Preliminaries

We will need various concepts from differential geometry. For more details one can consult chapter 2 of Wald.

Manifold

An n-dimensional smooth manifold M can be covered with a set of local coordinate charts $\{\mathcal{O}_\alpha, \psi_\alpha\}$ with $\mathcal{O}_\alpha$ open subsets in M and ψ_α maps from $\mathcal{O}_\alpha$ to open subsets of $\mathbb{R}^n$, in a 1-1 and onto manner. For any two sets $\mathcal{O}_\alpha$ and $\mathcal{O}_\beta$ that overlap, the composition maps $\psi_\alpha \circ \psi_\beta^{-1}$, which map an open subset of $\mathbb{R}^n$ to another, are taken to be C^∞.

Tangent Vectors

The key idea of a tangent vector is that it is a directional derivative. Let $\mathcal{F} = \{f : M \to \mathbb{R}, f \in C^\infty\}$. A vector V at a point $p \in M$ is defined to be a map $V : \mathcal{F} \to \mathbb{R}$ taking f to $V(f)$ which satisfies

1. Linearity: $V(af + bg) = aV(f) + bV(g)$, where $a, b \in \mathbb{R}$ and $f, g \in \mathcal{F}$.
2. Leibniz rule: $V(fg) = f(p)V(g) + g(p)V(f)$

In local coordinates we can write $V = V^\mu \frac{\partial}{\partial x^\mu}$, where $V^\mu \in \mathbb{R}$ are the components of V and $V(f) = V^\mu \partial_\mu f$. More precisely, $\{\partial x^\mu\}$ are a basis set of vectors and V^μ are the components with respect to this basis. Notice that $V = V^\mu \partial_\mu = V^\mu \frac{\partial x^{\nu'}}{\partial x^\mu} \partial_{\nu'} = V^{\nu'} \partial_{\nu'}$ gives the usual transformation law for the components.

A *vector field* on M is a specification of a vector at each point on M. The vector field is said to be smooth if $V(f)$, which is now a function on M, is smooth. This is equivalent to that statement that in a local coordinate patch, the components $V^\mu(x)$ are smooth functions of x.

Consider a smooth curve $\gamma : \mathbb{R} \to M$, taking λ to $\gamma(\lambda)$. Observe that for any function $f \in \mathcal{F}$ we have the composition $f \circ \gamma$ is a map from $\mathbb{R}$ to itself. Thus, for each point p that lies on the curve in M we can specify a vector V using the rule: $V(f) = \frac{d}{d\lambda}(f \circ \gamma)\big|_{\gamma^{-1}(p)}$. In local coordinates we have $V(f) = \frac{\partial f}{\partial x^\mu} \frac{dx^\mu}{d\lambda}$ and hence the components of the vector are given by $V^\mu = \frac{dx^\mu}{d\lambda} = \dot{x}^\mu$.

Conversely, given a vector field V we can construct the *integral curves* which have the property that one and only one curve passes through each point p and the tangent vector to the curve at p is $V(p)$.

A useful fact is that for a given vector field V it is possible to choose local coordinates such that $V = \frac{\partial}{\partial x^1}$ i.e. $V^\mu = (1, 0, \ldots, 0)$.

Tensors

We can define co-vectors or *one-forms* (more on forms later). The co-vectors at a point

$p \in M$ live in the vector space dual to the vector space of vectors at p. Since they live in the dual vector space they are linear maps taking vectors to the real numbers. The basis of co-vectors that are dual to the basis of vectors $\{\partial_\mu\}$ is denoted by $\{dx^\nu\}$, with the action giving δ^ν_μ. We can then write a general co-vector at a point as $W = W_\mu dx^\mu$, where W_μ are the components of the co-vector. The action of this on an arbitrary vector $V = V^\mu \partial_\mu$ is simply the contraction $V^\mu W_\mu$. A co-vector field is a specification of a co-vector at each point on M.

Tensors of type (r, s) have components $T^{\mu_1 \cdots \mu_r}{}_{\nu_1 \ldots \nu_s}$ We can define higher rank tensors by taking tensor products. Eg if S and T are two co-vectors then $W = S \otimes T$ is a tensor of type $(0, 2)$ with components $B_{\mu\nu} \equiv S_\mu T_\nu$.

Symmetrisation and antisymmetrisation Tensors with indices in the same position (i.e either up or down) can have symmetry properties. For example we say that $S_{\mu\nu}$ is symmetric if $S_{\mu\nu} = S_{\nu\mu}$. Similarly, $T^{\mu\nu}$ is symmetric if $T^{\mu\nu} = T^{\nu\mu}$. We also say that $A_{\mu\nu}$ (or $B^{\mu\nu}$) is anti-symmetric if $A_{\mu\nu} = -A_{\nu\mu}$ (or $B^{\mu\nu} = -B^{\nu\mu}$). Tensors with additional indices can be symmetric or anti-symmetric in some or all of the indices in the same position eg we could have $T^\mu{}_{\rho\sigma} = T^\mu{}_{\sigma\rho}$.

We can define symmetrisation and antisymmetrisation of the indices of a tensor T with two indices as follows:

$$T_{(\mu\nu)} \equiv \frac{1}{2}(T_{\mu\nu} + T_{\nu\mu}) \,, \qquad T_{[\mu\nu]} \equiv \frac{1}{2}(T_{\mu\nu} - T_{\nu\mu})$$

Clearly $T_{(\mu\nu)}$ is a symmetric tensor and $T_{[\mu\nu]}$ is an antisymmetric tensor (a two-form). Note that $T_{\mu\nu} = T_{(\mu\nu)} + T_{[\mu\nu]}$. If $S^{\mu\nu}$ is a symmetric tensor and $A^{\mu\nu}$ is an antisymmetric tensor, $S^{\mu\nu} = S^{(\mu\nu)}$ and $A^{\mu\nu} = A^{[\mu\nu]}$ and for an arbitrary tensor $T_{\mu\nu}$ we have $S^{\mu\nu} T_{\mu\nu} = S^{\mu\nu} T_{(\mu\nu)}$, $A^{\mu\nu} T_{\mu\nu} = A^{\mu\nu} T_{[\mu\nu]}$. If $S^{\mu\nu}$ is a symmetric tensor and $A_{\mu\nu}$ is an antisymmetric tensor then $S^{\mu\nu} A_{\mu\nu} = 0$.

For a tensor with three indices we can similarly define

$$T_{(\mu\nu\rho)} \equiv \frac{1}{3!}(T_{\mu\nu\rho} + T_{\mu\rho\nu} + T_{\rho\mu\nu} + T_{\rho\nu\mu} + T_{\nu\rho\mu} + T_{\nu\mu\rho}) \,,$$
$$T_{[\mu\nu\rho]} \equiv \frac{1}{3!}(T_{\mu\nu\rho} - T_{\mu\rho\nu} + T_{\rho\mu\nu} - T_{\rho\nu\mu} + T_{\nu\rho\mu} - T_{\nu\mu\rho})$$

Note, however, that $T_{\mu\nu\rho} \neq T_{(\mu\nu\rho)} + T_{[\mu\nu\rho]}$. We can also define (anti)symmetrisation on a subset of indices if desired eg $T_{\mu[\nu\rho]}$

Lie derivative

For a given vector field V we can define a Lie derivative $\mathcal{L}_V$ which acts on tensors. If T is a tensor of type (r, s) then $\mathcal{L}_V T$ is also a tensor of type (r, s). It is a linear map. The idea of the definition is that one is taking the derivative along the integral curves of V (see Wald). Instead of following that route, lets see it in action.

1. Acting on a function, f, we have $\mathcal{L}_V(f) = V(f)$.

2. Acting a vector field W, we have $\mathcal{L}_V W = [V,W]$, where $[V,W](f) \equiv V(W(f)) - W(V(f))$. In components we have $(\mathcal{L}_V W)^\nu = V^\mu \partial_\mu W^\nu - W^\mu \partial_\mu V^\nu$ (exercise)

Furthermore, $\mathcal{L}_V$ commutes with contraction and also satisfies the Leibniz rule: $\mathcal{L}_V(S \otimes T) = (\mathcal{L}_V S) \otimes T + S \otimes (\mathcal{L}_V T)$. To take an example, in components if we have $B_{\mu\nu} \equiv S_\mu T_\nu$ then $(\mathcal{L}_V B)_{\mu\nu} = (\mathcal{L}_V S)_\mu T_\nu + S_\mu (\mathcal{L}_V T)_\nu$. In fact, given the action on functions and vector fields, these properties are sufficient to define the action of $\mathcal{L}_V$ on any type of tensor. As an exercise you can verify that, for example,

$$\begin{aligned}(\mathcal{L}_V T)_\nu &= V^\mu \partial_\mu T_\nu + T_\mu \partial_\nu V^\mu \\ (\mathcal{L}_V B)_{\nu\rho} &= V^\mu \partial_\mu B_{\nu\rho} + B_{\mu\rho} \partial_\nu V^\mu + B_{\nu\mu} \partial_\rho V^\mu .\end{aligned} \tag{1}$$

Metric

We now assume that we have a metric $g_{\mu\nu}$ on the manifold M. This is a symmetric tensor that is non-degenerate everywhere on M. As such the inverse metric $g^{\mu\nu}$ also exists with the defining property that $g_{\mu\nu} g^{\nu\rho} = \delta^\rho_\mu$. The metric can be used to raise and lower indices of a tensor. eg Given a vector V^μ we can define a covector $V_\mu \equiv g_{\mu\nu} V^\nu$ and note that we will use the same letter to denote the vector and the co-vector. The metric can also be used to define the Levi-Civita covariant derivative ∇. For example, recall that $\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\rho} V^\rho$ and $\nabla_\mu W_\nu = \partial_\mu W_\nu - \Gamma^\rho_{\mu\nu} W_\rho$ where $\Gamma^\mu_{\nu\rho} = \Gamma^\mu_{\rho\nu}$ are the Christoffel symbols defined by

$$\Gamma^\mu_{\nu\rho} = \frac{1}{2} g^{\mu\sigma} \left(g_{\sigma\nu,\rho} + g_{\sigma\rho,\nu} - g_{\nu\rho,\sigma} \right) .$$

By definition we have $\nabla_\mu g_{\rho\sigma} = 0$. For an arbitrary vector field we also have

$$(\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu) V^\rho = R_{\mu\nu}{}^\rho{}_\sigma V^\sigma \tag{2}$$

where $R_{\mu\nu}{}^\rho{}_\sigma$ are the components of the Riemann tensor.

While the Lie derivative does not depend on the metric, and hence ∇, when a metric is present it can be useful to write the Lie derivative using ∇. For example, we have (exercise)

1. $\mathcal{L}_V(f) = V^\mu \nabla_\mu f$

2. $(\mathcal{L}_V W)^\nu = V^\mu \nabla_\mu W^\nu - W^\mu \nabla_\mu V^\nu$

Killing vector

A *Killing vector field* V has the property that $(\mathcal{L}_V g) = 0$. In components we can calculate

$$\begin{aligned}(\mathcal{L}_V g)_{\mu\nu} &= V^\rho \nabla_\rho g_{\mu\nu} + g_{\rho\nu} \nabla_\mu V^\rho + g_{\mu\rho} \nabla_\nu V^\rho \\ &= \nabla_\mu V_\nu + \nabla_\nu V_\mu \\ &= 2 \nabla_{(\mu} V_{\nu)}\end{aligned}$$

and hence a Killing vector is equivalent to the condition $\nabla_{(\mu}V_{\nu)} = 0$.

Working in local coordinates such that $V = \frac{\partial}{\partial x^1}$, from (??) the condition that V is a Killing vector is simply that $\frac{\partial}{\partial x^1} g_{\mu\nu} = 0$ (exericse). This gives a useful way to spot whether a metric admits any Killing vectors. One should be careful though, since the coordinates that the metric is presented in may not be of this type.

Geodesic motion of test particles

We now consider a spacetime (M, g) with g a Lorentzian metric. We will mostly be using units for which $G = \hbar = c = 1$ in this course. We are interested in a particle of rest mass m moving on a curve γ, with parameter λ, from point A to point B in M. In local coordinates the curve is specified via $x^\mu(\lambda)$.

The action for the test particle is determined by the proper time in moving from A to B:

$$I = m \int_{\tau_A}^{\tau_B} d\tau$$

where $d\tau^2 \equiv -ds^2 = -g_{\mu\nu} dx^\mu dx^\nu$. We can thus write

$$I[x^\mu(\lambda)] = m \int_{\lambda_A}^{\lambda_B} d\lambda [-g_{\mu\nu}(x(\lambda)) \dot{x}^\mu \dot{x}^\nu]^{1/2}$$

The test particle moves on a geodesic which extremises this action $\frac{\delta I}{\delta x^\mu(\lambda)} = 0$, where the variations are anchored at the end points: $\delta x(\lambda_A) = \delta x(\lambda_B) = 0$.

It is convenient to use an alternative action by introducing a new object along the curve, the "einbein", $e(\lambda) > 0$ via:

$$\hat{I}[x^\mu(\lambda), e(\lambda)] = \frac{1}{2} \int_{\lambda_A}^{\lambda_B} d\lambda [e^{-1} g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu - m^2 e]$$

This new action gives equivalent equations of motion. To prove this we note that we have the two variations $\frac{\delta \hat{I}}{\delta x^\mu(\lambda)} = 0$ and $\frac{\delta \hat{I}}{\delta e(\lambda)} = 0$ to impose. Now the latter equation can be solved for e as:

$$e = \frac{1}{m}[-g_{\mu\nu}(x(\lambda)) \dot{x}^\mu \dot{x}^\nu]^{1/2} = \frac{1}{m}\frac{d\tau}{d\lambda} \equiv e[x(\lambda)]$$

We also have that $\hat{I}[x^\mu, e[x]] = -I[x]$. So we calculate

$$\begin{aligned} -\frac{\delta I}{\delta x^\mu(\lambda)} &= \left.\frac{\delta \hat{I}}{\delta x^\mu(\lambda)}\right|_{e[x]} + \int_{\lambda_A}^{\lambda_B} d\lambda' \left.\frac{\delta \hat{I}}{\delta e(\lambda')}\right|_{e[x]} \frac{\delta e[x(\lambda')]}{\delta x(\lambda)} \\ &= \left.\frac{\delta \hat{I}}{\delta x^\mu(\lambda)}\right|_{e[x]} \end{aligned}$$

Thus, the condition for geodesics, $\frac{\delta I}{\delta x^\mu(\lambda)} = 0$, is equivalent to $\left.\frac{\delta \hat{I}}{\delta x^\mu(\lambda)}\right|_{e[x]} = 0$ combined with imposing $e = e[x]$, which completes the proof.

By explicit calculation of $\frac{\delta \hat{I}}{\delta x^\mu(\lambda)} = 0$ we thus find (exercise) that the condition for geodesics is

$$\frac{D}{d\lambda}\dot{x}^\mu = (e^{-1}\dot{e})\dot{x}^\mu, \qquad e = \frac{1}{m}\frac{d\tau}{d\lambda}$$

where $\frac{D}{d\lambda}\dot{x}^\mu \equiv \dot{x}^\rho \nabla_\rho \dot{x}^\mu = \ddot{x}^\mu + \Gamma^\mu_{\lambda\rho}\dot{x}^\lambda \dot{x}^\rho$. Now there is a freedom in the choice of the parametrisation, λ, of the curve which is equivalent to the choice of e. To obtain an *affinely parametrised* geodesic we choose $\dot{e} = 0$ which is equivalent to choosing $\lambda = a\tau + b$ where a, b are constants and such geodesics satisfy $\frac{D}{d\lambda}\dot{x}^\mu = 0$. It is worth noting that we can obtain the equations for an affinely parametrised geodesic using proper time, by varying the action $\int d\tau (\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} g_{\mu\nu})$ and separately imposing $\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} g_{\mu\nu} = -1$.

Recall that a vector V is said to be a *parallely transported along a curve* with tangent vector $T^\mu = \dot{x}^\mu$ if and only if $T^\nu \nabla_\nu V^\mu = f V^\mu$ for some function f. This is the same as $\frac{D}{d\lambda}V^\mu = f V^\mu$. Thus a geodesic is a curve that has a tangent vector that is parallely transported along it.

We note that we cannot use the action $I[x]$ when $m = 0$ but we can still use $\hat{I}[x, e]$ and hence $\hat{I}$ is more general. In this case we find we still have $\frac{D}{d\lambda}\dot{x}^\mu = (e^{-1}\dot{e})\dot{x}^\mu$ but $\frac{\delta \hat{I}}{\delta e} = 0$ now implies that $ds^2 = g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = 0$, i.e. it moves along a null curve.

To summarise, affinely parametrised geodesics satisfy $\frac{D}{d\lambda}\dot{x}^\mu = 0$. If $m = 0$ then $ds^2 = g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = 0$ and we have a null affinely parametrised geodesic. If $m \neq 0$ then $ds^2 = g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = -d\tau^2$ and $\lambda = a\tau + b$ giving a time-like affinely parametrised geodesic.

Killing vectors and conservation laws

Let V^μ be a Killing vector. Consider the shift in coordinates $x^\mu + \epsilon V^\mu$, where ϵ is infinitesimal. We calculate

$$\begin{aligned}
\delta \hat{I} &\equiv \hat{I}[x + \epsilon V, e] - \hat{I}[x, e] \\
&= \epsilon \int d\lambda e^{-1} \dot{x}^\mu \dot{x}^\nu V_{\mu;\nu} \\
&= \epsilon \int d\lambda e^{-1} \dot{x}^\mu \dot{x}^\nu V_{(\mu;\nu)} = 0
\end{aligned}$$

where to get the second lines requires some calculation. Noether' theorem implies that there is a conserved charge Q defined by $Q \equiv p_\mu V^\mu$ where $p_\mu \equiv \frac{\delta \mathcal{L}}{\delta \dot{x}^\mu} = e^{-1}\dot{x}^\nu g_{\mu\nu}$. As an exercise one can directly check that $\frac{d}{d\lambda}Q = 0$ and hence Q is indeed a conserved quantity.