CHAPTER 5

Space-times and geodesics

The 'happiest thought of Einstein's life' was a brilliant insight which led to the 'geometrization' of gravity. The physics you observe in a spaceship, with its engines switched off, far from any source of gravity, is the same as the physics you observe if you are freely falling in a lift under the influence of the earth's gravity. This suggests that particles which are acted upon by no force other than gravity should be regarded as freely falling, or essentially inertial. To incorporate the idea mathematically, we need an extension of the formalism of SR which is 'generally covariant'. Where the formalism of SR (4-vectors) was invariant under the Lorentz group (and even the Poincaré group), these transformations are essentially linear. In GR we need a set-up which allows for the 'physics' to be equivalent under arbitrary coordinate transformations.

In this chapter we shall introduce curved 4-dimensional space-times and study geodesics. The hypothesis is that freely falling particles follow such geodesics. We shall also show that at any event in space-time, 'local inertial coordinates can be introduced'. These are coordinates in which the coefficients are the same as the Minkowski metric up to second order.

In the next chapter we shall introduce curvature. This is the quantity (it is, unfortunately, a 4-index tensor) which allows us to tell whether space-time is actually curved or not. One remark about this here. The general equivalence principle described above: that the physics of the freely falling lift and the inertial spaceship should be the same—only applies on small scales of space and time, or 'locally' as we say in the trade. There is a difference in behaviour of freely falling particles in deep space as opposed to near the earth. According to Newton, the gravitational field near the earth is not uniform: in fact it is given by the famous inverse-square law. This means that if if Alice and Bob are freely falling near the earth, and if Alice is nearer to the Earth's surface than Bob, then her acceleration will be greater. In other words the *relative acceleration* of Alice, as measured by Bob, will be non-zero. In deep space, far from any stars, planets, or other gravitating objects, the relative acceleration of two observers will be zero. This is extremely subtle and is tied in with the subtle notion of curvature. More on that story later.

5.1. Curved space-time

DEFINITION 5.1.1. A *curved space-time* (or simply *space-time* $(\mathscr{M}, g)$ is a 4-dimensional manifold, with a given lorentzian metric g.

REMARK 5.1.2. For those of you who haven't yet mastered §4.7, you can think that $\mathscr{M}$ is just fancy notation for $\mathbb{R}^4$, or perhaps an open subset of it. But we are not allowed to use the vector-space structure of $\mathbb{R}^4$ in developing the formalism: everything we do from now on has to use the metric g and must work in any choice of local coordinates, or be 'generally covariant' to use somewhat old-fashioned terminology.

Recall that relative to a choice of (local) coordinates $x^a = (x^0, x^1, x^2, x^3)$, the metric has the form

$$ds^2 = g_{ab} dx^a dx^b. \tag{5.1.1}$$

Here the summation convention is in force and the components g_{ab} of g with respect to the coordinates x^a form a 4×4 symmetric matrix whose entries are smooth functions. To say the metric is lorentzian is to say that at any point x, the matrix $(g_{ab}(x))$ is invertible and has signature $(+, -, -, -)$.

REMARK 5.1.3. A closely related notion is that of a *riemannian manifold* of dimension n. This is a pair $(\mathscr{M}, g)$, where $\mathscr{M}$ is a manifold of dimension n and g is a riemannian metric, i.e.

$$\mathrm{d}s^2 = g_{ij}\mathrm{d}x^i\mathrm{d}x^j$$

where the components g_{ij} form a symmetric $n \times n$ positive-definite matrix. We shall use the case $n = 2$ to illustrate the theory.

REMARK 5.1.4. The 'right' way to think of a lorentzian or riemannian metric is an assignment $p \mapsto g(p)$, where $g(p)$ is a symmetric bilinear form on the tangent space $T_p\mathscr{M}$. For each p, $g(p)$ is positive-definite in the riemannian case and has signature $(+, -, \ldots, -)$ in the lorentzian case. Of course the assignment is supposed to be smooth as p varies in $\mathscr{M}$. To make this smoothness precise, we look at the components $g_{ab}(x)$ of the metric in some smooth local coordinate system, and insist that these be smooth.

We shall generally use indices $a, b, c, \ldots$ for space-time indices running from 0 to 3; and indices $i, j, k, \ldots$ for indices running from 1 to n (where n is usually 2 or 4 in examples).

Our metric is a symmetric $(0, 2)$-tensor. It has an inverse, with components denoted g^{ab}, which form a $(2, 0)$-tensor. The statement that g^{ab} is inverse to g_{ab} is written

$$g_{ab}g^{bc} = \delta_a^c. \tag{5.1.2}$$

EXAMPLE 5.1.5. The standard metric euclidean metric on $\mathbb{R}^2$ is

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2$$

The same metric in polar coordinates can be computed as follows: from

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

then

$$\mathrm{d}x = \cos\theta\,\mathrm{d}r - r\sin\theta\,\mathrm{d}\theta, \mathrm{d}y = \sin\theta\,\mathrm{d}r + r\cos\theta\,\mathrm{d}\theta.$$

Squaring and adding,

$$\begin{aligned} \mathrm{d}x^2 + \mathrm{d}y^2 &= (\cos\theta\,\mathrm{d}r - r\sin\theta\,\mathrm{d}\theta)^2 + (\sin\theta\,\mathrm{d}r + r\sin\theta\,\mathrm{d}\theta)^2 \\ &= \mathrm{d}r^2 + r^2\,\mathrm{d}\theta^2. \end{aligned}$$

Note that the coefficient of the cross-term $\mathrm{d}r\,\mathrm{d}\theta + \mathrm{d}\theta\,\mathrm{d}r$ is zero.

EXAMPLE 5.1.6. If

$$\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\theta^2,$$

then writing $x^1 = r, x^2 = \theta$,

$$g_{11} = 1, g_{12} = g_{21} = 0, g_{22} = r^2.$$

Hence

$$g^{11} = 1, g^{12} = g^{21} = 0, g^{22} = \frac{1}{r^2}.$$

This is because

$$g_{ij} = \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix} \text{ and } g^{ij} = \begin{bmatrix} 1 & 0 \\ 0 & r^{-2} \end{bmatrix}$$

are inverse to each other.

EXAMPLE 5.1.7. The 2D metric

$$\mathrm{d}s^2 = \mathrm{d}u\mathrm{d}v + \mathrm{d}v\mathrm{d}u$$

has components

$$g_{00} = g_{11} = 0, g_{01} = g_{10} = 1$$

if $u = x^0$ and $v = x^1$. In this case the components of g^{ij} are the same

$$g^{00} = g^{11} = 0, g^{01} = g^{10} = 1.$$

REMARK 5.1.8. The change of variables $t = u + v$, $x = u - v$ brings the previous metric into inertial form (diagonal, with one $+$ and one $-$ on the diagonal).

EXAMPLE 5.1.9. **Spherical polars** 3-dimensional spherical polar coordinates are given by

$$x = r\sin\theta\cos\varphi,\ y = r\sin\theta\sin\varphi,\ x = r\cos\theta. \tag{5.1.3}$$

Show that the euclidean metric

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2$$

takes the form

$$\mathrm{d}s^2 = \mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2). \tag{5.1.4}$$

5.1.1. Lorentzian metric at a point.

PROPOSITION 5.1.10. *Suppose that*

$$\mathrm{d}s^2 = g_{ab}\mathrm{d}x^a\mathrm{d}x^b$$

is a 4-dimensional lorentzian metric defined near the point p with coordinates $x^a(p) = 0$ (a coordinate system centred at p). Then there are coordinates y^a, also centred at p, such that

$$\widetilde{g}_{ab}(p) = \eta_{ab}.$$

Thus, very near any given point p of $\mathscr{M}$, the geometry of $\mathscr{M}$ is approximately the same as Minkowski space.

REMARK 5.1.11. We shall do better than this: in §5.5, we shall see that we can choose coordinates so that

$$\widetilde{g}_{ab}(\widetilde{x}) = \eta_{ab} + O(|\widetilde{x}|^2)$$

for small $\widetilde{x}$. Such a choice of coordinates will be called *local inertial coordinates*. They give the best approximating Minkowski space at the point with coordinates $\widetilde{x} = 0$.

PROOF. It is sufficient to make a change of coordinates which is linear:

$$x^a = J^a_b y^b.$$

Then by the transformation law for tensors of type $(0, 2)$,

$$\widetilde{g}_{ab} = g_{pq}J^p_a J^q_b$$

In matrix form this is

$$\widetilde{g} = J^t g J.$$

By the basic theorem about diagonalization of symmetric bilinear forms, K can be chosen to make $\widetilde{g}(0)$ diagonal, with diagonal entries ± 1. The signs are determined by the signature of g. If the latter is Lorentzian, this yields the Minkowski metric $\eta = \mathrm{diag}(1, -1, -1, -1)$. □

5.1.2. Timelike/spacelike/null.

DEFINITION 5.1.12. A tangent vector $X = X^a\partial_a$ is called *timelike* at p in $\mathscr{M}$ if

$$g(X, X)(p) = g_{ab}X^aX^b|_p > 0,$$

null if

$$g(X, X)(p) = g_{ab}X^aX^b|_p = 0$$

and *spacelike* if

$$g(X, X)(p) = g_{ab}X^aX^b|_p < 0.$$

The set of null vectors at a point p form a cone (in $T_p\mathscr{M}$) which are supposed to be tangent to photon worldlines through p.

5.2. Events and worldlines

As in the case of SR, the points of our 4-dimensional space $\mathscr{M}$ are called *events*, localized in time and space. Particles, observers and photons are described by worldlines, i.e. parameterized curves in $\mathscr{M}$. The following definition captures the idea that massive particles cannot travel faster than light.

DEFINITION 5.2.1. Let $(\mathscr{M}, g)$ be a 4-dimensional space-time. A paramaterized curve $\gamma(t)$ is called *timelike* if its tangent vector is timelike for every value of the parameter t,

$$g(\dot\gamma, \dot\gamma) > 0 \tag{5.2.1}$$

Similarly, a parameterized curve $\gamma(t)$ is called *null* if its tangent vector is *null* for every value of the parameter t,

$$g(\dot\gamma, \dot\gamma) = 0. \tag{5.2.2}$$

As in SR, where $(\mathscr{M}, g)$ reduces to $(\mathbb{M}, \eta)$, massive particles follow timelike curves in $\mathscr{M}$: this corresponds to the speed of the particle being everywhere less than the speed of light. Similarly photons follow null curves. Also as in Minkowski space, these curves are called *worldlines*.

HYPOTHESIS 5.2.2. A timelike curve γ is parameterized by proper time τ if

$$g\left(\frac{\mathrm{d}\gamma}{\mathrm{d}\tau}, \frac{\mathrm{d}\gamma}{\mathrm{d}\tau}\right) = 1. \tag{5.2.3}$$

Then τ is the time that would be shown on a clock with worldline $\gamma(\tau)$. More precisely, if Alice's worldline is $\gamma(\tau)$ and $p = \gamma(\tau_1)$ and $q = \gamma(\tau_2)$ are two events on her worldline, then her clock will show an elapsed time $\tau_2 - \tau_1$ between these two events if (5.2.3) holds.

Given any timelike curve, $\widetilde\gamma(u)$, there is always a reparameterization

$$\gamma(\tau) = \widetilde\gamma(u(\tau))$$

of the curve by proper time, cf. Proposition 1.3.4.

5.3. Geodesics

In special relativity, inertial observers were taken to travel at constant speed along straight lines in Minkowski space $\mathbb{M}$. One of the definitions of straight line is a curve which minimises the *energy* (cf. §1.4), amongst all those with fixed endpoints.

Using the space-time metric g on $\mathscr{M}$, we can do the same thing.

DEFINITION 5.3.1. The energy of a curve $\gamma : [t_0, t_1] \longrightarrow \mathscr{M}$ is defined to be

$$\mathscr{E}[\gamma] = \frac{1}{2}\int_{t_0}^{t_1} g(\dot\gamma(t), \dot\gamma(t))\,\mathrm{d}t. \tag{5.3.1}$$

A *geodesic* with end-points p and q is a curve which extremizes the energy among all curves with $\gamma(t_0) = p$, $\gamma(t_1) = q$.

HYPOTHESIS 5.3.2. In GR, freely falling particles (and free photons) follow geodesics, timelike for particles and null for photons. Here 'freely falling' means 'acted upon by no force except gravity'.

DEFINITION 5.3.3. Let $x^a = (x^0, x^1, x^2, x^3)$ be a given coordinate system, such that the metric coefficients are g_{ab},

$$\mathrm{d}s^2 = g_{ab}\mathrm{d}x^a\mathrm{d}x^b.$$

The *Christoffel symbols* of g_{ab} are defined by the formula

$$\Gamma^c_{ab} = \frac{1}{2}g^{cs}\left(\partial_a g_{bs} + \partial_b g_{as} - \partial_s g_{ab}\right). \tag{5.3.2}$$

REMARK 5.3.4. Note the symmetry of Γ in its two lower indices,

$$\Gamma^c_{ab} = \Gamma^c_{ba}. \tag{5.3.3}$$

THEOREM 5.3.5. *Let $\gamma(t)$ be a curve in $\mathscr{M}$ and suppose that in some coordinate system, it is given by $t \mapsto x^c(t)$. Then the Euler–Lagrange equations for $\mathscr{E}(\gamma)$ are equivalent to the equations*

$$\ddot{x}^c + \Gamma^c_{ab}\dot{x}^a\dot{x}^b = 0. \tag{5.3.4}$$

where Γ^c_{ab} are as in the previous definition. If $x^c(t)$ is a geodesic, then $g_{cd}\dot{x}^c\dot{x}^d$ is constant.

REMARK 5.3.6. The system of equations (5.3.4) are called the *geodesic equations*. They are frequently a more convenient way of getting at the Γs, as we shall see in examples.

PROOF. This is a calculus of variations problem with Lagrangian $L(x,\dot{x}) = \frac{1}{2}g(x)[\dot{x},\dot{x}] = \frac{1}{2}g_{ab}(x)\dot{x}^a\dot{x}^b$. We have

$$\frac{\partial L}{\partial \dot{x}^s} = g_{sb}\dot{x}^b, \quad \frac{\partial L}{\partial x^s} = \frac{1}{2}\partial_s g_{ab}\dot{x}^a\dot{x}^b \tag{5.3.5}$$

so

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}^s} = g_{sb}\ddot{x}^b + \frac{\partial g_{as}}{\partial x^b}\dot{x}^a\dot{x}^b. \tag{5.3.6}$$

Thus the Euler–Lagrange equations are

$$g_{sb}\ddot{x}^b + \frac{\partial g_{as}}{\partial x^b}\dot{x}^a\dot{x}^b - \frac{1}{2}\partial_s g_{ab}\dot{x}^a\dot{x}^b = 0 \tag{5.3.7}$$

Now multiply through by g^{cs} (and sum over s):

$$\ddot{x}^c + g^{cs}\left(\frac{\partial g_{as}}{\partial x^b} - \frac{1}{2}\partial_s g_{ab}\right)\dot{x}^a\dot{x}^b = 0 \tag{5.3.8}$$

Now

$$\frac{\partial g_{as}}{\partial x^b}\dot{x}^a\dot{x}^b = \frac{1}{2}\left(\frac{\partial g_{as}}{\partial x^b} + \frac{\partial g_{bs}}{\partial x^a}\right)\dot{x}^a\dot{x}^b \tag{5.3.9}$$

and substituting this into (5.3.8), taking into account the definition of the Γs, yields (5.3.4).

For the last part, the Lagrangian is homogeneous of degree 2 in the velocities, and so is conserved along a solution curve (Proposition 1.4.4: the potential energy part is zero in the case at hand.) □

Computing the Γs is the first step in computing the *curvature* of the metric, and computing the geodesic equations is needed to understand particle (and photon) motion in GR. We therefore give some worked examples. In each case, the Γs are read off the geodesic equations (5.3.4) rather than from the formula (5.3.2).

EXAMPLE 5.3.7. Minkowski space. This is

$$\mathrm{d}s^2 = \eta_{ab}\mathrm{d}x^a\mathrm{d}x^b.$$

The Lagrangian is

$$L = \frac{1}{2}\eta_{ab}\dot{x}^a\dot{x}^b$$

Then

$$\frac{\partial L}{\partial \dot{x}^a} = \eta_{ab}\dot{x}^b, \frac{\partial L}{\partial x^a} = 0. \tag{5.3.10}$$

Then the geodesic equations are

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial L}{\partial \dot{x}^a} = \eta_{ab}\ddot{x}^b = 0. \tag{5.3.11}$$

Thus the geodesic equations are $\ddot{x}^b = 0$, from which we read that all Γs are zero. The geodesics have the form

$$x^a(\tau) = Y^a + U^a\tau.$$

Of course, we already knew this.

EXAMPLE 5.3.8. (The 2D hyperbolic metric) This is

$$ds^2 = \frac{dx^2 + dy^2}{x^2} \tag{5.3.12}$$

The Lagrangian $L = (\dot{x}^2 + \dot{y}^2)/2x^2$. The Euler–Lagrange equations are

$$\frac{d}{d\tau}\frac{\dot{x}}{x^2} + \frac{1}{x^3}(\dot{x}^2 + \dot{y}^2) = 0; \quad \frac{d}{d\tau}\frac{\dot{y}}{x^2} = 0.$$

Rearranging, these become

$$\ddot{x} - \frac{\dot{x}^2}{x} + \frac{\dot{y}^2}{x} = 0, \quad \ddot{y} - 2\frac{\dot{x}\dot{y}}{x} = 0. \tag{5.3.13}$$

With $x^1 = x, x^2 = y$, these are supposed to be identical to the geodesic equations

$$\ddot{x}^1 + \Gamma^1_{ij}\dot{x}^i\dot{x}^j = 0, \quad \ddot{x}^2 + \Gamma^2_{ij}\dot{x}^i\dot{x}^j = 0. \tag{5.3.14}$$

Hence

$$\Gamma^1_{11} = -\frac{1}{x}, \Gamma^1_{22} = \frac{1}{x}, \Gamma^2_{12} = \Gamma^2_{21} = -\frac{1}{x}, \text{ others } = 0.$$

NB the factor of 2 in Γ^2_{12} coming from the symmetry $\Gamma^i_{12} = \Gamma^i_{21}$.

EXAMPLE 5.3.9. **The geodesics in 2D hyperbolic space** Rather than tackle the second-order equations (5.3.13) directly, it is better to work with conserved quantities. The first is the length of the velocity vector. If we assume τ is arc-length (i.e. $L = 1$), then

$$\dot{x}^2 + \dot{y}^2 = x^2. \tag{5.3.15}$$

The Lagrangian is also independent of y, and so $\partial L/\partial \dot{y}$ is constant along geodesics. So in addition to (5.3.15), we also have

$$\frac{\dot{y}}{x^2} = C \text{ (constant)}. \tag{5.3.16}$$

If $C = 0$, then $\dot{y} = 0$, so $y = y_0$ (constant). Subsituting this into (5.3.15) we get $dx/x = \pm d\tau$. Picking the $+$ sign, and integrating, we get $x = x_0 e^\tau$. Thus one family of geodesics are half-lines parallel to the x-axis, given by

$$(x(\tau), y(\tau)) = (x_0 e^\tau, y_0).$$

Note the non-standard parameterization of these straight lines. In particular, the 'boundary point' $(0, y_0)$ is infinitely far away in τ: we require $\tau \to -\infty$ for $(x, y) \to (0, y_0)$. This is not outrageous because the metric (5.3.12) blows up at these points with $x = 0$.

Now we have to deal with the case $C \neq 0$. For this, divide (5.3.15) through by $\dot{y}^2$,

$$\left(\frac{dx}{dy}\right)^2 + 1 = \frac{x^2}{\dot{y}^2} = \frac{1}{C^2 x^2}, \tag{5.3.17}$$

where we've used (5.3.16) and

$$\frac{\dot{x}}{\dot{y}} = \frac{dx}{dy}. \tag{5.3.18}$$

Rearranging (5.3.17), we obtain

$$\frac{Cx\,dx}{\sqrt{1 - C^2x^2}} = \pm dy$$

and this can be integrated to get

$$\pm\sqrt{1 - C^2x^2} = C(y - y_0)$$

for some constant of integration y_0. This can be rearranged to give

$$x^2 + (y - y_0)^2 = C^{-2}, x > 0 \tag{5.3.19}$$

i.e. a semicircle with diameter along the y axis. It is pleasing that as $C \to 0$, the radius C^{-1} tends to infinity and these semicircles approach the straight half-lines that we saw previously.

You can verify that

$$(x(\tau), y(\tau)) = (C^{-1}\operatorname{sech}\tau, y_0 + C^{-1}\tanh u) \tag{5.3.20}$$

is a parameterizaiton by arclength. (To obtain this, you eliminate x from (5.3.16) using (5.3.19), to get

$$\frac{\mathrm{d}y}{\mathrm{d}\tau} = C(C^{-2} - (y - y_0)^2)$$

and integrate this up to get y as a function of τ.)

Note that in this example it is relatively easy to obtain the geodesics as an implicit relation between x and y (5.3.19), whereas finding x and y as a function of τ is rather more involved. This is quite typical of the simple examples we shall see in this course.

Example 5.3.10. Minkowski space in polar coordinates. In spherical polars, the Minkowski metric takes the form

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \mathrm{d}r^2 - r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2). \tag{5.3.21}$$

The Lagrangian is

$$L = \frac{1}{2}\left(\dot{t}^2 - \dot{r}^2 - r^2\dot{\theta}^2 - r^2\sin^2\theta\dot{\varphi}^2\right) \tag{5.3.22}$$

Then

$$\frac{\partial L}{\partial \dot{t}} = \dot{t}, \quad \frac{\partial L}{\partial \dot{r}} = -\dot{r}, \quad \frac{\partial L}{\partial \dot{\theta}} = -r^2\dot{\theta}, \quad \frac{\partial L}{\partial \dot{\varphi}} = -r^2\sin^2\theta\dot{\varphi} \tag{5.3.23}$$

$$\frac{\partial L}{\partial t} = 0, \quad \frac{\partial L}{\partial r} = -r(\dot{\theta}^2 - \sin^2\theta\dot{\varphi}^2), \quad \frac{\partial L}{\partial \theta} = -r^2\sin\theta\cos\theta\dot{\varphi}^2, \quad \frac{\partial L}{\partial \varphi} = 0. \tag{5.3.24}$$

Thus the geodesic equations are

$$\ddot{t} = 0, \tag{5.3.25}$$

$$\ddot{r} - r(\dot{\theta}^2 + \sin^2\theta\dot{\varphi}^2) = 0, \tag{5.3.26}$$

$$\ddot{\theta} + \frac{2}{r}\dot{r}\dot{\theta} - \sin\theta\cos\theta\dot{\varphi}^2 = 0, \tag{5.3.27}$$

$$\ddot{\varphi} + \frac{2}{r}\dot{r}\dot{\varphi} + 2\cot\theta\dot{\theta}\dot{\varphi} = 0. \tag{5.3.28}$$

With $x^0 = t$, $x^1 = r$, $x^2 = \theta$, $x^3 = \varphi$, we read off that

$$\Gamma^1_{22} = -r, \qquad \Gamma^1_{33} = -r\sin^2\theta, \tag{5.3.29}$$

$$\Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{r}, \qquad \Gamma^2_{33} = -\sin\theta\cos\theta, \tag{5.3.30}$$

$$\Gamma^3_{13} = \Gamma^3_{31} = \frac{1}{r}, \qquad \Gamma^3_{23} = \Gamma^3_{32} = \cot\theta, \tag{5.3.31}$$

while all others are zero. Note again that there are factors of 2 between the Γs and the coefficients in the geodesic equations for the Γ^a_{bc} with $b \neq c$.

We will not go into finding the geodesics here as the moves you have to make were already described in Example 1.4.5. And you are urged to review that example now! Of course this is a bit different because of having the variable t as an additional coordinate in the problem. But since $\dot{t}$ is a constant, λ, say, the constancy of L means that

$$\dot{r}^2 + r^2(\dot{\theta}^2 + \sin^2\theta\dot{\varphi}^2) = 2L - 2\lambda^2$$

is a constant. Also $J = r^2\sin^2\theta\dot{\varphi}$ is a constant and we can restrict to equatorial curves where $\theta = \pi/2$ identically. Then we find that

$$\dot{r}^2 + r^2\dot{\varphi}^2 = 2L - 2\lambda^2 \text{ and } J = r^2\dot{\varphi} \tag{5.3.32}$$

are both constants just as in (1.4.18); the solution follows as there.

Remark 5.3.11. One important remark from this is that *the Christoffels can be non-zero even if the metric is the Minkowski metric.* Written in funny coordinates (here spherical polars) many of the Γs are non-zero, although the intrinsic geometry of the metric is unchanged.

5.4. A first look at the covariant derivative

If X and Y are two vector fields on $\mathscr{M}$, then we denote by $\nabla_X Y$ the vector field with components

$$(\nabla_X Y)^c = X^a \partial_a Y^c + \Gamma^c_{ab} X^a Y^b. \tag{5.4.1}$$

This is called the *covariant derivative* of Y with respect to X. In the previous sentence I slipped in the idea that $\nabla_X Y$ is a vector field. This is not obvious: neither of the two terms on the RHS in (5.4.1) transforms as a vector field under a change of coordinates. However, the combination does transform as a vector field:

THEOREM 5.4.1. *For any vector fields X and Y, $\nabla_X Y$ is again a vector field on $\mathscr{M}$.*

A proof will be given in the next chapter. You can also look in §5.1 of Woodhouse, GR, for a more computational proof.

5.4.1. Covariant derivative of a vector along a curve. Let $\gamma(\tau)$ be a curve in $\mathscr{M}$ with tangent vector

$$\dot{\gamma} = X. \tag{5.4.2}$$

If Y is another vector field defined along (or near) γ, then

$$\nabla_X Y = \nabla_{\dot{\gamma}}(Y) \tag{5.4.3}$$

is the *covariant derivative of Y along γ.*

On a curved space-time this is the best available replacement for the derivative of a vector along a curve in Minkowski space. In particular, if we take $Y = X = \dot{\gamma}$ we define the *acceleration* of the curve to be the vector field along γ

$$A = \nabla_{\dot{\gamma}}\dot{\gamma} \tag{5.4.4}$$

Then the LHS of the geodesic equation

$$\ddot{x}^c + \Gamma^c_{ab}\dot{x}^a\dot{x}^b = \dot{X}^c + \Gamma^c_{ab} X^a X^b = X^a \partial_a X^c + \Gamma^c_{ab} X^a X^b = \nabla_X X. \tag{5.4.5}$$

Hence

PROPOSITION 5.4.2. *The curve $\tau \mapsto \gamma(\tau)$ is a geodesic if and only its acceleration is zero.*

This is pleasing because it is a natural generalization of what happens in Minkowski space. In that case, the geodesics are of the form

$$\tau \mapsto X + U\tau$$

where X and U are constant vectors. Then the tangent vector is U and the acceleration is zero. And conversely any curve with zero acceleration in Minkowski space is of the above form and extremizes the energy.

DEFINITION 5.4.3. If $\gamma(\tau)$ is a curve in $\mathscr{M}$ with tangent vector $\dot{\gamma} = X$ then a vector field Y along γ is said to be parallel, parallel-transported, or parallel-propagated along γ if

$$\nabla_X Y = 0 \text{ along } \gamma. \tag{5.4.6}$$

Explicitly, in local coordinates, the parallel propagation equation (5.4.6) has the form

$$\dot{Y}^c + \Gamma^c_{ab}\dot{x}^a Y^b = 0. \tag{5.4.7}$$

Note that to be completely explicit, the Γs here are evaluated at the point $x^a(\tau)$. With the curve fixed, $x^a(\tau)$ and $\dot{x}^a(\tau)$ are known, and so (5.4.7) is a first-order system of equations for the unknown components $Y^b(\tau)$ as a functions of τ. Thus given any point τ_0 and a given tangent vector Z at $\gamma(\tau_0)$, there is a unique solution of (5.4.7) with initial condition $Y(\tau_0) = Z$. In this situation, we say that $Y(\tau_1)$ is obtained from Z by parallel transport along γ.

In Minkowski space we know when two vectors at different points are parallel (or point in the same direction). In a general curved space $\mathscr{M}$ there is no such global notion of parallelism, and the above is the best one can do.

5.4.2. Photons. A photon worldline in GR is a null geodesic $\tau \mapsto \gamma(\tau)$,

$$g(\dot{\gamma}, \dot{\gamma}) = 0.$$

The frequency 4-vector of a photon, K, is assumed to be a constant multiple of its velocity vector $\dot{\gamma}$. If U is an observer with velocity 4-vector U (i.e. $g(U,U) = 1$), and p is an event on the photon worldline and the observer's worldline, then the frequency measured by the observer is $g(p)[U(p), K(p)]$. Compare with analogous situation in SR, §3.3.

5.5. Local inertial coordinates

Let g be a lorentzian metric with components g_{ab} in local coordinates defined near $x^a = 0$. We have seen that we can suppose that $g_{ab}(0) = \eta_{ab}$, the standard Minkowski metric, by a linear change of coordinates. Then

$$g_{ab}(x) = \eta_{ab} + H_{abc}x^c + O(|x|^2) \tag{5.5.1}$$

for some set of numbers H_{abc} with

$$H_{abc} = H_{bac}. \tag{5.5.2}$$

Proposition 5.5.1. *The Christoffel symbols of* (5.5.1) *at $x = 0$ are*

$$\Gamma^c_{ab}(0) = \frac{1}{2}\eta^{cs}\left(H_{sab} + H_{sba} - H_{abs}\right). \tag{5.5.3}$$

Proof. This follows by substituting

$$\partial_s g_{ab}(0) = H_{abs} \tag{5.5.4}$$

into the formula for the Christoffels (5.3.2). □

In particular, geodesics do not have the simple form $\ddot{x}^a = 0$, even at $x = 0$, if $H \neq 0$. So such a coordinate system is not a particularly good approximation to an inertial coordinate system in Minkowski space.

However, a further change of coordinates can always be made to kill the coefficients H_{abc} at $x = 0$.

Theorem 5.5.2. *Let $(\mathscr{M}, g)$ be a space-time and let p be any point of $\mathscr{M}$. There exists a choice of coordinates x^a such that $x^a(p) = 0$ and*

$$g_{ab} = \eta_{ab} + O(|x|^2) \text{ for small } x. \tag{5.5.5}$$

Such coordinates are called *local inertial coordinates at p*. Note that *no statement is made about the form of the metric at other points near, but not equal to p.*

The notation is motivated as follows: from the Proposition, if the metric has the form (5.5.5) then $\Gamma^c_{ab} = 0$ with respect to these coordinates. Thus the geodesic equations

$$\ddot{x}^c + \Gamma^c_{ab}\dot{x}^a\dot{x}^b = 0$$

reduce to $\ddot{x}^c$ *at* $x^a = 0$, so at this point, at least, the equation is the same as for inertial worldlines in Minkowski space. In particular freely falling particles and free photons will have worldlines that appear straight in a very small neighbourhood of $x = 0$ in these special coordinates. This is the precise sense in which GR reduces to SR over small length- and time-scales, but it only works well relative to one of these inertial coordinate systems.

5.5.1. Proof of Theorem 5.5.2. We may suppose that

$$g_{ab} = \eta_{ab} + H_{abc}x^c$$

as the next term in the Taylor expansion is already of quadratic order. Consider the coordinate transformation

$$x^c = y^c - \frac{1}{2}G^c_{ab}y^a y^b \tag{5.5.6}$$

where the G^c_{ab} is an array of numbers to be determined, symmetric in the indices ab. The Jacobian of the transformation is

$$\frac{\partial \widetilde{x}^c}{\partial x^p} = J^c_p = \delta^c_p + j^c_p, \text{ where } j^c_p = G^c_{ap} y^a. \tag{5.5.7}$$

This is invertible when $y = 0$ so by the inverse function theorem, (5.5.6) has a smooth inverse as a mapping between a neighbourhood of $y = 0$ and a neighbourhood of $x = 0$.

Now we calculate, keeping only the leading terms (because we don't really care what's happening at $O(|y|^2)$ and beyond):

$$\begin{aligned} g_{pq} \mathrm{d}x^p \mathrm{d}x^q &= g_{pq}(x)(\mathrm{d}y^p - G^p_{ac} y^a \mathrm{d}y^c)(\mathrm{d}y^q - G^q_{bd} y^b \mathrm{d}y^d) \\ &= g_{pq} \mathrm{d}y^p \mathrm{d}y^q - G_{acq} y^a \mathrm{d}y^c \mathrm{d}y^q - G_{bdp} y^b \mathrm{d}y^p \mathrm{d}y^d + O(|y|^2) \end{aligned} \tag{5.5.8}$$

where we have set

$$G_{acp} = g_{pq} G^q_{ac}, \text{ so } G_{acp} = G_{cap}. \tag{5.5.9}$$

Hence

$$\widetilde{g}_{pq}(y) = g_{pq}(x(y)) - G_{apq} y^a - G_{aqp} y^a + O(|y^2|). \tag{5.5.10}$$

Now on the RHS we still have $g(x)$ and we need to write this in terms of y. The inverse to our transformation has the form

$$y^a = x^a + \frac{1}{2} G^a_{bc} x^b x^c + O(|x|^3) \tag{5.5.11}$$

as you can see by inserting (5.5.6) on the RHS of this equation. It follows that

$$g_{pq}(x(y)) = g_{pq}(y) + O(|y|^2) = \eta_{pq} + H_{pqr} y^r + O(|y|^2).$$

and so

$$\widetilde{g}_{pq}(y) = \eta_{pq} + H_{pqa} y^a - G_{apq} y^a - G_{aqp} y^a + O(|y|^2). \tag{5.5.12}$$

Thus we will get rid of the first order terms in y if we can choose the numbers G so that

$$G_{apq} + G_{aqp} = H_{pqa}. \tag{5.5.13}$$

(Here we have used the symmetry of the indices of G to neaten up the equation.) This is an equation for the array of number G in terms of the array of numbers H. In Lemma 5.5.9 below it is shown that this can always be solved, the solution being

$$G_{abc} = \frac{1}{2}(H_{acb} + H_{bca} - H_{abc}). \tag{5.5.14}$$

Note that this formula depends in an important way on the symmetry (5.5.9) of G. Inverting (5.5.9) we define

$$G^d_{ab} = \eta^{cd} G_{abc} = \frac{1}{2} \eta^{cd}(H_{acb} + H_{bca} - H_{abc}).$$

Thus G is uniquely determined by H, and the by the above calculations, the change of coordinates (5.5.6) gives metric components $\widetilde{g}_{ab}$ which satisfy the conditions of the Theorem. The proof is complete.

Remark 5.5.3. The array of numbers G^c_{ab} is nothing but $\Gamma^c_{ab}(0)$, the Christoffel symbols of the metric components g_{ab}, evaluated at $x = 0$ (cf. Proposition 5.5.1).

It remains to prove:

Lemma 5.5.4. *The equation* (5.5.13) *is solved by* (5.5.14),

$$G_{abc} = \frac{1}{2}(H_{cab} + H_{cba} - H_{abc})$$

Proof. One proof of this is simply to substitute (5.5.14) into (5.5.13) and see that it works. Namely, G is symmetric in its first two indices: simply switch them, and use the symmetry of H in its first two indices. And then

$$G_{abc} + G_{acb} = \frac{1}{2}(H_{cab} + H_{cba} - H_{abc} + H_{bac} + H_{bca} - H_{acb}). \tag{5.5.15}$$

Now use the symmetry of H in its first two indices to arrange the indices as far as possible in alphabetical order:

$$G_{abc} + G_{acb} = \frac{1}{2}(H_{acb} + H_{bca} - H_{abc} + H_{abc} + H_{bca} - H_{acb}) = H_{bca} \tag{5.5.16}$$

as required.

There is also a derivation of this formula in the Problem set, Problem 4.9. $\square$

5.6. A sneak preview of curvature

By Theorem 5.5.2, in inertial coordinates,

$$g_{ab} = \eta_{ab} + \frac{1}{2}P_{abcd}x^c x^d + O(|x|^3), \tag{5.6.1}$$

where P is an array of numbers with the symmetry properties

$$P_{abcd} = P_{bacd} = P_{abdc}. \tag{5.6.2}$$

We ask: is there a change of coordinates which can get rid of the P term here? The answer is no in general, but it is interesting to try.

It would be natural to try

$$\widetilde{x}^a = x^a + \frac{1}{6}W^a_{bcd}x^b x^c x^d \tag{5.6.3}$$

to change P. Note that here the array of numbers W satisfies

$$W^a_{bcd} \text{ is totally symmetric in } bcd \tag{5.6.4}$$

in the sense that

$$W^a_{bcd} = W^a_{cbd} = W^a_{bdc} = W^a_{dcb} \text{ etc.}$$

In Problem 4.8, you are invited to show that if g_{ab} is as in (5.6.1) and $\widetilde{x}$ and x are related by (5.6.3), then

$$\widetilde{g}_{ab} = \eta_{ab} + \frac{1}{2}\widetilde{P}_{abcd}\widetilde{x}^c\widetilde{x}^d + O(|\widetilde{x}|^3) \tag{5.6.5}$$

where

$$\widetilde{P}_{abcd} = P_{abcd} - W_{abcd} - W_{abdc} \tag{5.6.6}$$

Now I claim that this cannot be solved in general, because W just does not have enough parameters! For this, we need to do some counting.

5.6.1. Counting tensor components.

Definition 5.6.1. A tensor T of type $(0,m)$ in n dimensions is said to be totally symmetric if for the corresponding m-linear form we have

$$T(v_1,\dots,v_i,\dots,v_j,\dots,v_m) = T(v_1,\dots,v_j,\dots,v_i,\dots,v_m)$$

for any i and j. In components this is the same as saying

$$T_{p_1\dots p_i\dots p_j\dots p_m} = T_{p_1\dots p_j\dots p_i\dots p_m}.$$

Proposition 5.6.2. *The dimension of the vector space of all totally symmetric tensors of type* $(0,m)$ *in* n *dimensions is*

$$\binom{n+m-1}{m}. \tag{5.6.7}$$

Definition 5.6.3. A tensor of type $(0,m)$ in n dimensions is totally skew (or totally skew symmetric if the corresponding m-linear form has the property

$$T(v_1,\dots,v_i,\dots,v_j,\dots,v_m) = -T(v_1,\dots,v_j,\dots,v_i,\dots,v_m)$$

for any i and j.

PROPOSITION 5.6.4. *The dimension of the vector space of totally skew tensors of type* $(0, m)$ *in* n *dimensions is*

$$\binom{n}{m} \tag{5.6.8}$$

(In particular the dimension is 1 *if* $n = m$ *and* 0 *if* $m > n$*.)*

PROOF. We shall prove both of these together. By way of a warm-up let's make sure that we understand that

$$\text{The dimension of the space of all tensors of type } (0, m) \text{ in } n \text{ dimensions is } n^m \tag{5.6.9}$$

To see this, note that such a tensor has m indices. We can choose each one of these in n ways[1]. All choices are independent because there are no symmetry conditions. So the total number of possibilities is n^m. The number of these choices is equal to the dimension of the space of these tensors.

Let us move on to the proof of Proposition 5.6.4. Again we have a tensor with m indices. For ease of exposition, suppose that $m = 3$. Any component with two indices the same must be zero, because (for example)

$$T_{115} = -T_{115}$$

by switching the first two indices. So the only non-zero components of T have *distinct indices*. If we have a set of 3 distinct indices, say, 523, then we can use the skew symmetry to relate the T_{523} to T_{235}, where the indices are now in increasing order:

$$T_{523} = -T_{253} = T_{235}$$

(switching first the first two indices and then the last two). Thus *the number of independent components of a totally skew tensor of type* $(0, 3)$ *is equal to the number of unordered subsets of* 3 *elements of the set* $\{1, \ldots, n\}$. This is the binomial coefficient

$$\binom{n}{3}$$

The general case, with 3 replaced by m, works in the same way.

Finally let us prove Proposition 5.6.2. The big difference from the case of skew symmetry is that now that indices can take the same value without that component being zero. For a small number of indices (e.g. 3) it's possible to count by hand. There are n components where the indices are the same:

$$T_{111}, T_{222}, \ldots, T_{nnn}.$$

There are $n(n-1)$ with precisely two indices the same:

$$T_{112}, T_{113}, \ldots, T_{11n}; T_{221}, T_{223}, \ldots.$$

And there are $n(n-1)(n-2)/6$ where the indices are all distinct. Thus the total number of independent components of a totally symmetric tensor of type $(0, 3)$ in n dimensions is:

$$n + n(n-1) + \frac{1}{6}n(n-1)(n-2) = \frac{n(n+1)(n+2)}{6} \tag{5.6.10}$$

which checks with (5.6.7) if $m = 3$.

This approach can be generalized to tensor of any rank, but it's pretty messy. The following is the cunning way of doing it.

Consider a collection of indices on our totally symmetric tensor with m indices. It will consist of m_1 1's, m_2 2's, and so on, up to m_n n's. Here the m_j are allowed to be 0, but the constraint that the tensor is of type $(0, m)$ is

$$m_1 + m_2 + \cdots + m_n = m \tag{5.6.11}$$

This combinatorial problem can be visualised in the following way. Consider an arrangement of m coins and $n - 1$ pencils in a line, as in the example below:

[1] Since they vary from 1 to n

$$\bullet \quad \bullet \quad \bullet \,\Big|\, \bullet \quad \bullet \,\Big|\, \bullet \,\Big|\, \bullet \quad \bullet \,\Big|\Big|$$

Given such an arrangement, we count the coins to the left of the first pencil, and call that m_1. Then we count the coins between the first and second pencils, and call that m_2. Proceeding in this way, we get a collection of n integers $m_j \geqslant 0$, satisfying the constraint (5.6.11). (Note that $m_j = 0$ if two of the pencils are right next to each other, or if there is a pencil at the very beginning or the very end of the line.)

In the pictured configuration there are 5 pencils and 8 coins, and

$$m_1 = 3,\ m_2 = 2,\ m_3 = 1,\ m_4 = 2,\ m_5 = 0,\ m_6 = 0.$$

This would correspond to the component

$$T_{11122344}$$

of a totally symmetric tensor of type $(0, 8)$ in 6 dimensions. The number of arrangements of m coins and $n - 1$ pencils is the same as the number of ways of choosing m objects (the ones to be called coins) from a total of $m + n - 1$. This is the binomial coefficient (5.6.7). $\square$

5.6.2. Sneak preview of curvature, continued. We've seen that the coordinate transformation (5.6.3) allows us to change the array P_{abcd}, where

$$P_{abcd} = P_{bacd} = P_{abdc} \tag{5.6.12}$$

to

$$\widetilde{P}_{abcd} = P_{abcd} - W_{abcd} - W_{abdc} \tag{5.6.13}$$

where W is symmetric in its last 3 indices.

What is the dimension of the space of P's? P is symmetric in its first two indices and its third and fourth indices, but there is no other symmetry. So it is like a 2-index object, P_{IJ}, where I and J run over a basis of the space of symmetric 2-index tensors. If that dimension is N, the dimension of the space of Ps will be N^2. But we've seen that $N = n(n+1)/2$ (if the dimension is n—we work generally for the moment). Hence:

> In an n-dimensional manifold, the dimension of the space of Ps is
> $$\frac{1}{4}n^2(n+1)^2$$

On the other hand, the dimension of the space of Ws is n (for the extra index) times $n(n+1)(n+2)/6$ (the dimension of totally symmetric tensors of type $(0, 3)$).

> In an n-dimensional manifold, the dimension of the space of Ws is
> $$\frac{1}{6}n^2(n+1)(n+2).$$

Thus the dimension of the space of P's minus the dimension of the space of W's is

$$\begin{aligned} \frac{n^2(n+1)^2}{4} - \frac{n^2(n+1)(n+2)}{6} &= \frac{n^2(n+1)}{12}\left(3(n+1) - 2(n+2)\right) \\ &= \frac{n^2(n+1)}{12}(n-1) \\ &= \frac{1}{12}n^2(n^2-1) \end{aligned} \tag{5.6.14}$$

Since this number if positive for $n \geqslant 2$, it will be impossible to solve (5.6.13) to make $\widetilde{P} = 0$ in general. (Our calculation shows that $\widetilde{P} = 0$ is a system of linear equations with more equations than unknowns.)

Thus, while some components of P can be killed by coordinate transformations, there are others, in fact $n^2(n^2-1)/12$ of them, which cannot. These 'unkillable' components of P form a tensor called the *curvature* of g at the point $x=0$.

In fact, the *Riemann curvature tensor* at $x=0$ is built out of P in the following way:

$$R_{abcd} = \frac{1}{2}(P_{acbd} + P_{bdac} - P_{adbc} - P_{bcad}). \tag{5.6.15}$$

It can be checked that if P is changed to $\widetilde{P}$ as in (5.6.13), then the components of R do not change! Thus if this particular combination of components of P is non-zero then it cannot be killed by coordinate transformation.

These matters will be discussed much more extensively in the next chapter, where we shall see a different, but equivalent, definition of curvature.