1. Geodesics

We have seen that if we have a vector V^i at one point x then we van construct a vector at a neighbouring point $x+dx$ such that the new vector is the 'same' as the vector at the original point. This is called 'parallel transporting' the vector from x to $x+dx$. The change in the components of the vector was

$$\delta V^i = -\Gamma^i_{jk} V^j \delta x^k \tag{1.1}$$

Now suppose that we wanted to ask if a given curve was a 'straight line' on the manifold. Let us start with one point P on the curve, proceed in a given direction along the curve, and denote the distance along the curve from P by the symbol s. The vector

$$V^i(x) = \frac{dx^i}{ds}(x) \tag{1.2}$$

is obviously tangent to the curve at the point x. We also see that

$$V^i V_i = \frac{dx^i dx_i}{ds^2} = 1 \tag{1.3}$$

so V^i is a *unit* vector. We 'parallel transport' this vector to another point on the curve at $x+dx$, and ask: is the transported vector equal to tangent vector that was already defined at the point $x+dx$ through (1.2)? Clearly, if it is not, then the curve cannot be called 'straight'. If it is, then we can imagine that we have looked at a small patch around the point x, put coordinates on it that look locally Cartesian, and then transprted the vector to $x+dx$ as if we were keeping the components of the vector constant in Cartesian coordinates. Thus to have a 'straight line' we need

$$V^i(x+dx) = V^i(x) + \delta V^i = V^i(x) - \Gamma^i_{jk}(x) V^j(x) \delta x^k \tag{1.4}$$

We can rewrite this as

$$\frac{V^i(x+dx) - V^i(x)}{ds} = -\Gamma^i_{jk}(x) V^i(x) \frac{\delta x^k}{ds} \tag{1.5}$$

or

$$\frac{dV^i}{ds} + \Gamma^i_{jk} V^j V^k = 0 \tag{1.6}$$

If the tangent V defined for a curve satisfies this equation then that curve is a 'straight line' on the manifold and is called a *geodesic*.

There is one possibility that we did not discuss above: what happens if the vector V^i transported to $x + dx$ happens to be parallel to the tangent at $x + dx$, but is not of unit length after the transport? Will we call the curve a geodesic in that case?

In fact we can easily see that this will *not* happen. Upon parallel transport, a vector may change its components but will not change its length:

$$\begin{aligned}\delta(V^iV^jg_{ij}) &= 2(\delta V^i)V^jg_{ij} + V^iV^j\delta g_{ij} = -2\Gamma^i_{mn}V^m\delta x^nV^jg_{ij} + V^iV^jg_{ij,n}\delta x^n \\ &= -g^{ik}[g_{km,n} + g_{kn,m} - g_{mn,k}]g_{ij}V^mV^j\delta x^n + V^iV^jg_{ij,n}\delta x^n \\ &= -[g_{jm,n} + g_{jn,m} - g_{mn,j}]V^mV^j\delta x^n + V^iV^jg_{ij,n}\delta x^n \\ &= 0\end{aligned} \tag{1.7}$$

where in the last but one step we observed that two of the terms in the box bracket cancelled since taken together they were antisymmetric in m, j while the quantity V^mV^j was symmetric.

In fact, more generally, we see that the inner product between two vectors will also not change under parallel transport:

$$\begin{aligned}\delta(V^iW^jg_{ij}) &= (\delta V^i)W^jg_{ij} + V^i(\delta W^j)g_{ij} + V^iW^j\delta g_{ij} \\ &= -\Gamma^i_{mn}V^m\delta x^nW^jg_{ij} - \Gamma^j_{mn}W^m\delta x^nV^ig_{ij} + V^iW^jg_{ij,n}\delta x^n \\ &= -\frac{1}{2}g^{ik}[g_{km,n} + g_{kn,m} - g_{mn,k}]g_{ij}V^mW^j\delta x^n \\ &\quad - \frac{1}{2}g^{jk}[g_{km,n} + g_{kn,m} - g_{mn,k}]g_{ij}W^mV^i\delta x^n + V^iW^jg_{ij,n}\delta x^n \\ &= -\frac{1}{2}[g_{jm,n} + g_{jn,m} - g_{mn,j}](V^mW^j + V^jW^m)\delta x^n + V^iW^jg_{ij,n}\delta x^n \\ &= 0\end{aligned} \tag{1.8}$$

2. Covariant derivative

Recall that the motivation for defining a connection was that we should be able to compare vectors at two neighbouring points. Suppose we are given a *vector field* - that is, a vector $V^i(x)$ at each point x. Then we can compute the derivative of this vector field. Thus we take two points, with coordinates x^i and $x^i + \delta x^i$. The vector at x has components $V^i(x)$. The vector at $x + dx$ has components

$$V^i(x + \delta x) = V^i(x) + \frac{\partial V^i}{\partial x^k}\delta x^k \tag{2.1}$$

To compare this latter vector with the vector at x we transport the vector at x also to the point $x + dx$. This transported vector has components

$$V^i(x) - \Gamma^i_{jk}(x)V^j(x)\delta x^k \tag{2.2}$$

Thus the 'true' change in the vector between x and $x + dx$ is

$$\frac{\partial V^i}{\partial x^k}\delta x^k + \Gamma^i_{jk}(x)V^j(x)\delta x^k = [\frac{\partial V^i}{\partial x^j} + \Gamma^i_{jk}(x)V^j(x)]\delta x^k \tag{2.3}$$

Thus we may define a 'true' derivative, called the *covariant derivative* by

$$V^i_{;k} \equiv \frac{\partial V^i(x)}{\partial x^j} + \Gamma^i_{jk}(x)V^j(x) \tag{2.4}$$

the first term takes into account the fact that the components of V are changing, while the second removes the part of this change that is simply due to the fact that the coordinates themselves are changing.

How do we find the covariant derivative of a *covariant* vector W_i? One way is to redrive the connection for a covariant vector just the way we did for a contravariant vector, by starting with flat space where transport is trivial, and changing to curvilinear coordinates. But we can instead use the fact that we found above in (1.8), namely that the dot product between two contrvariant vectors is unchanged by parallel transport:

$$0 = \delta(V^i W^j g_{ij}) = \delta(V^i W_i) = (\delta V^i)W_i + V^i(\delta W_i) = -\Gamma^i_{kl}V^k\delta x^l W_i + V^i(\delta W_i) \tag{2.5}$$

Since this relation must hold true for all possible V^i, we can set to zero the coefficient of V^k for each k in the above equation:

$$\delta W_k = \Gamma^i_{kl}W_i\delta x^l \tag{2.6}$$

Thus the connection is related to that involved in transporting the contravariant vector, but note that the overall sign is different, and that the contraction of indices is different. The notation of up and down indices for contravariant and covariant vectors helps us to keep track of how the indices should be contracted here. Recall however that though V^i and W_i are vectors, Γ^i_{jk} does not transform like a tensor, so the index structure is a pnemonic and not a tensor index contraction.

We can now define a covariant derivative for covariant vectors, by following the same chain of reasoning that we followed for contravariant vectors. We find

$$W_{i;k} \equiv \frac{\partial W_i}{\partial x^k} - \Gamma^l_{ik} W_l \tag{2.7}$$

Let us now compute the covariant derivative of the metric tensor:

$$\begin{aligned} g_{ij;k} &= g_{ij,k} - \Gamma^l_{ik} g_{lj} - \Gamma^l_{jk} g_{il} \\ &= g_{ij,k} - \frac{1}{2} g^{lm}[g_{mi,k} + g_{mk,i} - g_{ik,m}] g_{lj} - \frac{1}{2} g^{lm}[g_{mj,k} + g_{mk,j} - g_{jk,m}] g_{li} \\ &= g_{ij,k} - \frac{1}{2}[g_{ji,k} + g_{jk,i} - g_{ik,j}] - \frac{1}{2}[g_{ij,k} + g_{ik,j} - g_{jk,i}] \\ &= 0 \end{aligned} \tag{2.8}$$

This fact is very significant, since we find

$$V_{i;k} = (V^i g_{ij})_{;k} = V^i_{;k} g_{ij} + V^i g_{ij;k} = V^i_{;k} g_{ij} \tag{2.9}$$

Thus we can covariantly differentiate a contravariant vector, and then lower the contravariant index, or lower the index first and then compute the covariant derivative - in either case we will get the same answer. thus the operation of raising and lowering indices commutes with the operation of taking a covariant derivative. This fact will substantially reduce the kinds of geometric objects that we can make by starting with a given tensor and taking covariant derivatives.

3. 4-velocity

In Newtonian mechanics velocity is a vector with three space components: $\frac{dx^i}{dt}$, $i = 1, 2, 3$. But with special relativity we needed to treat time on the same footing as space. In this case what shall we use as the denominator in the expression for the velocity? Along the world line of a particle, if we take to infinitesimally separated points, then a coordinate independent quantity is the 'proper distance' ds between the points:

$$ds^2 = dt^2 - dx_1^2 - dx_2^2 - dx_3^2 \tag{3.1}$$

The components of 4-velocity (called 4-velocity since it has four components) are defined as

$$U^0 = \frac{dt}{ds}, \quad U^1 = \frac{dx^1}{ds} \quad U^2 = \frac{dx^2}{ds} \quad U^3 = \frac{dx^3}{ds} \tag{3.2}$$

If the particle is moving slowly, then $dx/dt << 1$, and

$$ds = dt[1 - (\frac{d\vec{x}}{dt})^2]^{1/2} = dt + O(\vec{v}^2) \tag{3.3}$$

Thus in the limit where we expect Newtonian physics to be valid, we get

$$U^0 \approx 1, \quad U^1 \approx v^1, \quad U^2 \approx v^2, \quad U^3 \approx v^3 \tag{3.4}$$

and the four velocity U has the same data as the usual velocity $\vec{v}$. In curvilinear coordinates on flat space and more generally in curved spacetime, we will similarily have

$$ds^2 = g_{ij}(x)dx^i dx^j \tag{3.5}$$

where at each point x the metric g_{ij} represents one timelike and three spacelike directions. The 4-velocity is

$$U^i = \frac{dx^i}{ds} \tag{3.6}$$

We will often call the 4-velocity just the velocity.

We observe that

$$U^i U_i = \frac{dt^2 - (d\vec{x})^2}{ds^2} = 1 \tag{3.7}$$

Thus U is a timelike 4-vector. We can see that it is a vector of contravariant type, since it is defined through the separation between two points on spacetime. Thus U has its index written 'up'.

4. Raising and lowering indices

Suppose we have a contravariant vector V^i at some point x of the manifold, and a covariant vector W_i at the same point. Let us form the object

$$f = V^i W_i \tag{4.1}$$

If we transform to new coordinates we see that

$$f' = V'^i W'_i = \frac{\partial \xi^i}{\partial x^m} \frac{\partial x^n}{\xi^i} V^m W_n = \delta^n_m V^m W_n = V^m W_m = f \tag{4.2}$$

thus f transforms as a function (which we also called a scalar); it has no free indices, and we see that when indices are 'contracted' in the above fashion then we do not need to worry about the transformations that act on such indices - the transformations cancel each other out. If V, W were vector fields, we would obtain a function $f(x)$ over spacetime b carrying out the contraction (4.1) at each point.

We had seen that the metric tensor had two covariant indices. Let us start with a contravariant vector V^i and form the contraction

$$g_{ij}V^j \tag{4.3}$$

It is easy to check that this transforms to a new coordinate frame as

$$(g'_{ij}V'^j) = \frac{\partial x^m}{\partial \xi^i}\frac{\partial x^n}{\xi^j}\frac{\partial \xi^j}{\partial x^p}g_{mn}V^p = \frac{\partial x^m}{\partial \xi^i}\delta^n_p g_{mn}V^p = \frac{\partial x^m}{\partial \xi^i}(g_{mp}V^p) \tag{4.4}$$

Thus again only the free index participates in the transformation law, but the transformation law of the quantity (4.4) is that of a *covariant* vector; the two lower indices of the metric and one upper index of V have resulted in an object with one lower index. We denote this object by the same symbol V, but with a lower index:

$$V_i \equiv g_{ij}V^j \tag{4.5}$$

The fact that we use the symbol V still is justified because if we assume that we are given the metric, then the information in V^i is contained in the information in V_i and vice versa: we can get

$$g^{ij}V_j = g^{ij}g_{jk}V^k = \delta^i_k V^k = V^i \tag{4.6}$$

which converts a covariant vector V_i to a contravariant. vector. Thus the metric and the inverse metric can be used to raise and lower indices on vectors.

5. The Newtonian limit

With this formalism of geodesics on curved manifolds we should be able to reproduce in some limit the law of Gravitation in Newton's theory. The key difference between general relativity and Newtonian mechanics is of course the fact that the latter is not relativistic; thus we should look for a limit where all particle velocities are low compared to the speed of light, or with $c = 1$, small compared to unity. Thus $\vec{v}$ will be the small parameter in our approximation.

Consider the geodescic equation (1.6). Note that for $i = 1, 2, 3$

$$\frac{dU^i}{ds} = \frac{d}{ds}\frac{dx^i}{ds} \approx \frac{d^2x^i}{dt^2} \tag{5.1}$$

Further,

$$U^0 \approx 1, \quad U^i << 1, \quad i = 1, 2, 3 \tag{5.2}$$

Thus we get

$$\frac{d^2x^i}{dt^2} \approx -\Gamma^i_{00} \tag{5.3}$$

Now we assume that the spacetime is nearly flat, so that the metric is

$$g_{ij} = \eta_{ij} + h_{ij}, \quad h_{ij} << 1 \tag{5.4}$$

We also assume that the metric is static, so that no component depends on the coordinate $x^0 = t$. Then

$$\Gamma^i_{00} = -\frac{1}{2}g_{00,i} = -\frac{1}{2}h_{00,i}, \quad i = 1, 2, 3 \tag{5.5}$$

and the geodesic equation becomes

$$\frac{d^2x^i}{dt^2} \approx \frac{1}{2}h_{00,i}, \quad i = 1, 2, 3 \tag{5.6}$$

We should compare this to the equation expected from Newton's theory

$$\frac{d^2x^i}{dt^2} = -\phi_{,i} \tag{5.7}$$

where ϕ is the gravitational potential. Thus we see that we need to identify

$$h_{00} \approx -2\phi \tag{5.8}$$

For a point mass source M we have

$$\phi = -\frac{GM}{r} \tag{5.9}$$

Thus the metric must have

$$g_{00} \approx (1 - \frac{2GM}{r}) \tag{5.10}$$

Indeed, the exact metric for a point mass source is the Schwarzschild metric

$$ds^2 = (1 - \frac{2GM}{r})dt^2 - \frac{dr^2}{1 - \frac{2GM}{r}} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{5.11}$$

Let us compute the period of circular orbits in this metric. Let the orbit have radius r_0. Then

$$\frac{d^2r}{ds^2} = \frac{d}{ds}\frac{dr}{ds} = \frac{dU^r}{ds} = 0 \tag{5.12}$$

But by the geodesic equation,

$$\frac{dU^r}{ds} = -\Gamma^r_{ij}U^iU^j \tag{5.13}$$

But the only components of the 4-velocity that are nonvanising are $U^\theta = \frac{d\theta}{ds}, U^0 = \frac{dt}{ds}$, since only the coordinates θ, t change along the orbit. Thus we find

$$0 = \Gamma^r_{\theta\theta}U^\theta U^\theta + \Gamma^r_{00}U^0U^0 + 2\Gamma^r_{r\theta}U^\theta U^0 \tag{5.14}$$

We have

$$\Gamma^r_{\theta 0} = 0, \quad \Gamma^r_{\theta\theta} = -(1 - \frac{2GM}{r})r, \quad \Gamma^r_{00} = (1 - \frac{2GM}{r})\frac{GM}{r^2} \tag{5.15}$$

From the geodesic equation we find

$$\frac{U^\theta}{U^0} = \frac{d\theta}{dt} = (\frac{GM}{r^3})^{1/2} \tag{5.16}$$

Thus the period of an orbit of radius r_0 is

$$T = 2\pi(\frac{r_0^3}{GM})^{1/2} \tag{5.17}$$

In Newtonian gravity we would find the angular velocity by equating the centripetal acceleration to the gravitational force

$$(\frac{d\theta}{dt})^2 r = \frac{GM}{r^2} \tag{5.18}$$

which also gives the time period (5.17). While for large r where the metric is close to the flat one and the particle velocity in circular orbits is slow we would expect the time periods to be approximately the same between the general relativistic treatment and the Newtonian treatment, we see here that by a coincidence the expressions are in fact identical for all r. However as we will see now this does not mean that we can have such circular orbits for all r in the relativistic theory, unlike the case for the Newtonian theory.

6. Gravitational redshift

Suppose we have a person who stays at a fixed radius $r > 2M$ in the spacetime with metric (5.11). Let this person emit some periodic signal - for example it might be a light wave with a frequency ν, or he may just spray bullets from a gun at a fixed interval. We assume that in his own frame these periodic events have a time separation ΔT. Let the light wave or the sequence of bullets reaches a person standing fixed at some other point, say at $r' > 2M$. What will be the percieved interval between the periodic events for the person at r'?

We assume that when the person at r says that the events have a separation ΔT then he means that the proper time along his world line between two successive events is ΔT. Thus

$$ds = (1 - \frac{2GM}{r})^{1/2}\Delta t = \Delta T \tag{6.1}$$

Since the spacetime is static, the intervals between the events at r' will be fixed in the sense that they will occur at the same separations Δt. But this gives for the proper time along the worldline of the observer at r'

$$ds' = (1 - \frac{2GM}{r'})^{1/2}\Delta t = (\frac{1 - \frac{2GM}{r'}}{1 - \frac{2GM}{r}})^{1/2}\Delta T \tag{6.2}$$

Thus the frequency of a light wave would appear to be *lower* that the one emitted to an observer who sits at a radius $r' > r$. This effect is called *gravitational redshift*. It is somewhat different from the doppler shift that we encounter in studying sound waves r light waves. If we move at a velocity v compared to the source emitting a sound wave, then the fractional change in the frequency of the sound wave will be appriciable if we move with a speed that is of the order of the sound speed. but this same speed would give a very small fractional change for the frequency of light waves: to get a significant chage for light we would have to move with a speed comparale to the speed of *light*. But the gravitational redshift that we discussed changes the frequency of any motion by the same proportionality factor, once we fix the positions r and r'. Thus gravitational redshift is a property of the curved spacetime itself, and is a central feature of general relativity.

7. Curvature

Let us take the second derivative of the vector field that we had above. Then we get an expression

$$[V^a{}_{;c}]_{;d} \equiv V^a{}_{;cd} \tag{7.1}$$

If we had been computing ordinary partial derivatives of the components V^a with respect to the coordinates ξ^b, ξ^c then these partial derivatives would commute

$$\frac{\partial^2 V^a}{\partial \xi^c \partial \xi^d} = \frac{\partial^2 V^a}{\partial \xi^d \partial \xi^c} \tag{7.2}$$

But the covariant derivatives in fact do not commute in general. We have

$$V^a{}_{;c} = V^a{}_{,c} + \Gamma^a_{cf} V^f \tag{7.3}$$

$V^a{}_{;c}$ is a tensor with one contravariant index and one covariant index. So we have

$$\begin{aligned} V^a{}_{;cd} &= [V^a{}_{,c} + \Gamma^a_{cf} V^f]_{,d} + \Gamma^a_{df} V^f{}_{;c} - \Gamma^f_{cd} V^a{}_{;f} \\ &= V^a{}_{,cd} + \Gamma^a_{cf,d} V^f + \Gamma^a_{cf} V^f{}_{,d} + \Gamma^a_{df} V^f{}_{;c} - \Gamma^f_{cd} V^a{}_{;f} \end{aligned} \tag{7.4}$$

As written, the above expression has both ordinary partial derivatives of V as well as some covariant derivatives of V. We could convert them all to ordinary partial derivatives plus some connection terms, but what we want to do is to compare the above expression with the two covariant derivatives taken in the reverse order. Thus we have

$$V^a{}_{;dc} = V^a{}_{,dc} + \Gamma^a_{df,c} V^f + \Gamma^a_{df} V^f{}_{,c} + \Gamma^a_{cf} V^f{}_{;d} - \Gamma^f_{dc} V^a{}_{;f} \tag{7.5}$$

If we take the difference of the above two relations, we will find by (7.2) that the first terms on the RHS will cancel. The last terms on the RHS cancel as well, since $\Gamma^a_{cd} = \Gamma^a_{dc}$. There are two terms left in each expression with first order derivatives of V, but of these one term is a partial derivative while the other is a covariant derivative. Expanding the covariant derivative involved here, we get

$$\begin{aligned} V^a{}_{;cd} - V^a{}_{;dc} &= \Gamma^a_{cf,d} V^f - \Gamma^a_{df,c} V^f + \Gamma^a_{dg} \Gamma^g_{fc} V^f - \Gamma^a_{cg} \Gamma^g_{fd} V^f \\ &= [\Gamma^a_{cf,d} - \Gamma^a_{df,c} + \Gamma^a_{dg} \Gamma^g_{fc} - \Gamma^a_{cg} \Gamma^g_{fd}] V^f \\ &\equiv -R^a{}_{fcd} V^f \end{aligned} \tag{7.6}$$

Note the remarkable fact that even though the second order covariant partial derivatives of V do not commute, the difference between the derivatives taken in the two different orders is an expression that involves only V and not any of its derivatives. The effect of the noncommutation has been summarised in the *Riemann curvature tensor*

$$R^a{}_{bcd} = \Gamma^a_{bd,c} - \Gamma^a_{bc,d} + \Gamma^a_{cf}\Gamma^f_{bd} - \Gamma^a_{df}\Gamma^f_{bc} \tag{7.7}$$

It is not evident from the expression above that this quantity should be a tensor, since it involves the connection, which is not itself a tensor, and further there are ordinary partial derivatives of this connection. But let us consider geometrically what R^a_{bcd} signifies. In the above calculation we had assumed that V was a vector field, i.e. a vector at each point of the space. But as we saw the final expression defining the curvature through (7.6) did not involve any derivatives of V, so we did not really need to know how the vector field changed from point to point, which suggests that we should be able to define $R^a{}_{bcd}$ using only a vector that is assumed to exist at one point x (where we wish to define the tensor). Thus take a vector V^a at x, and parallel transport it to a point $x + \delta x$. then the components of the transported vector will be

$$V^a - \Gamma^a_{fc}(x)V^f\delta x^c \tag{7.8}$$

where we have written explicitly the fact that the connection is evaluated at x. Now we transport this vector further to a point $x + \delta x + \tilde{\delta}x$. Then the components of the vector will be

$$\begin{aligned} &[V^a - \Gamma^a_{fc}(x)V^f\delta x^c] - \Gamma^a_{gd}(x+\delta x)[V^g - \Gamma^g_{fc}(x)V^f\delta x^c]\tilde{\delta}x^d \\ &\quad = [V^a - \Gamma^a_{fc}(x)V^f\delta x^c] - [\Gamma^a_{gd}(x) + \Gamma^a_{gd,k}\delta x^k][V^g - \Gamma^g_{fc}(x)V^f\delta x^c]\tilde{\delta}x^d \end{aligned} \tag{7.9}$$

Now suppose we had done the transports in the other order - first moved the vector to $x + \tilde{\delta}x$ and then to $x + \tilde{\delta}x + \delta x$ which is the same point as the one reached before. then we would get

$$\begin{aligned} &[V^a - \Gamma^a_{fd}(x)V^f\tilde{\delta}x^d] - \Gamma^a_{gc}(x+\tilde{\delta}x)[V^g - \Gamma^g_{fd}(x)V^f\tilde{\delta}x^d]\delta x^c \\ &\quad = [V^a - \Gamma^a_{fd}(x)V^f\tilde{\delta}x^d] - [\Gamma^a_{gc}(x) + \Gamma^a_{gc,k}\tilde{\delta}x^k][V^g - \Gamma^g_{fd}(x)V^f\tilde{\delta}x^d]\delta x^c \end{aligned} \tag{7.10}$$

Subtracting (7.10) from (7.10) we get for the difference of the change in V between the two paths

$$[-\Gamma^a_{bd,c} + \Gamma^a_{bc,d} + \Gamma^a_{df}\Gamma^r_{bc} - \Gamma^a_{cf}\Gamma^r_{bd}]V^b\delta x^c\delta\tilde{x}^d = -R^a{}_{bcd}V^b\delta x^c\delta\tilde{x}^d \tag{7.11}$$

So we see that curvature describes the difference in parallel transport along different paths.