is shifted in position by $\delta\phi_{\rm prec}$. Then, we have that

$$\Delta\phi = \delta\phi_{\rm prec}-2\pi,$$

where we use 2π to make the Newtonian prediction give $\Delta\phi = 0$.

We follow a similar tack as for light deflection, but we must take $K = 1$ as we are dealing with time-like objects. So, the effective potential is now

$$V_{\text{eff}} = \frac{\ell^2}{2r^2}\left(1 - \frac{r_{\text{s}}}{r}\right) - \frac{r_{\text{s}}}{2r}.$$

We also use (6.6)

$$\frac{d\phi}{dr} = \frac{1}{r^2}\frac{\ell}{\sqrt{2}(E - V_{\text{eff}})^{1/2}}, \tag{6.12}$$

where the energy for time-like objects is

$$E = \frac{\varepsilon^2 - 1}{2}.$$

Then, we write, as before,

$$\Delta\phi = 2\int_{r_{\text{min}}}^{r_{\text{max}}} dr \frac{d\phi}{dr},$$

which is just

$$\Delta\phi = 2\int_{r_{\text{min}}}^{r_{\text{max}}} dr \frac{\ell}{r^2}\left[\sqrt{2}(E - V_{\text{eff}})^{1/2}\right]^{-1},$$

putting in the effective potential,

$$\Delta\phi = 2\int_{r_{\text{min}}}^{r_{\text{max}}} dr \frac{\ell}{r^2}\left[2E - \frac{\ell^2}{r^2}\left(1 - \frac{r_{\text{s}}}{r}\right) + \frac{r_{\text{s}}}{r}\right]^{-1/2}.$$

If we now take the ℓ inside the square-root, and use the expression for E, then

$$\Delta\phi = 2\int_{r_{\text{min}}}^{r_{\text{max}}} dr \frac{1}{r^2}\left[\frac{\varepsilon^2}{\ell^2} - \frac{1}{\ell^2} - \frac{1}{r^2}\left(1 - \frac{r_{\text{s}}}{r}\right) + \frac{r_{\text{s}}}{r\ell^2}\right]^{-1/2}.$$

Let us rewrite the square-rooted bit slightly,

$$\frac{\varepsilon^2}{\ell^2} - \frac{1}{r^2}\left(1 - \frac{r_{\text{s}}}{r}\right) - \frac{1}{\ell^2}\left(1 - \frac{r_{\text{s}}}{r}\right).$$

Let us change integration variables,

$$u \equiv \frac{1}{r},$$

hence,

$$\Delta\phi = 2\int_{u_{\text{min}}}^{u_{\text{max}}} du \left[\frac{\varepsilon^2}{\ell^2} - u^2(1 - r_{\text{s}}u) - \frac{1}{\ell^2}(1 - r_{\text{s}}u)\right]^{-1/2}.$$

If we take out a common factor,

$$\Delta\phi = 2\int_{u_{\min}}^{u_{\max}} du(1-r_{\rm s}u)^{-1/2}\left[\frac{\varepsilon^2}{\ell^2}(1-r_{\rm s}u)^{-1}-\frac{1}{\ell^2}-u^2\right]^{-1/2}.$$

We now expand out, but we must take to higher order within the expression on the right,

$$\Delta\phi = 2\int_{u_{\min}}^{u_{\max}} du\left(1+\frac{r_{\rm s}u}{2}\right)\left[\frac{\varepsilon^2}{\ell^2}\left(1+r_{\rm s}u+r_{\rm s}^2u^2\right)-\frac{1}{\ell^2}-u^2\right]^{-1/2},$$

collecting terms,

$$\Delta\phi = 2\int_{u_{\min}}^{u_{\max}} du\left(1+\frac{r_{\rm s}u}{2}\right)\left[\frac{\varepsilon^2}{\ell^2}\left(1+r_{\rm s}u\right)-\frac{1}{\ell^2}-u^2\left(1-\frac{\varepsilon^2 r_{\rm s}^2}{\ell^2}\right)\right]^{-1/2},$$

thus,

$$\begin{aligned}\Delta\phi = 2\left(1+\frac{\varepsilon^2 r_{\rm s}^2}{2\ell^2}\right)&\int_{u_{\min}}^{u_{\max}} du\left[\frac{\varepsilon^2}{\ell^2}(1+r_{\rm s}u)-\frac{1}{\ell^2}-u^2\right]^{-1/2}\\ &+r_{\rm s}\int_{u_{\min}}^{u_{\max}} du u\left[\frac{\varepsilon^2}{\ell^2}(1+r_{\rm s}u)-\frac{1}{\ell^2}-u^2\right]^{-1/2}.\end{aligned}$$

Now, by looking up the integrals, the first gives π, the second $\frac{\pi}{2}(u_{\min}+u_{\max})$. Now, the integrand on the second integral has poles at the integration limits. Therefore, one can easily see that the sum of the roots of the integrand, is

$$\frac{\varepsilon^2}{\ell^2}r_{\rm s},$$

and therefore

$$\Delta\phi = 2\pi\left(1+\frac{\varepsilon^2 r_{\rm s}^2}{2\ell^2}\right)+\frac{\pi\varepsilon^2 r_{\rm s}^2}{2\ell^2}.$$

Hence, we read off

$$\delta\phi_{\rm prec} = \frac{3\pi r_{\rm s}^2}{2\ell^2} = \frac{6\pi G^2M^2}{\ell^2}.$$

Now, in getting a handle on how big this is, we appeal to standard ellipse-theory. The result of which allows us to write the angular momentum ℓ in terms of the semi-major axis a of the orbit, and the eccentricity e,

$$\ell^2 = GMa(1-e^2).$$

Hence, the precession angle reads

$$\delta\phi_{\rm prec} = \frac{6\pi GM}{a(1-e^2)}. \tag{6.13}$$

See Table (6.1) for a comparison of the prediction and observations of these precession angles.

Planet	**GR Prediction (per century)**	**Observation**
Mercury	$43''$	$43.1 \pm 0.5''$
Venus	$8.6''$	$8.4 \pm 4.8''$
Earth	$3.8''$	$5.0 \pm 1.2''$

Table 6.1: The GR prediction of, and experimental observation of, the perihelion precession of various planets. The agreement is one of the most convincing experimental "proofs" of general relativity.

6.4 Black Holes

Let us consider what the mass and radius is, of a gravitating body for whom the escape velocity is the speed of light. That is, what is M, R for which $v_{\text{esc}} = c$?

Recall that the Newtonian expression for total energy is

$$E_N = \frac{1}{2}mv^2 - \frac{GMm}{r},$$

so that rearranging into the familiar form

$$\frac{1}{2}\left(\frac{dr}{dt}\right)^2 = \frac{E_N}{m} - \left(-\frac{GM}{r}\right), \quad v = \frac{dr}{dt},$$

we see the presence of the effective potential. Now, escape velocity is when $E_N = 0$, which corresponds to

$$v_{\text{esc}}^2 = \frac{2GM}{R},$$

which we require to be c^2, which, under the units of $c = 1$, is just the statement that

$$R = 2GM = r_{\text{s}}.$$

That is, we seem to have derived the Schwarzschild radius (which was a GR result) using Newtonian mechanics. This is actually just a coincidence, as we have neglected both SR and GR (i.e. no mention of mass-energy in the above derivation).

Let us return to the Schwarzschild metric, with the assumption that θ, ϕ are constant. Then, it reads

$$ds^2 = \left(1 - \frac{r_{\text{s}}}{r}\right) dt^2 - \left(1 - \frac{r_{\text{s}}}{r}\right)^{-1} dr^2.$$

Notice that this expression has two singularities. One at $r = r_{\text{s}}$ and one at $r = 0$.

Now, it is not immediately obvious whether these singularities are an artifact of how we have constructed our coordinate system, or if they are "true singularities". So, a way of finding this out, would be to construct a quantity that is invariant of coordinate system. Such a quantity is of course a scalar. Now, we want a scalar that is dependent upon the geometry of the system. Such quantities are the contracted Riemann and Ricci tensors, and the Ricci scalar. Now, experience has shown us that the best test is the Riemann tensor, in the form

$$R^{\alpha\beta\nu\mu}R_{\alpha\beta\nu\mu} = \frac{6r_{\rm s}^2}{r^6}.$$

That is, we see that this coordinate-system independent quantity does not have a singularity as $r \to r_{\rm s}$, but does have one for $r \to 0$.

Therefore, we see that $r \to r_{\rm s}$ is a removable axis singularity, whereby we can change coordinates so that the metric does not retain the singularity; and that $r \to 0$ is an essential singularity. Now, although we shall not go into it at all, a quantum theory of gravity will be able to "sort out" this essential singularity.

6.4.1 Null Geodesics

Let us consider the case $\ell = 0$, and $ds^2 = 0$. Then, the metric is just

$$\left(1 - \frac{r_{\rm s}}{r}\right) dt^2 - \left(1 - \frac{r_{\rm s}}{r}\right)^{-1} dr^2 = 0,$$

which trivially rearranges into

$$\left(\frac{dr}{dt}\right)^2 = \left(1 - \frac{r_{\rm s}}{r}\right)^2,$$

which is just

$$\frac{dr}{dt} = \pm\left(1 - \frac{r_{\rm s}}{r}\right).$$

Notice that this is the radial geodesic. So, we can solve this,

$$\begin{aligned} t &= \pm \int \frac{dr}{1 - r_{\rm s}/r} \\ &= \pm \int \frac{r\,dr}{r - r_{\rm s}} \\ &= \pm \int dr \left(1 + \frac{r_{\rm s}}{r - r_{\rm s}}\right) \\ &= \pm \left[r + r_{\rm s} \ln|r - r_{\rm s}| + \text{const}\right] \\ &= \pm \left[r + r_{\rm s} \ln\left|\frac{r}{r_{\rm s}} - 1\right| + \text{const}\right]. \end{aligned}$$

Now, we define the *tortoise coordinate*

$$r_* \equiv r + r_{\rm s} \ln\left|\frac{r}{r_{\rm s}} - 1\right|, \tag{6.14}$$

so that the geodesic reads

$$t = \pm r_* + \text{const}$$

Now, notice that for flat space, $r_{\mathrm{s}} \to 0$. Hence, the geodesics read

$$t = \pm r + \text{const}.$$

Hence, we denote this as

$$u = t - r, \quad v = t + r,$$

so that lines of $u = \text{const}$ and $v = \text{const}$ define the null geodesics. See Figure (6.4) for these lines.

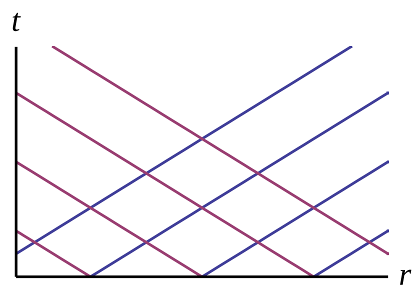


Figure 6.4: Null geodesics for flat space. Blue (left to right) lines are $u = \text{const}$, Red (right to left) lines are $v = \text{const}$. Photons move on these lines, and massive particles move within a light cone, defined by the lines. That is, the light cone is defined at an intersection of lines of $v = \text{const}$ and $u = \text{const}$; where the particles future is everything above that point, within that cone, and its past is everything below that point, within the cone.

We say that u, v are the light-cone coordinates for flat space. Hence, for flat space, the metric is

$$ds^2 = dt^2 - dr^2 - r^2(d\theta^2 + \sin^2 d\phi^2)$$

Now, notice that

$$t = \frac{1}{2}(u + v), \quad r = \frac{1}{2}(v - u).$$

Also that

$$dr = \frac{dr}{du}du + \frac{dr}{dv}dv = \frac{1}{2}(dv - du), \quad dt = \frac{1}{2}(dv + du).$$

Therefore, the metric reads

$$ds^2 = dudv - r^2(d\theta^2 + \sin^2 d\phi^2),$$

which is no longer diagonal.

6.4.2 Eddington-Finkelstein Coordinates

Now, let us return computing the null geodesics, but for curved space. We shall still use the light-cone coordinates,

$$u = t - r_*, \quad v = t + r_*,$$

with the tortoise coordinate

$$r_* = r + r_s \ln\left|\frac{r}{r_s} - 1\right|. \tag{6.15}$$

From which we can compute

$$\frac{dr_*}{dr} = \frac{r}{r - r_s} \quad \Rightarrow \quad dr^2 = \left(1 - \frac{r_s}{r}\right)^2 dr_*^2.$$

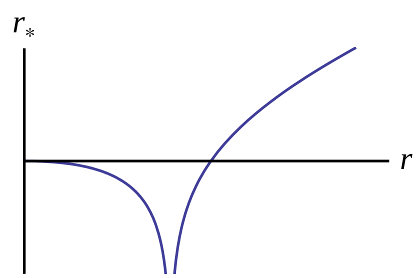


Figure 6.5: The tortoise coordinate (6.15). The position of r_s is obvious.

Now, the Schwarzschild metric may be written as (where we are suppressing the angular part)

$$ds^2 = \left(1 - \frac{r_s}{r}\right)\left[dt^2 - \frac{dr^2}{\left(1 - \frac{r_s}{r}\right)^2}\right],$$

which, using our derived relation for dr_*^2. is

$$ds^2 = \left(1 - \frac{r_s}{r}\right)\left[dt^2 - dr_*^2\right].$$

This, in terms of u, v is just

$$ds^2 = \left(1 - \frac{r_s}{r}\right) du dv.$$

Notice that this metric is no longer singular at $r = r_s$, but is still singular at $r = 0$.

With reference to Figure (6.6), we can see the geodesics for curved spacetime. We have plotted the lines $u = \text{const}$ and $v = \text{const}$. The interesting things to note from the plot:

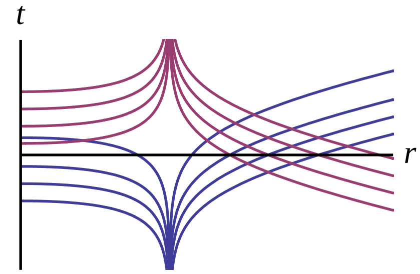


Figure 6.6: The null geodesics for curved spacetime. Blue lines are $u =$ const and red lines are $v =$ const. The future direction, for a light cone, is that were a red line is on the left, and a blue line on the right. Notice that for $r > r_s$, all future cones are pointing upwards, and that at $r < r_s$, all future cones point leftwards.

- As r decreases towards r_s, the angle between a u and v line decrease. This means that the future (and past) light cone of a particle becomes sharper. This means that "stuff" must be closer to the particle for it to influence the particle, as the particle gets closer to the Schwarzschild radius.

- As a particle crosses $r = r_s$, light cones flip 90°, and point towards the t-axis. That is, the future of the particle can only be for motion towards the origin. That is, the particle can never escape.

Therefore, we have seen that as a particle crosses the Schwarzschild radius, its light cone gets tilted so that its future is always within the Schwarzschild radius. That is, particles can get into this region, but never out.

Thus, we see that $r = r_s$ is some sort of membrane which allows one-way travel. This is the event horizon.

Therefore, we see how black holes "work". We have only considered stationary black holes. To consider rotating holes, one must analyse the Kerr metric, which we shall not do here.

Hawking Radiation Now, classically, particles cannot escape from a black hole, as we have just seen. However, quantum mechanically, they can tunnel out. According to quantum field theory, there is a "sea" of particle-anti-particle pairs being created and annihilated all the time, in vacuum (i.e. there is no true vacuum). Now, suppose one of these pairs were created on the event horizon, so that one of the particle gets created inside the horizon, one out side. Then, as the particle inside cannot get out (it is inside the horizon), then it cannot annihilate with the one that was created outside the horizon. Therefore, the particle outside the horizon can escape. Now, the energy to create the particle-anti-particle pair came from the vacuum inside the horizon. Therefore, by the particle escaping, energy is removed from the black hole, and over time, the black hole evaporates. This is called *Hawking radiation.*

To properly understand this radiation requires a huge amount of QFT, which we shall not go into here.

This effect can be conceived in a rather tamer environment. Consider two metal plates, which posses opposite electric charge, where the space between the plates is “vacuum”. Now, the energy density due to the electric field may be ramped up so that it is high enough to create an electron-positron pair from the vacuum. This experiment, as far as I am aware, has not been done, but it is conceivable to see that it could (if the idea of a sea of virtual particles is correct).

7 The Friedmann-Robertson-Walker Universe

We shall abbreviate the above name to FRW.

Now, we can start to consider the geometry of our universe. Historically, there were two theories for the universe.

The *FRW universe* was one based upon the cosmological principle: "*Our universe is homogenous and isotropic.*" This means that the universe is pretty much the same everywhere you look, and in any direction. That is, the ensemble properties of the universe are invariant under both translation and rotation.

The competing theory was that of a *steady state universe*, proposed in 1948 by F.Hoyle, H.Bondi and T.Gold. The steady state theory was a more "perfect" version of the cosmological principle, by imposing a condition that the universe be invariant under time as well as translation/rotation. This means that the universe looks the same at any time.

The main differences between the theories are that the FRW universe started, and then expanded, whereas the steady state universe "always has been". At the time these two theories were proposed, the church preferred FRW, with scientists preferring steady state.

The FRW universe model predicts some background radiation from the beginning event (i.e. the big bang), in the form of the cosmic microwave background (CMB). The CMB signature was predicted by Gamow and Alpher, and was observed by Penzias and Wilson. Therefore, providing evidence for the FRW universe.

The standard model of cosmology, today, uses the FRW model of the universe.

7.1 The FRW Metric

Schur's theorem (which we state without proof) states a globally isotropic n-dimensional manifold ($n > 2$) has a constant curvature k, and that the Riemann tensor has the form

$$R_{\mu\nu\alpha\beta} = k(g_{\mu\alpha}g_{\nu\beta} - g_{\mu\beta}g_{\nu\alpha}).$$

Following this, one can construct a isotropic metric,

$$ds^2 = dt^2 - a^2(t)d\sigma^2, \tag{7.1}$$

where $d\sigma^2$ is the line element for $3-$dim space, and $a(t)$ is the *scale factor*. We define the Hubble parameter, noting its present value,

$$H \equiv \frac{\dot{a}}{a}, \quad H_0 = 73\,\text{km/sec/Mpc};$$

where it is important to note that an overdot here denotes derivative with respect to coordinate time t. Furthermore, the metric actually looks like

$$ds^2 = dt^2 - a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right]. \tag{7.2}$$

Then, by a suitable coordinate transformation, the curvature constant k can take on one of 3 values,

$$k = \begin{cases} 1 & \text{closed} \\ 0 & \text{flat} \\ -1 & \text{open} \end{cases} \tag{7.3}$$

So, consider the values of k, to see how the actually correspond to the above "claimed" geometries.

Closed Space Consider setting $k = 1$, and the transformation

$$r = \sin\chi \quad \Rightarrow \quad dr = \cos\chi d\chi,$$

so that the metric looks like

$$ds^2 = dt^2 - a^2(t)\left[\frac{\cos^2\chi d\chi^2}{1-\sin^2\chi} + \sin^2\chi\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right],$$

which simplifies trivially down to

$$ds^2 = dt^2 - a^2(t)\left[d\chi^2 + \sin^2\chi\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right].$$

Now, consider taking a slice through θ. That is, set $\theta = \pi/2$, then one finds that

$$ds^2 = dt^2 - a^2(t)\left[d\chi^2 + \sin^2\chi d\phi^2\right],$$

where it is clear that the bracketed quantity is the line element of the 2-sphere. That is,

$$d\chi^2 + \sin^2\chi d\phi^2 \quad \Rightarrow \quad \text{sphere.}$$

Open Space Consider setting $k = -1$, and the coordinate transformation

$$r = \sinh\chi.$$

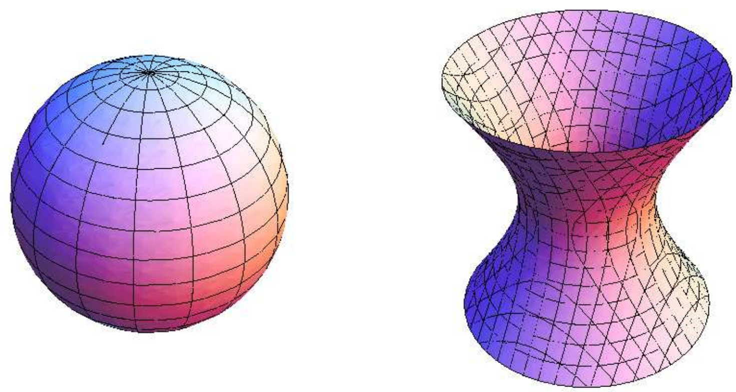
Then, under a completely analogous manner as before, we get the line element

$$ds^2 = dt^2 - a^2(t)\left[d\chi^2 + \sinh^2\chi d\phi^2\right],$$

where we now notice that

$$d\chi^2 + \sinh^2\chi d\phi^2 \quad \Rightarrow \quad \text{hyperboloid.}$$

That is, $k = -1$ corresponds to a geometry based upon the surface of a hyperboloid.



(a) Sphere - Closed (b) Hyperboloid - Open

Figure 7.1: A visualisation of closed and open geometries.

Flat Space Let us set $k = 0$, and the transformation

$$r = \chi,$$

then, we have the line element

$$ds^2 = dt^2 - a^2(t)\left[d\chi^2 + \chi^2\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right],$$

if we set $\theta = \pi/2$ again, then the square-bracketed quantity is just

$$d\chi^2 + \chi^2 d\phi^2.$$

This line element is just that of plane polars, which is flat. Hence, we see that $k = 0$ corresponds to flat space,

These correspondences of k with a particular geometry will become much clearer later on.

The standard way to write the FRW metric, in light of these coordinate transformations, is

$$\begin{aligned} ds^2 &= dt^2 - a^2(t)\left[d\chi^2 + \left\{\begin{array}{c}\sin^2\chi \\ \chi^2 \\ \sinh^2\chi\end{array}\right\}\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right], \\ k &= \left\{\begin{array}{c}+1 \\ 0 \\ -1\end{array}\right\}. \end{aligned} \qquad (7.4)$$

7.2 Geodesics & Christofell Symbols

We can compute the geodesics, and read off the Christofell symbols, from the effective Lagrangian formed from the FRW metric (7.2)

$$L_{\text{eff}} = \dot{t}^2 - a^2(t)\left[\frac{1}{1-kr^2}\dot{r}^2 + r^2\left(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2\right)\right],$$

where an overdot denotes derivative with respect to the affine parameter, λ, say. Now, one will need to use the following relation

$$\begin{aligned}\dot{a} &= \frac{da}{d\lambda}\\ &= \frac{\partial a}{\partial t}\frac{\partial t}{\partial\lambda}\\ &= a'\dot{t}, \quad a' \equiv \frac{da}{dt}.\end{aligned}$$

Upon careful computation, one finds the four geodesics:

$$\begin{aligned}\ddot{t} - \frac{aa'}{1-kr^2}\dot{r}^2 - aa'r^2\dot{\theta}^2 - aa'r^2\sin^2\theta\dot{\phi}^2 &= 0,\\ \ddot{r} + \frac{kr^2}{1-kr^2}\left(\frac{2a^2-1}{a^2}\right)\dot{r}^2 + 2\frac{a'}{a}\dot{t}\dot{r} - r(1-kr^2)\dot{\theta}^2 - r\sin^2\theta(1-kr^2)\dot{\phi}^2 &= 0,\\ \ddot{\theta} + 2\frac{a'}{a}\dot{t}\dot{\theta} + \frac{2}{r}\dot{r}\dot{\theta} - \sin\theta\cos\theta\dot{\phi}^2 &= 0,\\ \ddot{\phi} + 2\frac{a'}{a}\dot{t}\dot{\phi} + 2\frac{\sin^2\theta}{r}\dot{r}\dot{\phi} + 2\cot\theta\dot{\theta}\dot{\phi} &= 0.\end{aligned}$$

This allows us to read off the non-zero components of the Christofell symbols;

$$\begin{aligned}\Gamma^t{}_{rr} = -\frac{aa'}{1-kr^2}, \quad \Gamma^t{}_{\theta\theta} = -aa'r^2, \quad \Gamma^t{}_{\phi\phi} &= -aa'r^2\sin^2\theta,\\ \Gamma^r{}_{rr} = \frac{kr^2}{1-kr^2}\left(\frac{2a^2-1}{a^2}\right), \quad \Gamma^r{}_{tr} = \frac{a'}{a}, \quad \Gamma^r{}_{\theta\theta} &= -r(1-kr^2),\\ \Gamma^r{}_{\phi\phi} &= -r\sin^2\theta(1-kr^2),\\ \Gamma^\theta{}_{t\theta} = \frac{a'}{a}, \quad \Gamma^\theta{}_{r\theta} = \frac{1}{r}, \quad \Gamma^\theta{}_{\phi\phi} &= -\sin\theta\cos\theta,\\ \Gamma^\phi{}_{t\phi} = \frac{a'}{a}, \quad \Gamma^\phi{}_{r\phi} = \frac{\sin^2\theta}{r}, \quad \Gamma^\phi{}_{\theta\phi} &= \cot\theta.\end{aligned}$$

Notice that using the definition of the Hubble parameter, $H = a'/a$, we see that

$$\Gamma^r{}_{tr} = \Gamma^\theta{}_{t\theta} = \Gamma^\phi{}_{t\phi} = H.$$

This is the only section in which the derivative with respect to the affine parameter will be used; hence, an overdot from hereon denotes derivative with respect to coordinate time t.

7.3 Cosmology in the FRW Universe

We now wish to consider what happens to spacetime, in the FRW Universe. To do so, we shall need the Ricci tensor corresponding to the FRW metric, and some energy-momentum tensor.

So, following from the FRW metric, (7.2), one can compute the associated components of the Ricci tensor. Doing so, one finds

$$
\begin{aligned}
R_{00} &= -\frac{3\ddot{a}}{a}, && (7.5)\\
R_{0i} &= R_{i0} = 0, && (7.6)\\
R_{ij} &= -\left(\frac{2k}{a^2} + \frac{\ddot{a}}{a} + \frac{2\dot{a}^2}{a^2}\right) g_{ij}. && (7.7)
\end{aligned}
$$

The metric $g_{\mu\nu}$ is the FRW metric, which we note can be written as

$$g_{00} = 1, \quad g_{ij} = -a^2(t)\mathrm{diag}\left((1-kr^2)^{-1}, r^2, r^2\sin^2\theta\right).$$

Recall Einstein's equation, in the form

$$R_{\mu\nu} = 8\pi G\left(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T\right), \quad T \equiv g^{\mu\nu}T_{\mu\nu}.$$

We now use *Weyl's postulate* which is that our Universe is a perfect fluid. A perfect fluid is one for whom there is no heat conduction or viscosity.

Recall that the general energy-momentum tensor is given by

$$T_{\mu\nu} = (\rho + P)u_\mu u_\nu - Pg_{\mu\nu},$$

where P is the pressure of the fluid, and ρ the density. Hence, its trace is

$$T = (\rho + P)u^\mu u_\mu - Pg^\mu{}_\mu = \rho + P - 4P,$$

that is,

$$T = \rho - 3P.$$

Infact, this result can be obtained in a slightly easier way. Recall that in the comoving frame of the fluid, $u^\mu = (1,0,0,0)$, then the energy-momentum tensor is diagonal,

$$T_{\mu\nu} = \mathrm{diag}(\rho, -P, -P, -P).$$

Hence, its trace is just the sum of its components, $T = \rho - 3P$.

So, let us compute the bracketed bit of the Einstein equation,

$$
\begin{aligned}
T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T &= (\rho + P)u_\mu u_\nu - Pg_{\mu\nu} - \frac{1}{2}g_{\mu\nu}(\rho - 3P)\\
&= (\rho + P)u_\mu u_\nu - \frac{1}{2}g_{\mu\nu}(\rho - P).
\end{aligned}
$$

Hence, the Einstein equation reads

$$R_{\mu\nu} = 8\pi G \left((\rho + P) u_\mu u_\nu - \frac{1}{2} g_{\mu\nu} (\rho - P) \right).$$

Now, consider the comoving frame of the fluid, then we have that

$$T_{\mu\nu} = \text{diag}\left(\rho, -P g_{ij}\right), \quad T = \rho - 3P,$$

and thus that

$$T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T = \frac{1}{2} \text{diag}\left(\rho + 3P, g_{ij}(P - \rho)\right).$$

Hence,

$$T_{00} - \frac{1}{2} g_{00} T = \frac{1}{2}(\rho + 3P),$$

so, the 00-component of the Einstein equation, using (7.5) is

$$-\frac{3\ddot{a}}{a} = 8\pi G \frac{1}{2}(\rho + 3P),$$

trivially rearranging into

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P). \tag{7.8}$$

This is known as *Raychauhuri's equation.*

Similarly, suppose we took the ij-part of the Einstein equation, using (7.7), then

$$-\left(\frac{2k}{a^2} + \frac{\ddot{a}}{a} + \frac{2\dot{a}^2}{a^2} \right) g_{ij} = -8\pi G \frac{1}{2} g_{ij}(\rho - P),$$

from which we cancel out the metric g_{ij},

$$\frac{2k}{a^2} + \frac{\ddot{a}}{a} + \frac{2\dot{a}^2}{a^2} = 4\pi G(\rho - P).$$

Let us then insert Raychauhuri's equation for the middle term on the LHS,

$$\frac{2k}{a^2} - \frac{4\pi G}{3}(\rho + 3P) + \frac{2\dot{a}^2}{a^2} = 4\pi G(\rho - P).$$

This can then be rearranged easily enough into

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}. \tag{7.9}$$

This is known as the *Friedmann equation.* It is common to notate

$$\frac{\dot{a}}{a} \equiv H,$$

so that the Friedmann equation reads

$$H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}.$$

In deriving these two equations, we jumped around a bit between comoving frames. These equations describe the expansion of the universe, in the comoving frame of the fluid.

Let us see where the continuity equation

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

can get us. So, this is just

$$\partial_\nu T^\nu{}_\mu + \Gamma^\nu{}_{\nu\alpha} T^\alpha_{\ \mu} - \Gamma^\nu{}_{\mu\alpha} T^\alpha_{\ \nu} = 0,$$

where

$$T^\mu_{\ \nu} = \mathrm{diag}(\rho, -P, -P, -P).$$

Now, the Christofell symbols relevant are

$$\Gamma^t{}_{tt} = 0, \quad \Gamma^\theta{}_{t\theta} = \Gamma^\phi{}_{t\phi} = \Gamma^r{}_{tr} = H.$$

Now, let us take the $\mu = t$-component of the continuity equation,

$$\partial_\nu T^\nu{}_t + \Gamma^\nu{}_{\nu\alpha} T^\alpha_{\ t} - \Gamma^\nu{}_{t\alpha} T^\alpha_{\ \nu} = 0,$$

that is,

$$\partial_t T^t{}_t - \partial_i T^i{}_t + \Gamma^\nu{}_{\nu\alpha} T^\alpha_{\ t} - \Gamma^\nu{}_{t\alpha} T^\alpha_{\ \nu} = 0.$$

Now, the second term above is zero, as the energy-momentum tensor is diagonal. Hence, if we write that $T^\mu_{\ \nu} = \delta^\mu_\nu T^\mu{}_\nu$, then

$$\partial_t T^t{}_t + \Gamma^\nu{}_{\nu\alpha} \delta^\alpha_t T^\alpha_{\ t} - \Gamma^\nu{}_{t\alpha} \delta^\alpha_\nu T^\alpha_{\ \nu} = 0,$$

which is just

$$\partial_t T^t{}_t + \Gamma^\nu{}_{\nu t} T^t{}_t - \Gamma^\nu{}_{t\nu} T^\nu{}_\nu = 0.$$

Now, the only non-zero Christofell symbols of the form $\Gamma^\nu{}_{\nu t}$ are those $\Gamma^i{}_{it}$. Hence,

$$\partial_t T^t{}_t + \Gamma^i{}_{it} T^t{}_t - \Gamma^i{}_{ti} T^i{}_i = 0.$$

Therefore, with reference to the above Christofell symbols, we see that this is just

$$\partial_t \rho + 3H\rho + 3HP = 0,$$

which is

$$\dot{\rho} = -3H(\rho + P), \tag{7.10}$$

which is known as the *energy conservation equation*, or the fluid equation.

Hence, the three important equations we have derived, for a Universe in a perfect fluid:

- The Raychaudhuri equation:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P). \tag{7.11}$$

- The Friedmann equation:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}. \tag{7.12}$$

- The fluid equation:

$$\dot{\rho} = -3\frac{\dot{a}}{a}(\rho + P). \tag{7.13}$$

All three equations are dependent upon the others, so that in solving them, one must use all three. Infact, using any two, one can derive the third.

7.3.1 Species Evolution & Densities

The components to the fluid are called "species". That is, we could conceive that the fluid is composed of matter, radiation and possibly some other "stuff" (which we shall come to later).

Notice that we can write the fluid equation as

$$a\frac{\partial \rho}{\partial a} = -3(\rho + P).$$

Then, this can be solved, for the evolution of ρ as a function of scale factor a. We now consider three cases. We shall consider how the density of a particular species evolves, as a function of scale factor, if only that species exists in the Universe.

Matter Dominated FRW Universe Consider a Universe that is filled solely with matter. For matter, there is no associated pressure. Hence, $P_{\rm m} = 0$, and the fluid equation becomes

$$a\frac{\partial \rho_{\rm m}}{\partial a} = -3\rho_{\rm m},$$

integrating,

$$-3\int \frac{da}{a} = \int \frac{d\rho_{\rm m}}{\rho_{\rm m}} \quad \Rightarrow \quad -3\ln a = \ln \rho_{\rm m},$$

which is just

$$\rho_{\rm m} = \frac{\rho_{\rm m,0}}{a^3}, \tag{7.14}$$

whereby $\rho_{m,0}$ the (constant) initial density of matter.

Radiation Dominated FRW Universe Radiation has the equation of state

$$\rho_{\mathrm{r}} = 3P_{\mathrm{r}},$$

which may be derived from black-body radiation theory. Hence using this, the fluid equation reads

$$a\frac{\partial \rho_{\mathrm{r}}}{\partial a} = -4\rho_{\mathrm{r}},$$

integrating as before results in

$$\rho_{\mathrm{r}} = \frac{\rho_{\mathrm{r},0}}{a^4}. \tag{7.15}$$

Vacuum Dominated FRW Universe The equation of state for vacuum is

$$\rho_{\mathrm{V}} = -P_{\mathrm{V}},$$

so that the fluid equation reads

$$\dot{\rho}_{\mathrm{V}} = 0,$$

hence, we see that $\rho_{\mathrm{V}} = \text{const}$.

Critical Density Recall the Friedmann equation, but let us set $k = 0$ (i.e. flat),

$$H^2 = \frac{8\pi G}{3}\rho.$$

Then, let us define this ρ to be ρ_{crit}, so that

$$\rho_{\mathrm{crit}} = \frac{3H^2}{8\pi G}. \tag{7.16}$$

That is, ρ_{crit} is the density required to make the Universe flat. If we take the present value of the Hubble parameter to be

$$H_0 = 100h\,\mathrm{km\,s^{-1}Mpc^{-1}},$$

then the critical density should have value (if measured today),

$$\rho_{\mathrm{crit}} = 10.54h^2\mathrm{keV\,cm^{-3}}.$$

We use the notation that a subscript “0” denotes the present value of a quantity. In particular, we define

$$a_0 \equiv 1;$$

the present value of the scale factor is unity.

Normalised Energy Densities Let us suppose that there are four species present in the Universe: matter, radiation, vacuum and curvature. Let us now define

$$\Omega_{\rm m} \equiv \frac{\rho_{\rm m,0}}{\rho_{\rm crit}}, \quad \Omega_{\rm r} \equiv \frac{\rho_{\rm r,0}}{\rho_{\rm crit}}, \quad \Omega_{\rm V} \equiv \frac{\rho_{\rm V,0}}{\rho_{\rm crit}}, \quad \Omega_{\rm k} \equiv -\frac{k}{H_0^2 a_0^2}. \tag{7.17}$$

That is, the Ω_i are called the *normalised energy densities* of the species; they represent the current fraction of that species, in terms of the critical density. We impose the condition

$$\Omega_{\rm m} + \Omega_{\rm r} + \Omega_{\rm V} + \Omega_{\rm k} = 1,$$

as the Universe appears to be flat, by measurement. The matter species is composed of both baryonic and dark matter, radiation is composed of both photons and neutrinos. We tend to call the vacuum species the cosmological constant, so that $\Omega_{\rm V} = \Omega_\Lambda$. See Table (7.1) for the current values of various quantities.

Quantity	**Current Accepted Value**
$\Omega_{\rm m}$	0.24
$\Omega_{\rm b}$	0.04
$\Omega_{\rm DM}$	0.20
$\Omega_{\rm r}$	< 0.01
$\Omega_{\rm k}$	$\ll 0.05$
Ω_Λ	0.7

Table 7.1: Various quantities, as a fraction of $\rho_{\rm crit}$.

7.4 Age of the FRW Universe

Let us return to the Friedmann equation

$$H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2},$$

if we the divide through by H_0^2,

$$\frac{H^2}{H_0^2} = \frac{8\pi G}{3H_0^2}\rho - \frac{k}{a^2 H_0^2},$$

and the last expression on the RHS multiply/divide by a_0^2, to give

$$\frac{H^2}{H_0^2} = \frac{8\pi G}{3H_0^2}\rho - \frac{k}{a_0^2 H_0^2}\frac{a_0^2}{a^2}.$$

Now, we notice the presence of our definitions of $\rho_{\rm crit}$ and $\Omega_{\rm k}$, so that

$$\frac{H^2}{H_0^2} = \frac{\rho}{\rho_{\rm crit}} + \frac{\Omega_{\rm k}}{a^2},$$

after using that $a_0 = 1$. We now insert our derived evolutions of the various species ρ_i,

$$\begin{aligned} \frac{H^2}{H_0^2} &= \frac{1}{\rho_{\text{crit}}}\left(\frac{\rho_{\text{m},0}}{a^3} + \frac{\rho_{\text{r},0}}{a^4} + \rho_{V,0}\right) + \frac{\Omega_{\text{k}}}{a^2} \\ &= \frac{\Omega_{\text{m}}}{a^3} + \frac{\Omega_{\text{r}}}{a^4} + \Omega_{\text{V}} + \frac{\Omega_{\text{k}}}{a^2}. \end{aligned}$$

Hence, if we set $H = H_0$, and $a = a_0$, then we have

$$\Omega_{\text{m}} + \Omega_{\text{r}} + \Omega_{\text{V}} + \Omega_{\text{k}} = 1.$$

So, let us write our expression back in terms of the scale factor, so that

$$\left(\frac{\dot{a}}{a}\right)^2 = H_0^2\left[\frac{\Omega_{\text{m}}}{a^3} + \frac{\Omega_{\text{r}}}{a^4} + \Omega_{\text{V}} + \frac{\Omega_{\text{k}}}{a^2}\right],$$

or,

$$\frac{\dot{a}}{a} = H_0\left[\frac{\Omega_{\text{m}}}{a^3} + \frac{\Omega_{\text{r}}}{a^4} + \Omega_{\text{V}} + \frac{\Omega_{\text{k}}}{a^2}\right]^{1/2},$$

multiplying through by a, and pulling inside the square-root,

$$\dot{a} = H_0\left[\frac{\Omega_{\text{m}}}{a} + \frac{\Omega_{\text{r}}}{a^2} + \Omega_{\text{V}}a^2 + \Omega_{\text{k}}\right]^{1/2}. \tag{7.18}$$

Now, consider that

$$t_0 = \int_0^{t_0} dt = \int_0^{a_0=1} \frac{dt}{da}da = \int_0^1 \frac{da}{\dot{a}}.$$

Hence, we have that

$$t_0 = \frac{1}{H_0}\int_0^1 da\left[\frac{\Omega_{\text{m}}}{a} + \frac{\Omega_{\text{r}}}{a^2} + \Omega_{\text{V}}a^2 + \Omega_{\text{k}}\right]^{-1/2}. \tag{7.19}$$

Therefore, this expression will give us the age of the Universe.

7.4.1 Age of Matter Dominated Universe

So, let us assume that $\Omega_{\text{m}} = 1$ (all other species are zero). Hence, the present age of the Universe may be given by

$$\begin{aligned} t_0 &= \frac{1}{H_0}\int_0^1 da\, a^{1/2} \\ &= \frac{2}{3H_0}. \end{aligned}$$

Also notice that in the matter dominated universe, (7.18) looks quite simple,

$$\dot{a} = H_0 \sqrt{\Omega_{\rm m}} a^{-1/2},$$

which is easily solved to give

$$a \propto t^{2/3}. \tag{7.20}$$

That is, if the Universe is matter dominated, then the scale factor evolves in time as $t^{2/3}$.

Another curious result, is that for a vacuum dominated Universe, $\dot{a} \propto a$, which implies that

$$a \propto e^t,$$

that is, in a vacuum dominated Universe, the scale factor grows exponentially with time.

7.4.2 Age of Matter & Curvature Dominated Universe

Here, we have a mixture of two species, such that

$$\Omega_{\rm r} = \Omega_{\rm V} = 0, \quad \Omega_{\rm m} + \Omega_{\rm k} = 1.$$

Let us introduce a rescaling of time, known as *conformal time*, whereby

$$a d\eta = dt.$$

Hence,

$$\eta = \int_0^t \frac{dt}{a} = \int_0^a \frac{da}{a\dot{a}}.$$

Notice that in writing this, we have that $\eta = \eta(a)$. We should then be able to invert it, so that $a = a(\eta)$. Notice that if we use conformal time, the FRW metric (7.2) can be written in the form

$$ds^2 = a^2(t)\left[d\eta^2 - \frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right] \sim a^2(t) g_{\mu\nu} dx^\mu dx^\nu.$$

That is, we have a conformal transformation of the metric. This is why we call η conformal time.

Now, (7.18) in our model is

$$\dot{a} = H_0 \left[\frac{\Omega_{\rm m}}{a} + \Omega_{\rm k}\right]^{1/2},$$

hence,

$$a\dot{a} = H_0 \left[\Omega_{\rm m} a + \Omega_{\rm k} a^2\right]^{1/2}.$$

Therefore, using this,

$$\eta = \frac{1}{H_0}\int_0^a \frac{da}{\sqrt{\Omega_{\rm m} a + \Omega_{\rm k} a^2}}.$$

To integrate this, we complete the square, giving

$$\eta = \frac{1}{H_0}\int_0^a da \left\{\Omega_{\rm k}\left[\left(a + \frac{\Omega_{\rm m}}{2\Omega_{\rm k}}\right)^2 - \frac{\Omega_{\rm m}^2}{4\Omega_{\rm k}^2}\right]\right\}^{-1/2}.$$

If we then define

$$x \equiv \frac{2\Omega_{\rm k}}{\Omega_{\rm m}} a + 1,$$

then we see that we can write

$$\begin{aligned}\eta &= \frac{1}{H_0}\int_a^x dx \frac{2\Omega_{\rm k}}{\Omega_{\rm m}}\left\{\Omega_{\rm k}\left[\frac{\Omega_{\rm m}^2}{4\Omega_{\rm k}^2}(x^2 - 1)\right]\right\}^{-1/2} \\ &= \frac{1}{H_0\sqrt{\Omega_{\rm k}}}\int_1^x \frac{dx}{\sqrt{x^2 - 1}},\end{aligned}$$

where we look up the value of the integral,

$$\int_1^x \frac{dx}{\sqrt{x^2-1}} = \cosh^{-1} x.$$

Hence,

$$\eta = \frac{1}{H_0\sqrt{\Omega_{\rm k}}}\cosh^{-1} x.$$

Therefore,

$$x = \cosh\left(\eta H_0 \sqrt{\Omega_{\rm k}}\right).$$

Hence,

$$a(\eta) = \frac{\Omega_{\rm m}}{2\Omega_{\rm k}}\left[\cosh\left(\eta H_0\sqrt{\Omega_{\rm k}}\right) - 1\right].$$

Writing $\Omega_{\rm k} = 1 - \Omega_{\rm m}$, then this reads

$$a(\eta) = \frac{\Omega_{\rm m}}{2(1-\Omega_{\rm m})}\left[\cosh\left(\eta H_0\sqrt{1-\Omega_{\rm m}}\right) - 1\right], \quad \Omega_{\rm k} > 0. \tag{7.21}$$

Clearly, this only holds for $\Omega_{\rm k} > 0$. If $\Omega_{\rm k} < 0$, then the cosh becomes a cosine, and we have

$$a(\eta) = \frac{\Omega_{\rm m}}{2(\Omega_{\rm m}-1)}\left[1 - \cos\left(\eta H_0\sqrt{\Omega_{\rm m}-1}\right)\right], \quad \Omega_{\rm k} < 0. \tag{7.22}$$

With reference to Figure (7.2), we see the two different types of Universes. It is clear from the analytic forms of the evolution of scale factor with conformal time, $a(\eta)$, that $\Omega_{\rm k} > 0$

Figure 7.2: A visualisation of closed and open universes. Closed has $\Omega_k < 0$, and open $\Omega_k > 0$. The former is just a sinusoidal-oscillation, the latter an exponential expansion.

corresponds to an exponential increase in scale factor (7.21), and $\Omega_k < 0$ an oscillatory scale factor (7.22). Also, from the definition of Ω_k,

$$\Omega_k = -\frac{k}{H_0^2 a_0^2},$$

we see that

$$\Omega_k > 0 \quad \Rightarrow \quad k < 0 \quad \Rightarrow \quad \text{open}, \tag{7.23}$$

$$\Omega_k < 0 \quad \Rightarrow \quad k > 0 \quad \Rightarrow \quad \text{closed}. \tag{7.24}$$

which are in agreement of our previous statements of open and closed Universes. So,

An oscillatory Universe will have a definite (conformal) time when it ends, when $a(\eta)$ hits the axis again,

$$\cos\left(\eta_{\text{tot}} H_0 \sqrt{\Omega_m - 1}\right) = 1 \quad \Rightarrow \quad \eta_{\text{tot}} = \frac{2\pi}{\sqrt{\Omega_m - 1} H_0}.$$

Hence, the actual total time is given by

$$t_{\text{tot}} = \int_0^{\eta_{\text{tot}}} d\eta\, a(\eta),$$

which easily evaluates to

$$t_{\text{tot}} = \frac{\pi \Omega_m}{(\Omega_m - 1)^{3/2} H_0}.$$

Therefore, we have an expression for the total possible age of the Universe, if the Universe has a closed geometry. Hence, a small non-zero k is sufficient to control the future "fate" of the Universe. That is, the Universe will either end up exponentially growing (the "heat death"), or will crunch back on itself (the "big crunch").

7.5 Light in the FRW Universe

Consider the FRW metric, where we shall ignore all angular terms;

$$ds^2 = dt^2 - a^2(t)\frac{dr^2}{1-kr^2}.$$

Now, assuming flatness, for light (i.e. null geodesics, $ds^2 = 0$), we have that the metric reduces to

$$dt = a(t)dr.$$

Therefore, consider

$$R = \int_0^R dr = \int_{t_o}^{t_e} \frac{dt}{a(t)}.$$

That is, the distance between two points that have photons sent between them. We have that t_e is the time of emission of the photon, and t_o the time of observation. Now, we shall assume that this distance is unchanged, for pulses sent slightly after this first set, so that

$$R = \int_{t_o+\delta t_o}^{t_e+\delta t_e} \frac{dt}{a(t)}.$$

Therefore, we have that

$$\int_{t_o+\delta t_o}^{t_e+\delta t_e} \frac{dt}{a(t)} = \int_{t_o}^{t_e} \frac{dt}{a(t)}.$$

Now, the only non-zero contribution to this (via a general calculus mid-point theorem) is

$$\frac{\delta t_o}{a(t_o)} - \frac{\delta t_e}{a(t_e)} = 0,$$

which easily rearranges to

$$\frac{\delta t_o}{\delta t_e} = \frac{a(t_o)}{a(t_e)}.$$

Now, we can express the LHS as a ratio of frequencies (by units), so that

$$\frac{\nu_e}{\nu_o} = \frac{a(t_o)}{a(t_e)} \equiv 1 + z.$$

Hence, we arrive at a standard relation in cosmology,

$$\frac{\nu_e}{\nu_o} = 1 + z. \tag{7.25}$$

This is always > 0. Therefore, we see that the ratio of received frequency and "sent" frequency (i.e. the frequency that the light was, when it was sent by the object) is dependent upon the *redshift* z that the light was emitted. This quantity z is just the ratio of the scale factors when the light was received, to when it was emitted. Hence, we see that the further away something is, the frequency we see light emitted by it drops. That is, the wavelength increases. Hence, this is called the *cosmological redshift effect*. This is a different effect from gravitational redshift, because gravitational redshift occurred due to different distances from a gravitating body.

$$\begin{aligned} \text{Expansion of Universe} \quad &\Rightarrow \quad \text{Cosmological redshift,} \\ \text{Different distances up gravitational potential} \quad &\Rightarrow \quad \text{Gravitational redshift.} \end{aligned}$$

To get a handle on the numbers invloved, consider that the most distant quasar is at $z \approx 6.6$, and that recombination is at $z \approx 10^3$.

Notice that we can write

$$z = \frac{\nu_e - \nu_o}{\nu_o} = \frac{a_o - a_e}{a_e}.$$

Also, recall that (non-relativistic) redshift is related to the velocity of the object,

$$z = \frac{v}{c} = \frac{\delta a}{a}.$$

Hence, notice that we may compute

$$\frac{\delta a}{a} = \frac{\delta a/\delta t}{a}\delta t = H\frac{R}{c}.$$

Therefore,

$$v = HR.$$

This is Hubble's law, as derived from first principles from the FRW metric.

7.6 Flatness Problem

Now, there are problems with the FRW Universe.

Recall that the fraction, today, of curvatures contribution to the total density of the Universe is $\Omega_{\mathrm{k},0} < 10\%$. Also recall that we defined

$$\Omega_{\mathrm{k}}(t) \equiv \frac{k}{H(t)a^2(t)},$$

where t is the time at which we are measuring. Hence, let us compute,

$$\frac{\Omega_{\mathrm{k}}(t_0)}{\tilde{\Omega}_{\mathrm{k}}(tr)} = \frac{k/H_0^2 a_0^2}{k/H_r^2 a_r^2};$$

the ratio of the curvature contributions today and in the radiation dominated epoch. This easily reduces to

$$\frac{\Omega_{\mathrm{k}}(t_0)}{\tilde{\Omega}_{\mathrm{k}}(t_r)} = \frac{\dot{a}_r^2}{H_0^2 a_0^2}.$$

Now, recalling that the scale factor, in the radiation dominated epoch, depends upon time as

$$a_r = a_0 \left(\frac{t_r}{t_0}\right)^{1/2} \quad \Rightarrow \quad \dot{a}_r = \frac{a_0}{2t_0}\left(\frac{t_r}{t_0}\right)^{-1/2},$$

Also, recall that the Hubble parameter, in the radiation dominated epoch, goes as

$$H_0 = \frac{1}{2t_0}.$$

Hence,

$$\dot{a}_r = H_0 a_0 \left(\frac{t_r}{t_0}\right)^{-1/2}.$$

And therefore,

$$\frac{\Omega_{\rm k}(t_0)}{\tilde{\Omega}_{\rm k}(t_r)} = \frac{t_0}{t_r}.$$

Putting some typical numbers in, one sees that

$$\frac{\Omega_{\rm k}(t_0)}{\tilde{\Omega}_k(t_r)} \approx \frac{10^{17}\text{secs}}{10^{-43}\text{secs}} = 10^{60}.$$

Hence,

$$\Omega_{\rm k}(t_0) = 10^{60}\tilde{\Omega}_{\rm k}(t_r).$$

That is, the value of $\Omega_{\rm k}$ is 10^{60} times what it was in the radiation epoch! This requires a very small (so called "fine-tuning") curvature in the early epoch, so that the Universe could be 10^{60} times more curved now, than it was.

This fine-tuning required is called the *flatness problem*.

7.6.1 Inflation

One way to "solve" the flatness problem, is to introduce the concept of inflation. If we allow an epoch before the radiation domination, that was vacuum dominated (recall that $a_V(t) = a_i e^{Ht}$). In this case, we can compute that

$$\frac{\Omega_{\rm k}(t_r)}{\tilde{\Omega}_{\rm k}(t_i)} = \frac{a_i^2}{a_r^2},$$

after assuming that $H_i \approx H_r$. This gives

$$\frac{\Omega_{\rm k}(t_r)}{\tilde{\Omega}_{\rm k}(t_i)} = e^{-2H(t_r - t_i)},$$

a number we require to be less than 10^{-60}. Therefore, we require

$$\mathcal{N}_e \equiv 2H(t_r - t_i) > 60.$$

That is, we require the number of e-folds to be about 60, in order for us to observe the flatness that we do today.

Basically, this idea of inflation gives a mechanism by which the Universe is able to stretch and flatten out, very quickly. Infact, inflation also aids in explaining the observed homogeneity of the Universe.

8 The General Theory of Relativity: Discussion

We have now come to a place whereby all the mathematical groundwork has been laid, for a "wordy" discussion about the general theory of relativity.

Before general relativity (or at least a few hundred years before Einstein, as general relativity went through a few people before Einstein, in various forms), gravity was some force that was present between two bodies having mass. As this was so, things that don't have mass don't interact with gravity. This means that things like photons are not affected by gravity, and that photons are not capable of generating a gravitational field. Also, the structure of spacetime was that space is flat, and time is just something to be moved through, at a constant rate; where the rate is the same for all observers.

General relativity somewhat starts off by letting space and time mix: spacetime. The whole collection of bits of spacetime is then what we call a manifold; further to this, allowing a meaning to the term "distance" in a manifold, we introduce a metric. We call a manifold (collection of points) that has a metric, a Riemannian manifold. We "used" to think of spacetime as being flat (Pythagoras' theorem for distances between two points). A flat spacetime is described by a metric with constant components; taking the derivative of any one of them, with respect to any coordinate, is zero. Now, general relativity introduces the idea that a metric has components that depend on position. This means that in order to find the distance between two points, you not only have to know where the points are, but where you are relative to the origin of the coordinate system. This is in contrast with only needing to know the relative positions of the two points.

When one computes the derivative of something, one is computing the rate of change of something in a particular direction. Now, when one did this in a flat spacetime, the derivative of the metric didn't do anything: its derivative was zero. In a position dependent metric, this is no longer true. One finds that there is an extra bit, added onto the differential of something, that is proportional to the derivative of the metric. That this extra bit exists, is directly due to the metric being position dependent. Therefore, various combinations of this metric (in the form of differential with respect to various coordinates), will give us a handle on the geometry of the manifold. A slightly curious thing is that a manifold does not require a higher dimension in which to curve. Usually, when one imagines a ball (as an example), one can see that the surface of the ball is curved round, through three dimensions, but the surface of the ball itself is two dimensional. Manifolds do not require this extra dimension (to those within the manifold itself) in which to curve.

Mathematically, we carry around the "extra bits" of the differential in the Christofell symbols; and the "various combinations" of the differentials of the metric in the Riemann tensor.

Now, something that lives in a manifold, and moves in that manifold, will move along some sort of curve; which is fairly obvious. Now, the motion of something, with respect to a stationary observer, can be determined. In a flat spacetime, something will move along lines that are determined by Newton's equations of motion. In a curved spacetime (i.e. a spacetime that doesn't have all zero components of its Riemann tensor), the curves that things move along are changed; and the amount that they are changed by is proportional to the Christofell symbols. These curves are geodesics. A geodesic, in flat spacetime, with no external forces (such as a rocket boost, or magnetic fields), is a straight line. A corresponding geodesic, in curved spacetime is curved. This curvature of the movement free "thing" is due to the curvature of the spacetime.

So, this far in our discussion, we have seen that if a manifold is curved, then things don't tend to move in straight lines within the manifold. That is, the geodesics are curved lines. A way to imagine this, is to envisage a cube threaded with a 3D grid; beads move along the gridlines, but the gridlines are not straight. This is only an analogy, as the real geodesics are 4D. Then, we must consider what it is that does the curving. What thing, in a manifold, causes it to be curved?

The proposal of Einstein is that all forms of matter and energy (even though they are essentially the same) curve spacetime. The proposal equates the distribution of "stuff" (i.e. the things that do the curving, things that have mass & energy) with the geometry of the spacetime. That is, the distribution of mass-energy with combinations of the metric. This means that the more energy you put in a given place, the more the spacetime is curved (and hence the more curvy geodesics get). The distribution of mass-energy is carried around in the energy-momentum tensor, and the geometry in the Einstein tensor.

This curvature of spacetime, due to the distribution of mass-energy is the "main idea" of general relativity.

Some of the consequences of this general theory include the "ability" of massless things, which have energy, to interact with gravity. This is because the massless things move through the spacetime, and gravity is just the curvature of spacetime. This allows the geodesic of a photon to be curved. Notice that this is in contrast with the previous flat spacetime we started off discussing. This gives the so-called "light deflection" effect. Another consequence is that things at a different distance from the centre of a body doing the curving (the so-called gravitating mass), experience difference rates of passage through time. This is because the position-dependent metric has different values at different positions (obviously). An example of this, is that if we synchronize two clocks, on the surface of the earth, then take one up from the surface of the earth, and leave one on the surface, they will tell different times when brought back together.