3 Parallel transport and geodesics

3.1 Differentiation along a curve

As a prelude to parallel transport we consider another form of differentiation: differentiation along a curve. A curve is a parametrized path through spacetime: \( x(\lambda) \), where \( \lambda \) is a parameter that varies smoothly and monotonically along the path. The curve has a tangent vector \( \vec{V} \equiv d\vec{x}/d\lambda = (dx^\mu/d\lambda) \vec{e}_\mu \). Here one must be careful about the interpretation: \( x^\mu \) are not the components of a vector; they are simply 4 scalar fields. However, \( \vec{V} = d\vec{x}/d\lambda \) is a vector (i.e. a tangent vector in the manifold).

If we wish, we could make \( \vec{V} \) a unit vector (provided \( \vec{V} \) is non-null) by setting \( d\lambda = |d\vec{x} \cdot d\vec{x}|^{1/2} \) to measure path length along the curve. However, we will impose no such restriction in general.

Now, suppose that we have a scalar field \( f_\lambda \) defined along the curve. We define the derivative along the curve by a simple extension of equations (36) and (38) of the first set of lecture notes:

\[
\frac{df}{d\lambda} \equiv \nabla_\lambda f \equiv \langle \nabla f, \vec{V} \rangle = V^\mu \partial_\mu f , \quad \vec{V} \equiv \frac{d\vec{x}}{d\lambda}.
\]  

(21)

We have introduced the symbol \( \nabla_\lambda \) for the directional derivative, i.e. the covariant derivative along \( \vec{V} \), the tangent vector to the curve \( x(\lambda) \). This is a natural generalization of \( \nabla_\mu \), the covariant derivative along the basis vector \( \vec{e}_\mu \).

For the derivative of a scalar field, \( \nabla_\lambda \) involves just the partial derivatives \( \partial_\mu \). Suppose, however, that we differentiate a vector field \( \vec{A}_\lambda \) along the curve. Now the components of the gradient \( \nabla_\mu A^\nu \) are not simply the partial derivatives but also involve the connection. The same is true when we project the gradient onto the tangent vector \( \vec{V} \) along a curve:

\[
\frac{d\vec{A}}{d\lambda} \equiv \frac{D A^\mu}{d\lambda} \vec{e}_\mu \equiv \nabla_\lambda \vec{A} \equiv \langle \nabla \vec{A}, \vec{V} \rangle = V^\nu (\nabla_\mu A^\nu) \vec{e}_\mu = \left( \frac{dA^\mu}{d\lambda} + \Gamma^\mu_{\kappa\nu} A^\kappa V^\nu \right) \vec{e}_\mu .
\]  

(22)
We retain the symbol $\nabla_V$ to indicate the covariant derivative along $V$ but we have introduced the new notation $D/d\lambda = V^\mu \nabla_\mu \neq d/d\lambda = V^\mu \partial_\mu$.

### 3.2 Parallel transport

The derivative of a vector along a curve leads us to an important concept called parallel transport. Suppose that we have a curve $x(\lambda)$ with tangent $\vec{V}$ and a vector $\vec{A}(0)$ defined at one point on the curve (call it $\lambda = 0$). We define a procedure called parallel transport by defining a vector $\vec{A}(\lambda)$ along each point of the curve in such a way that $D\vec{A}^\mu /d\lambda = 0$:

$$\nabla_V \vec{A} = 0 \iff \text{parallel transport of } \vec{A} \text{ along } \vec{V}. \quad (23)$$

Over a small distance interval this procedure is equivalent to transporting the vector $\vec{A}$ along the curve in such a way that the vector remains parallel to itself with constant length: $A(\lambda + \Delta \lambda) = A(\lambda) + O(\Delta \lambda)^2$. In a locally flat coordinate system, with the connection vanishing at $x(\lambda)$, the components of the vector do not change as the vector is transported along the curve. If the space were globally flat and we used rectilinear coordinates (with vanishing connection everywhere), the components would not change at all no matter how the vector is transported. This is not the case in a curved space or in a flat space with curvilinear coordinates because in these cases the connection does not vanish everywhere.

### 3.3 Geodesics

Parallel transport can be used to define a special class of curves called geodesics. A geodesic curve is one that parallel-transports its own tangent vector $\vec{V} = d\vec{x}/d\lambda$, i.e., a curve that satisfies $\nabla_V \vec{V} = 0$. In other words, not only is $\vec{V}$ kept parallel to itself (with constant magnitude) along the curve, but locally the curve continues to point in the same direction all along the path. A geodesic is the natural extension of the definition of a “straight line” to a curved manifold. Using equations (22) and (23), we get a second-order differential equation for the coordinates of a geodesic curve:

$$\frac{DV^\mu}{d\lambda} = \frac{dV^\mu}{d\lambda} + \Gamma^\mu_{\alpha\beta} V^\alpha V^\beta = 0 \quad \text{for a geodesic}, \quad V^\mu \equiv \frac{dx^\mu}{d\lambda}. \quad (24)$$
Indeed, in locally flat coordinates (such that the connection vanishes at a point), this is the equation of a straight line. However, in a curved space the connection cannot be made to vanish everywhere. A well-known example of a geodesic in a curved space is a great circle on a sphere.

There are several technical points worth noting about geodesic curves. The first is that $\vec{V} \cdot \vec{V} = g(\vec{V}, \vec{V})$ is constant along a geodesic because $d\vec{V}/d\lambda = 0$ (eq. 24) and $\nabla_V g = 0$ (metric compatibility with gradient). Therefore, a geodesic may be classified by its tangent vector as being either timelike ($\vec{V} \cdot \vec{V} < 0$), spacelike ($\vec{V} \cdot \vec{V} > 0$) or null ($\vec{V} \cdot \vec{V} = 0$). The second point is that a nonlinear transformation of the parameter $\lambda$ will invalidate equation (24). In other words, if $x^\mu(\lambda)$ solves equation (24), $y^\mu(\lambda) \equiv x^h(\xi(\lambda))$ will not solve it unless $\xi = a\lambda + b$ for some constants $a$ and $b$. Only a special class of parameters, called affine parameters, can parametrize geodesic curves.

The affine parameter has a special interpretation for a non-null geodesic. We deduce this relation from the constancy along the geodesic of $\vec{V} \cdot \vec{V} = (d\vec{x} \cdot d\vec{x})/(d\lambda^2) \equiv a$, implying $ds = ad\lambda$ and therefore $s = a\lambda + b$ where $s$ is the path length ($ds^2 = g_{\mu\nu}dx^\mu dx^\nu$). For a non-null geodesic ($\vec{V} \cdot \vec{V} \neq 0$), all affine parameters are linear functions of path length (or proper time, if the geodesic is timelike). The linear scaling of path length amounts simply to the freedom to change units of length and to choose any point as $\lambda = 0$. Note that originally we imposed no constraints on the parameterization. However, the solutions of the geodesic equation automatically have $\lambda$ being an affine parameter. There is no fundamental reason to use an affine parameter; one could always take a solution of the geodesic equation and reparameterize it or eliminate the parameter altogether by replacing it with one of the coordinates along the geodesic. For example, for a timelike trajectory, $x^t(t)$ is a perfectly valid description and is equivalent to $x^\mu(\lambda)$. But the spatial components as functions of $t = x^0$ clearly do not satisfy the geodesic equation for $x^\mu(\lambda)$.

Another interesting point is that the total path length is stationary for a geodesic:

$$\delta \int_A^B ds = \delta \int_A^B \left| g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \right|^{1/2} d\lambda = 0$$  \hfill (25)
if $\lambda$ is an affine parameter. The $\delta$ refers to a variation of the integral arising from a variation of the curve, $x^\mu(\lambda) \to x^\mu(\lambda) + \delta x^\mu(\lambda)$, with fixed endpoints. The metric components are considered here to be functions of the coordinates. The variational principle is discussed in section 2 of the 8.962 notes “Hamiltonian Dynamics of Particle Motion,” where it is shown that stationary path length implies the geodesic equation (24) if the parameterization is affine. Equation (25) is invariant under reparameterization, so its stationary solutions are a broader class of functions than the solutions of equation (24). In general, the tangent vector of the stationary solutions are not normalized: $|\vec{V} \cdot \vec{V}|^{1/2} = Q(\lambda) \neq \text{constant}$, implying that $\lambda$ is not affine. It is easy to show that any stationary solution may be reparameterized, $\lambda \to \tau$ through $d\tau/d\lambda = Q(\lambda)$, and that the resulting curve $x^\mu(\lambda(\tau))$ obeys the geodesic equation with affine parameter $\tau$. This transformation replaces the unnormalized tangent vector $\vec{V}$ by $\vec{V}/Q(\lambda)$. For an affine parameterization, the tangent vector must always have constant length.

Equation (25) is a curved space generalization of the statement that a straight line is the shortest path between two points in flat space.

### 3.4 Integrals of motion and Killing vectors

Equation (24) is a set of four second-order nonlinear ordinary differential equations for the coordinates of a geodesic curve. One may ask whether the order of this system can be reduced by finding integrals of the motion. An integral, also called a conserved quantity, is a function of $x^\mu$ and $V^\mu = dx^\mu/d\lambda$ that is constant along any geodesic. At least one integral always exists: $\vec{V} \cdot \vec{V} = g_{\mu\nu}V^\mu V^\nu$. (For an affine parameterization, $\vec{V} \cdot \vec{V}$ is constant along the curve.) Are there others? Sometimes. One may show that equation (24) may be rewritten as an equation of motion for $V_\mu \equiv g_{\mu\nu}V^\nu$, yielding

$$
\frac{dV_\mu}{d\lambda} = \frac{1}{2}(\partial_\mu g_{\alpha\beta})V^\alpha V^\beta .
$$

Consequently, if all of the metric components are independent of some particular coordinate $x^\mu$, the corresponding component of the tangent one-form is constant along the geodesic. This result is very useful in reducing the amount of integration needed to construct geodesics for metrics with high symmetry. However, the condition $\partial_\mu g_{\alpha\beta} = 0$ is coordinate-dependent. There is an equivalent coordinate-free test for integrals, based on the existence of special vector fields $\vec{K}$ call Killing vectors. Killing vectors are, by definition, solutions of the differential equation

$$
\nabla_\mu K_\nu + \nabla_\nu K_\mu = 0 .
$$

(The Killing vector components are, of course, $K^\mu = g^{\mu\nu}K_\nu$.) The Killing equation (27) usually has no solutions, but for highly symmetric spacetime manifolds there may be
one or more solutions. It is a nice exercise to show that each Killing vector leads to the integral of motion

\[ \langle \vec{V}, \vec{K} \rangle = K^\mu V_\mu = \text{constant along a geodesic}. \]  

(28)

Note that if one of the basis vectors (for some basis) satisfies the Killing equation, then the corresponding component of the tangent one-form is an integral of motion. The test for integrals implied by equation (26) is a special case of the Killing vector test when the Killing vector is simply a coordinate basis vector.

The discussion here has focused on geodesics as curves. The notes “Hamiltonian Dynamics of Particle Motion” interprets them as worldlines for particles because, as we will see, a fundamental postulate of general relativity is that, in the absence of non-gravitational forces, particles move along geodesics. Given this fact, we are free to choose units of the affine parameter \( \lambda \) so that \( dx^\mu/d\lambda \) is the 4-momentum \( P^\mu \), normalized by \( \vec{P} \cdot \vec{P} = -m^2 \) for a particle of mass \( m \) (instead of \( dx^\mu/d\lambda = V^\mu, \vec{V} \cdot \vec{V} = -1 \)). Thus, the tangent vector, denoted \( \vec{V} \) above, is equivalent to the particle 4-momentum vector. The affine parameter \( \lambda \) then measures proper time divided by particle mass. Although one might fear this makes no sense for a massless particle, in fact it is the only way to affinely parameterize null geodesics because the proper time change \( d\tau \) vanishes along a null geodesic so \( dx^\mu/d\tau \) is undefined. For a massless particle, one takes the limit \( m \to 0 \) starting from the solution for a massive particle, with the result that \( d\lambda = d\tau/m \) is finite as \( m \to 0 \).