1 Basics of Differentiable Manifolds

A differentiable manifold $M$ is a topological space which is locally homeomorphic to an open subset of $\mathbb{R}^n$ and whose transition functions are smooth maps between these open subsets. Using the transition functions we can give coordinates around a point in the manifold. These coordinates will usually be denoted $\{x_1(\cdot), x_2(\cdot), \ldots, x_n(\cdot)\}$ so that $x(p) = (0, \ldots, 0)$.

The tangent space at each point $p \in M$, denoted $T_pM$, is an $n$-dimensional vector space spanned by the vectors $\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\}$. The inner product on this tangent space is given by the metric $g$ at the point $p$.

For any vectors $X, Y \in T_pM$ the inner product is denoted $g(X, Y, p)$. For any point $q$ in a neighborhood of $p$ the metric can be written as a matrix with coefficients $g_{ij}(q) = g\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}, q\right)$. We require the functions $g_{ij}$ to vary smoothly in the coordinate neighborhood on which they’re defined.

The metric can be used to measure the length of differentiable curves $\gamma : [a, b] \to M$ with velocity vector fields $\dot{\gamma}(t) \in T_{\gamma(t)}M$

$$L(\gamma) = \int_a^b \sqrt{g(\dot{\gamma}(t), \dot{\gamma}(t), \gamma(t))} \, dt$$

A natural question to ask is what curves minimize length between fixed endpoints $\gamma(a), \gamma(b)$ in $(M, g)$.

2 The Euler-Lagrange Equation of the Length Functional

For a given differentiable manifold the tangent bundle $TM$ is a manifold whose points are $(p, X)$ where $X \in T_pM$. If $M$ is $n$-dimensional $TM$ is $2n$-dimensional, accounting for both the dimension of $M$ and $T_pM$. A basic example is for the circle $M = S^1$ which has $T_pM = \mathbb{R}$ at every point, and $TM = S^1 \times \mathbb{R}$.

We define a Lagrangian on $M$ to be a smooth function on the tangent bundle $TM$. Thus $L(\gamma)$ is a Lagrangian as it only depends on the points $(\gamma(t), \dot{\gamma}(t)) \in TM$. We now seek to find the critical points of the length functional $L$ by computing its Euler-Lagrange Equation.

Let $\gamma_\epsilon : [a, b] \to M$ be a one-parameter family of curves with fixed endpoints $\gamma_\epsilon(a) = \gamma(a), \gamma_\epsilon(b) = \gamma(b), \gamma_0 = \gamma$, and $\gamma_\epsilon(t)$ $C^\infty$ in $\epsilon$ and $t$.

If $\gamma$ is a critical point of the length function $L$ then for any one-parameter family $\gamma_\epsilon$, $\frac{d}{d\epsilon} \bigg|_{\epsilon=0} L(\gamma_\epsilon) = 0$.

To compute the above derivative and derive the Euler-Lagrange equation we first define coordinates $\{x_1, \ldots, x_n\}$ on an open neighborhood $U_p \subset M$ of a point $p$ and coordinates $\{p_1, \ldots, p_n\}$ on the tangent space so that for an arbitrary Lagrangian $\int F(\gamma, \dot{\gamma})$ we can compute

$$\frac{d}{d\epsilon} \bigg|_{\epsilon=0} \int_a^b \left( \sum_i \frac{\partial F}{\partial x_i} \frac{\partial \gamma_i}{\partial \epsilon} \bigg|_{\epsilon=0} + \sum_i \frac{\partial F}{\partial p_i} \frac{\partial \gamma_i}{\partial \epsilon} \bigg|_{\epsilon=0} \right) \, dt$$

After integrating by parts
\[ = \int_a^b \sum_i \left. \frac{\partial \gamma_i^1}{\partial \epsilon} \right|_{\epsilon=0} \left( \frac{\partial F}{\partial x_i} - \frac{d}{dt} \frac{\partial F}{\partial p_i} \right) \, dt \]

Thus \( \frac{d}{d\epsilon} \bigg|_{\epsilon=0} F(\gamma_\epsilon) = 0 \) imples

\[ \sum_i \left. \frac{\partial F}{\partial x_i} - \frac{d}{dt} \frac{\partial F}{\partial p_i} \right|_{\epsilon=0} = 0 \]

which is the Euler-Lagrange equation of \( \int F(\dot{\gamma}, \gamma) \). Now we compute the Euler-Lagrange equation for the specific case of the energy functional \( E \) which has the same critical points as the length functional \( L \).

\[ E(\gamma, \dot{\gamma}) = g_{kl} \dot{\gamma}^k(t) \dot{\gamma}^l(t) \]

suppressing the summation over repeated indices. The Euler-Lagrange equation becomes:

\[
\sum_i \frac{\partial g_{kl}}{\partial x_i} \dot{\gamma}^i \dot{\gamma}^l = \frac{d}{dt} \left( \sum_i g_{il} \dot{\gamma}^l + \sum_i g_{kl} \dot{\gamma}^k \right) = \sum_i g_{il} \ddot{\gamma}^l + g_{kl} \ddot{\gamma}^k + \frac{\partial g_{il}}{\partial x_m} \dot{\gamma}^m \dot{\gamma}^l + \frac{\partial g_{kl}}{\partial x_m} \dot{\gamma}^m \dot{\gamma}^k
\]

By using the symmetry that \( g_{ij} = g_{ji} \), reindexing and combining the \( \ddot{\gamma} \) terms, and rewriting \( \frac{\partial g_{ij}}{\partial x_k} = g_{kj,i} \) we get

\[
2 \sum_i g_{il} \ddot{\gamma}^l = \sum_i (g_{mk,i} - g_{im,k} - g_{ik,m}) \ddot{\gamma}^k \dot{\gamma}^m
\]

Thus a curve \( \gamma \) is a critical point of the energy functional \( E \), and thus a critical point of the length functional \( L \), if it is a solution to this differential equation. Next we will show that this is the same geodesic equation that arises in differential geometry.

### 3 Geodesics from a Riemannian Geometry Viewpoint

Let \((M, g)\) be a Riemannian manifold. The geodesic equation can also be derived by considering what it means for a curve to have zero acceleration in a manifold. The vector field \( \dot{\gamma} \) is a well defined first derivative for a curve \( \gamma \), but in order to have a well defined second derivative we must define a connection on \( M \) which will allow us to differentiate one vector field along another. An affine connection on \( M \) is a map \( \nabla : TM \times TM \rightarrow TM \) satisfying two conditions.

1) \( \nabla_{aX + bY} Z = a\nabla_X Z + b\nabla_Y Z \)

2) \( \nabla_X (aY) = a\nabla_X Y + da(X)Y = a\nabla_X Y + X(a)Y \)

for \( X, Y, Z \) sections of \( TM \) and \( a, b \) smooth functions on \( M \)

The Levi-Civita connection or Riemannian connection is an affine connection with the added properties of being symmetric and compatible with the metric.
3) \nabla_X Y - \nabla_Y X = [X, Y] 
4) Xg(X, Y) = g(\nabla_X Y, X) + g(Y, \nabla_X Z)

With the Levi-Civita connection we can define a vector field $X$ to be parallel along a curve $\gamma$ if it satisfies $\nabla_{\dot{\gamma}}(t)X(t) = 0$. Because of the compatibility of $\nabla$ with the metric $g$, the inner product $g(X(t), Y(t))$ of two vector fields which are parallel along $\gamma$ is constant.

In order to facilitate computations with the connection we define the action of $\nabla$ on the basis for the tangent space in local coordinates $\{\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\}$ which we will furthermore denote $\{\partial_1, \ldots, \partial_n\}$.

$$\nabla_{\partial_i} \partial_j = \Gamma^k_{ij} \partial_k$$
$$\Gamma_{ijk} = g_{kl} \Gamma^l_{ij}$$

$\Gamma^k_{ij}$ are called Christoffel symbols of the 2nd kind and $\Gamma_{ijk}$ are called Christoffel symbols of the 1st kind.

**Lemma 3.1.** The Levi-Civita connection is unique.

**Proof.** We will use the metric compatibility and apply this to basis vector fields.

$$\partial_k g(\partial_i, \partial_j) = g_{i,j,k}$$
$$= g(\nabla_{\partial_k} \partial_i, \partial_j) + g(\partial_i, \nabla_{\partial_k} \partial_j)$$
$$= g(\Gamma^l_{ki} \partial_i, \partial_j) + g(\partial_i, \Gamma^l_{kj} \partial_l)$$
$$= \Gamma^l_{ki} g_{lj} + \Gamma^l_{kj} g_{il}$$
$$= \Gamma_{kij} + \Gamma_{kji}$$

By the symmetry of the connection

$$0 = [\partial_i, \partial_j]$$
$$= \nabla_{\partial_i} \partial_j - \nabla_{\partial_j} \partial_i$$
$$= (\Gamma^k_{ij} - \Gamma^k_{ji}) \partial_k$$

$$\implies \Gamma^k_{ij} = \Gamma^k_{ji}$$

and summing covariant derivatives of the metric

$$g_{i,j,k} + g_{j,k,i} - g_{k,i,j} = \Gamma_{kij} + \Gamma_{kji} + \Gamma_{ijk} + \Gamma_{ikj} - \Gamma_{jki} - \Gamma_{jik}$$
$$= 2\Gamma_{ijk}$$

So the Christoffel symbols are determined by the metric and the connection is unique. \qed

Now we define geodesics from a Riemannian geometry viewpoint as curves $\gamma : [a, b] \rightarrow M$ which have zero acceleration, or equivalently, parallel velocity vector fields $\dot{\gamma}$. Thus, the geodesic equation becomes $\nabla_{\dot{\gamma}} \dot{\gamma}$. In local coordinates, written in terms of the Christoffel symbols of the Riemmanian connection this becomes:

$$\gamma(t) = (x^1(t), \ldots, x^n(t))$$
$$\dot{\gamma}(t) = (\dot{x}^1(t), \ldots, \dot{x}^n(t))$$
\[
\n\nabla_\gamma \dot{\gamma} = \nabla_{\dot{x}^i \partial_i \dot{x}^j \partial_j} = \dot{x}^i \dot{x}^j \Gamma^k_{ij} \partial_k + \ddot{x}^k \partial_k = (\dot{x}^i \dot{x}^j \Gamma^k_{ij} + \ddot{x}^k) \partial_k = 0
\]

So the geodesic equation is
\[
\dot{\gamma}^i \dot{\gamma}^j \Gamma^k_{ij} + \ddot{\gamma}^k = 0
\]

To show this is the same as the geodesic equation which we arrived at by variational methods
\[
0 = 2 \dot{\gamma}^i \dot{\gamma}^j \Gamma^k_{ij} + 2 \ddot{\gamma}^k = 2 g_{lk} \ddot{\gamma}^k + 2 \Gamma^k_{ij} \dot{\gamma}^i \dot{\gamma}^j = 2 g_{lk} \ddot{\gamma}^k - (g_{li,j} - g_{ij,l} - g_{jl,i}) \dot{\gamma}^i \dot{\gamma}^j
\]

which agrees with our variational method up to a change of indices.

4 Jacobi’s Theorem

Next we seek to know when a critical point of the length functional is in fact a minimizer. All geodesics exist and are length minimizing in a small ball around any point. We describe these geodesics by the exponential map. Let \( \gamma \) be a geodesic in \( M \) such that \( \gamma(0) = p, \dot{\gamma}(0) = X \). Then \( \gamma \) has constant speed \( \sqrt{g(X,X)} \). The exponential map \( \exp_p : T_p M \to M \) is defined:
\[
\exp_p(X) = \gamma(1) \quad \exp_p(tX) = \gamma(t)
\]

To find when a geodesic is length minimizing, we must understand the critical points of \( \exp_p \), so we need to know when \( d\exp_p \) has a nontrivial kernel. Thus we define a point \( q = \exp_p(tA) \) of a geodesic \( \gamma \) to be conjugate to \( p \) along \( \gamma \) if it is a critical value of \( \exp_p \). With this definition we can state the main theorem regarding length minimizing geodesics.

Theorem 4.1. If \( \gamma \) is a minimizing geodesic then none of its interior points are conjugate points.

Before we prove this theorem we must relate conjugate points to variations of a geodesic. Let \( \gamma_\epsilon(t) \) be a variation such that \( \gamma_0(t) = \gamma(t), \gamma_\epsilon(0) = \gamma(0), \) and \( \gamma_\epsilon(1) = \gamma(1) \). Our strategy is to assume the existence of a conjugate point along \( \gamma \) and show that this implies the existence of a variation \( \gamma_\epsilon \) such that \( \frac{\partial^2}{\partial \epsilon^2} L(\gamma_\epsilon) < 0 \). Such a variation would show \( \gamma \) to in fact be a locally maximizing geodesic, thus proving the theorem.
Decomposing a vector field $X$ choose such we can rewrite as a map $R\colon TM \times TM \to TM$ defined as

$$R(X,Y)Z = (\nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]})Z$$

which we can rewrite as a map $R: TM \times TM \times TM \times TM \to \mathbb{R}$ defined as

$$R(X,Y,Z,W) = g(R(X,Y)Z,W)$$

Now the second term of the integral can be rewritten as

$$\frac{1}{|\gamma|^3} \left[ (\nabla_X \nabla_\gamma, \gamma) + |\nabla X \gamma|^2 \right] = \frac{1}{|\gamma|^3} \left[ -\langle \nabla_\gamma \nabla_X \gamma, \gamma \rangle + \langle R(\gamma)X, \gamma \rangle + |\nabla_X \gamma|^2 \right]

= -\frac{d}{dt} \langle \nabla_X \gamma, \gamma \rangle + \langle R(\gamma)X, \gamma \rangle + |\nabla_X \gamma|^2$$

And if we rescale $\gamma$ to have constant unit speed the second variational formula for the length functional becomes

$$\frac{\partial^2}{\partial \epsilon^2} \bigg|_{\epsilon=0} L(\gamma_\epsilon) = \int_a^b \left( \frac{d}{dt} \langle X, \gamma \rangle \right)^2 + \langle R(\gamma)X, \gamma \rangle + |\nabla_X \gamma|^2 \, dt$$

Now we consider normal fields $X$ such that $\langle X(t), \dot{\gamma}(t) \rangle = 0$ for any $t$. If $X$ is normal along $\gamma$ then the second variational formula becomes

$$I_\gamma (X) = \frac{\partial^2}{\partial \epsilon^2} \bigg|_{\epsilon=0} L(\gamma_\epsilon) = \int_a^b R(\gamma, \dot{\gamma}, X, \dot{\gamma}) + |\nabla_X \dot{\gamma}|^2 \, dt$$

Now we will seek minimizers of $I$ and prove they have negative values when $\gamma$ contains conjugate points. We consider $I_\gamma$ over all vector fields along $\gamma$, which are sections of $\gamma^* TM$, vanishing at 0 and 1. $\gamma^* TM \simeq [0,1] \times \mathbb{R}^n$ and we can make this explicit by choosing a frame for $T_p \mathbb{M}$ which varies along $\gamma$. First choose $\{E_1, \ldots, E_n\}$, an orthonormal frame for $T_{\gamma(0)} \mathbb{M}$ and extend $E_i(t) \in T_{\gamma(t)} \mathbb{M}$ by the solution to $\nabla_{\dot{\gamma}} E_i(t) = 0, E_i(0) = E_i$. It can be shown that $E_i(t)$ is an orthonormal frame at each point of $\gamma$.

Decomposing a vector field $X$ along $\gamma$ as $X = \sum q_i(t) E_i(t)$, and $\dot{q}(t) = \sum a_i E_i(t)$. Then $\nabla_\gamma X = \sum \dot{q}_i(t)^2 E_i(t)$.

Now the second variational formula for $X$ along $\gamma$ is

$$I_\gamma X = \int_a^b \sum q_i^2 + R_{ijkl} q_i a_j q_k a_l \, dt$$
where $R_{ijkl} = R(E_i, E_j, E_k, E_l)$.

Using the Euler-Lagrange equation for functionals of this type that we derived previously we find that the Euler-Lagrange equation for $I$ is

$$2\ddot{q}_i = -R_{ijkl} a_j q_k a_l - R_{mjkl} q_m a_j a_l$$

And using that $\sum \ddot{q}_i E_i = \nabla \dot{\gamma} \nabla \dot{\gamma} X$ we get the Jacobi Equation

$$\nabla \dot{\gamma} \nabla \dot{\gamma} X + R(\dot{\gamma}, X) \dot{\gamma} = 0$$

When $X$ satisfies this equation it is called a Jacobi field. We seek to find a normal Jacobi field $X$ with $I(X) < 0$ along a geodesic $\gamma$ with a conjugate point, which will provide the contradiction needed to prove Theorem 4.1. But first we must show that any Jacobi field $X$ with $X(0) = 0$ and $X(1) = 0$ can be obtained from a variation so that showing $I(X) < 0$ is sufficient for proving the theorem which concerns variations rather than Jacobi fields.

**Proposition 4.2.** There is a bijection between variations $\gamma_t(t)$ which are composed of geodesics and have a fixed start point $\gamma(0)$ and Jacobi fields $X(t)$ with $X(0) = 0$.

**Proof.** First we’ll show that a variation of geodesics with a fixed start point induces a Jacobi field $X$ with $X(0) = 0$. Such a variation can be defined as $\gamma_t(t) = \exp_{p} t A(\epsilon)$ with $A(\epsilon) \in T_p M$ and $A(0) = \dot{\gamma}(0)$. The induced Jacobi field is $X(t) = \gamma_t(t) = \frac{d\exp_{p} t A(\epsilon)}{dt} (t A'(0))$. Clearly $X(0) = 0$.

Now we’ll show that $\gamma_t'(t)$ is a Jacobi field.

$$\nabla \gamma \nabla \gamma' + R(\gamma', \gamma') \gamma = 0$$

$$\iff \nabla \gamma \nabla \gamma' + R(\gamma', \gamma') \gamma = 0$$

$$\iff \nabla \gamma \nabla \gamma' + R(\gamma', \gamma') \gamma = 0$$

$$\iff \nabla \gamma \nabla \gamma' = 0$$

But the last line being equal to 0 is exactly the geodesic equation, so $\gamma_t'(t)$ is a Jacobi field.

Next we’ll show that a Jacobi field $X$ with $X(0) = 0$ gives rise to a variation of geodesics with a fixed start point. But first we’ll give an argument for why for each variation with fixed start point $\phi(t, \epsilon) = \gamma_t(t) = \exp_{p} t A(\epsilon)$ which gives rise to a variational field $X(t) = \frac{\partial \phi}{\partial \epsilon}(t, 0)$ has its covariant derivative $\nabla \gamma X(0) = A'(0)$. This relies on the fact that

$$\nabla \gamma X(t) = \nabla \gamma \left( \frac{\partial \phi}{\partial \epsilon} \bigg|_{\epsilon=0} \right) = (\nabla \gamma \gamma_t'(t)) |_{\epsilon=0}$$

Where in the last equality I used the relation $\nabla \gamma \frac{\partial \phi}{\partial \epsilon} = \nabla \gamma \frac{\partial \phi}{\partial x}$ between partial and covariant derivatives on the surface which is the image of $\phi$. To see why this relation is true choose a local coordinate $x$ for $M$ near a point on the image of $\phi$ and compute

$$\nabla \gamma \left( \frac{\partial \phi}{\partial \epsilon} \right) = \nabla \gamma \left( \frac{\partial x^i}{\partial \epsilon} \frac{\partial}{\partial x^i} \right)$$

$$= \frac{\partial^2 x^i}{\partial \epsilon \partial \epsilon} \frac{\partial}{\partial x^i} + \frac{\partial x^i}{\partial \epsilon} \nabla \gamma \frac{\partial x^j}{\partial x^i} \frac{\partial}{\partial x^j}$$

$$= \frac{\partial^2 x^i}{\partial \epsilon \partial \epsilon} \frac{\partial}{\partial x^i} + \frac{\partial x^i}{\partial \epsilon} \frac{\partial x^j}{\partial \epsilon} \nabla \gamma \frac{\partial}{\partial x^j}$$

And the symmetry of the connection gives $\nabla \frac{\partial x^i}{\partial x^j} = \nabla \frac{\partial x^i}{\partial x^j}$ so a computation of $\nabla \gamma \frac{\partial \phi}{\partial x}$ gives the same result.
Thus $\nabla_\gamma X(t) = (\nabla_\gamma \dot{\gamma}_e(t))|_{\epsilon = 0}$ and $\nabla_\gamma X(0) = (\nabla_\gamma A(\epsilon))|_{\epsilon = 0}$. But since at $t = 0$ the the curve $\epsilon \mapsto \gamma_\epsilon(0)$ is constant, the covariant derivative is just a partial derivative with respect to $\epsilon$ and $\nabla_\gamma X(0) = A'(0)$. Now that we’ve established that $\nabla_\gamma X(0) = A'(0)$, we continue the proof by considering an arbitrary Jacobi field with initial value 0.

Given such an $X$ let $B = \nabla_\gamma X(0)$. Now define $A(\epsilon) = \dot{\gamma}(0) + \epsilon B$ and consider the variation $\gamma_\epsilon(t) = \text{exp}_{p_0} tA(\epsilon)$. The induced Jacobi field is $Y(t) = \dot{\gamma}_\epsilon(t) = d\text{exp}_{p_0}|_{tA(\epsilon)}(tB)$. But $Y(0) = X(0) = 0$ and $\nabla_\gamma Y(0) = B = \nabla_\gamma X(0)$ by the argument given above, so by the uniqueness of ordinary differential equations $X(t) = Y(t)$ and $X(t)$ is induced by a variation of geodesics with a fixed start point. \hfill \QED

Before proving the existence of a Jacobi field with $I(X) < 0$, we will first prove a helpful proposition.

**Proposition 4.3.** If $X$ is a Jacobi field such that $\langle X(0), \dot{\gamma}(0) \rangle = \langle X(1), \dot{\gamma}(1) \rangle = 0$ then $X$ is a normal Jacobi field, and if furthermore $X(0) = X(1) = 0$ then $I(X) = 0$.

**Proof.** To show that $X$ is normal we must show $\langle \dot{\gamma}, X \rangle = 0$. First we note

$$\frac{d}{dt} \langle \dot{\gamma}, X \rangle = \langle \ddot{\gamma}, \nabla_\gamma X \rangle$$

And taking a further derivative

$$\frac{d^2}{dt^2} \langle \dot{\gamma}, \nabla_\gamma X \rangle = \langle \dddot{\gamma}, \nabla_\gamma \nabla_\gamma X \rangle = -\langle \dot{\gamma}, R(\dot{\gamma}, \dot{\gamma}) \dot{\gamma} \rangle = -R(\dot{\gamma}, \dot{\gamma}, X, \dot{\gamma}) = 0$$

So $\langle \dot{\gamma}, X \rangle = at + b$ for $a, b \in \mathbb{R}$ and by using the conditions on the endpoints we can show that $\langle \dot{\gamma}, X \rangle = 0$, so $X$ is normal. Thus, the equation for $I(\dot{\gamma})$ reduces to:

$$I(X) = \int_0^1 |\nabla_\gamma X|^2 - R(X, \dot{\gamma}, X, \dot{\gamma}) \, dt$$

$$= \int_0^1 \langle \nabla_\gamma X, \nabla_\gamma X \rangle + \langle X, \nabla_\gamma \nabla_\gamma X \rangle \, dt$$

$$= \int_0^1 \frac{d}{dt} \langle X, \nabla_\gamma X \rangle \, dt$$

$$= \langle X, \nabla_\gamma X \rangle \big|_0^1 = 0$$

as the conditions on the endpoints are the same. \hfill \QED

Now we have all of the tools necessary to prove Theorem 4.1 which states that if $\gamma$ is a minimizing geodesic then none of its interior points are conjugate points.

**Proof.** We will prove the theorem by contradiction, assuming there exists a conjugate point at $\gamma(t_0)$ and showing that $\gamma$ cannot be length minimizing. To this end we will show there exists a normal Jacobi field $Y$ along $\gamma$, which by Proposition 4.2 is the variational field of some variation $\gamma_\epsilon(t)$, such that the second variation of $Y$, $I(Y) < 0$ which would imply $\gamma$ is locally length maximizing.

First we show the existence of a normal Jacobi field such that $X(0) = 0$ and $X(t_0) = 0$, where $\gamma(t_0)$ is the conjugate point to $\gamma(0)$. Since $\gamma(t_0)$ is conjugate, there exists $A(\epsilon) \in T_{p_0}M$ such that $A'(0) \in \ker \left( \text{exp}_{p_0} |_{t_0A} \right)$, and defining $X(t) = d\text{exp}_{p_0}|_{tA(\epsilon)}(tA'(t))$ gives the desired Jacobi field. And by Proposition 4.3 it is normal on $[0, t_0]$ and has $I(X) = 0$ on that interval.

Around every point in a Riemannian manifold the exponential map is a diffeomorphism from a small ball in the tangent space to a small normal neighborhood in the manifold. Thus there exists $\delta > 0$ such that
the exponential map is a diffeomorphism on $B(0, \delta) \subset T_{\gamma(t_0)}M$. Thus $exp_{\gamma(t_0)}$ cannot have any critical points in the interval, so none of the points on $\gamma|_{[t_0-\delta,t_0+\delta]}$ are conjugate to $\gamma(t_0 - \delta)$. Now we seek to prove the existence of a Jacobi field $Z(t)$ on $[t_0 - \delta, t_0 + \delta]$ such that $Z(t_0 - \delta) = X(t_0\delta)$ and $Z(t_0 + \delta) = 0$. We first note that since the Jacobi equation is a system of $n$ second-order ordinary differential equations, its solution set is $2n$-dimensional. Next, we consider the linear map $\alpha : X \mapsto (X(t_0 - \delta), X(t_0 + \delta))$ from the $2n$-dimensional space of Jacobi fields to the $2n$-dimensional space $T_{\gamma(t_0-\delta)}M \oplus T_{\gamma(t_0+\delta)}M$. It is injective as it is a homomorphism of vector spaces and the only Jacobi field which has both endpoints 0 is the unique solution $X(t) = 0$ of the Jacobi equation with those initial conditions. But since the dimension of the domain and codomain are both $2n$ it is surjective as well. Thus there exists a Jacobi field $Z$ with the prescribed endpoints. By Proposition 4.3, $Z$ is normal.

Before we define our final Jacobi field which will satisfy $I(Y) < 0$, we will first show that the Jacobi field $Z$ defined above is minimal in the sense that $I(Z) < I(W)$ where $W$ is any other piecewise smooth field along $\gamma$. Since $Z$ is normal, it can be written as an $\mathbb{R}$-linear combination of of $\{W_1, \ldots, W_{n-1}\}$, a basis for the subspace of normal Jacobi fields. A computation shows that for a smooth field $W = f^iW_i$, where now $f^i$ are smooth functions, $I$ can be written.

$$I(W) = \int_0^1 \left| f^iW_i \right|^2 \, dt + f^i(1)f^j(1) \langle \nabla_{\gamma}W_i, W_j \rangle_{t=1}$$

And since $Z$ is normal, the $f^i$s are constant, and the formula reduces to.

$$I(Z) = f^i(1)f^j(1) \langle \nabla_{\gamma}W_i, W_j \rangle_{t=t_0+\delta}$$

Thus for any smooth field $W$ which agrees with $Z$ on $t_0 - \delta$ and $t_0 + \delta$ we have

$$I(W) - I(Z) = \int_{t_0 - \delta}^{t_0 + \delta} \left| f^iW_i \right|^2 \, dt \geq 0$$

And so $Z$ is strictly minimal amongst smooth fields with the same endpoints.

Now define the Jacobi field $Y$ on all of $\gamma$ by

$$Y(t) = \begin{cases} X(t) & t \in [0, t_0 - \delta] \\ Z(t) & t \in [t_0 - \delta, t_0 + \delta] \\ 0 & t \in [t_0 + \delta, 1] \end{cases}$$

And we compute

$$I_0^1(Y) = I_0^{t_0+\delta}(Y) = I_0^{t_0-\delta}(X) + I_0^{t_0+\delta}(Z) = I_0^{t_0}(X) - I_0^{t_0-\delta}(X) + I_0^{t_0+\delta}(Z) = I_{t_0-\delta}^{t_0+\delta}(Z) - I_{t_0-\delta}^{t_0}(X)$$

And $Z$ is minimal by the argument above, so for a field defined

$$W(t) = \begin{cases} X(t) & t \in [t_0 - \delta, t_0] \\ 0 & t \in [t_0, t_0 + \delta] \end{cases}$$

$W$ and $Z$ have the same endpoints, so $I(Z) < I(W)$, and since $I_{t_0-\delta,t_0}^{t_0}(X) = I_{t_0-\delta}^{t_0}(W)$, it follows that $I_0^1(Y) < 0$, contradicting the fact that $\gamma$ was length minimizing.

□
Exercise 11.3 Prove that parallel translation with respect to a Riemannian connection is an isometric mapping of tangent spaces.

Proof. Let $X(t)$ and $Y(t)$ be parallel vector fields along $\gamma$ so that $\nabla_{\dot{\gamma}(t)} X(t) = \nabla_{\dot{\gamma}(t)} Y(t) = 0$. To prove parallel translation is an isometry we must show $\langle X(t), Y(t) \rangle$ is constant.

$$\frac{d}{dt} \langle X(t), Y(t) \rangle = \langle \nabla_{\dot{\gamma}(t)} X(t), Y(t) \rangle + \langle X(t), \nabla_{\dot{\gamma}(t)} Y(t) \rangle = 0$$

Thus parallel translation is an isometry. □

Exercise 11.4 Prove that the matrix for the Lagrangian $L$ on $TM$ with local coordinates $(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ given by

$$\left[ \begin{array}{c} \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \end{array} \right]$$

does not depend on the choice of local coordinates.

Proof. Let $\phi$ be a change of coordinates such that $\phi(q) = p$. This induces a change of coordinates on the tangent space by $\dot{p} = (d\phi)^{-1} \dot{q}$ and more specifically $\dot{p}_i = (d\phi^{-1})_{ji} \dot{q}_j$. Thus $\frac{\partial \dot{p}_i}{\partial q^j} = (d\phi^{-1})_{ji}$ and $\frac{\partial^2 \dot{p}_i}{\partial q^j \partial q^k} = 0$. Under the change of coordinates $L$ changes as

$$L(q, \dot{q}) = L(\phi^{-1}(p), d\phi(\dot{p})) = G(p, \dot{p})$$

And the partial derivatives of $L$ change as

$$\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial G}{\partial \dot{p}_k} \frac{\partial \dot{p}_k}{\partial \dot{q}_i}$$
Proof. Consider a reparameterization \( s = g(t) \) for the geodesic \( \gamma(s) \). Now put the reparameterized curve \( \gamma(g(t)) \) into the geodesic equation.

\[
\frac{d^2}{dt^2} \gamma^i(g(t)) + \Gamma^i_{kj}(\gamma(g(t))) \frac{d}{dt} \gamma^k(t) \frac{d}{dt} \gamma^j(t) = \hat{\gamma}^i(g(t))(g'(t))^2 + \hat{\gamma}^i(g(t))g''(t) + \Gamma^i_{kj}(\gamma(g(t))\hat{\gamma}^j(g(t))\hat{\gamma}^k(g(t))(g'(t))^2 = \hat{\gamma}^i(s)g''(t) + \left[ \hat{\gamma}^i(s) + \Gamma^i_{kj}(\gamma(s)\hat{\gamma}^j(s)\hat{\gamma}^k(s)) \right] g''(t) = \hat{\gamma}^i(s)g''(t)
\]

Where the last equality is because \( \gamma(s) \) is a geodesic. Since \( \hat{\gamma}^i(s) \neq 0 \), \( \gamma(g(t)) \) is a geodesic iff \( g''(t) = 0 \) iff \( g(t) = at + b \) for \( a, b \in \mathbb{R} \) and \( a \neq 0 \).

Exercise 27.1 Show that for any piecewise smooth field \( X \) on the geodesic \( \gamma \), the inequality

\[
I^0_1(X^\perp) \leq I^0_0(X)
\]

holds, where \( X^\perp \) is the normal component of the field \( X \) (i.e., \( X(t) - X^\perp(t) \) is collinear to the vector \( \dot{\gamma}(t) \) for every \( t \)).

Proof. Assuming \( \gamma : [a, b] \to M \) is a unit speed geodesic, let \( \{ \gamma(t) = A_1, A_2, \ldots, A_n \} \) be an orthonormal basis for \( T_{\gamma(t)}M \). Parallel transporting each of these vectors along \( \gamma \) based on the unique solution to the differential equation \( \nabla_{\dot{\gamma}(t)} X_i(t) = 0, X_i(0) = A_i \) yields an orthonormal frame along \( \gamma \) with \( X_1(t) = \dot{\gamma}(t) \).

Thus, any smooth vector field \( X \) along \( \gamma \) can be written as \( X = \sum_{i=1}^{n} f^i X_i \) for \( f^i \) smooth functions. Then

\[
X^\perp = \sum_{i=2}^{n} f^i X_i
\]

as \( X^\perp \) has no component parallel to \( X_1 = \dot{\gamma} \). On smooth vector fields \( I(x) \) is given by

\[
\int_a^b \left[ (\nabla_{\dot{\gamma}} X, \nabla_{\dot{\gamma}} X) - R(X, \dot{\gamma}, \dot{\gamma}) - (\nabla_{\dot{\gamma}} X, \dot{\gamma}) \right] dt
\]

Examining the first term in the integral we see
\[ \langle \nabla_X X, \nabla_Y Y \rangle = \langle \nabla_{f^i X, i}, \nabla_{f^j X, j} \rangle \]
\[ = \langle f^i X, f^i \nabla_X X, f^j X, f^j \nabla_X X \rangle \]
\[ = \sum |f^i|^2 \]

And comparing \( X \) to \( X^\perp \) we see
\[ \int_a^b \langle \nabla_X X, \nabla_X X \rangle \, dt - \int_a^b \langle \nabla_X X^\perp, \nabla_X X^\perp \rangle \, dt = \int_a^b |f^1|^2 \, dt \]

Examining the second term in the integral, since \( R(X, Y, Z, Z) = 0 \) by the symmetries of the curvature tensor, it follows that \( R(f^i X, i, f^1 X, 1) = R(f^i X, i, f^1 X, 1, X_1) = 0 \). Thus \( R(X, i, X, i) = R(X^\perp, i, X^\perp, i) \) and their difference contributes nothing to the integral.

Examining the third term in the integral we see
\[ \langle \nabla_X X, i \rangle^2 = \langle f^i X, i \rangle^2 \]
\[ = |f^1|^2 \]

And comparing \( X \) to \( X^\perp \) we see
\[ - \int_a^b \langle \nabla_X X, i \rangle^2 \, dt - \left( - \int_a^b \langle \nabla_X X^\perp, i \rangle^2 \, dt \right) = - |f^1|^2 \]

So together
\[ I(X) - I(X^\perp) = \int_a^b |f^1|^2 - |f^1|^2 \, dt = 0 \]

\( \Box \)

**Exercise 5** Prove that the Jacobi field \( Z \) from the proof of Theorem 4.1 from the notes on Lectures 1-4 is normal.

**Proof.** The Jacobi field \( Z(t) \), which is defined on \([t_0 - \delta, t_0 + \delta]\) has the property that \( Z(t_0 - \delta) = X(t_0 - \delta) \) where \( X \) is a normal Jacobi field on \([0, t_0]\), and \( Z(t_0 + \delta) = 0 \). Thus \( \langle Z(t_0 - \delta), \dot{Z}(t_0 - \delta) \rangle = \langle X(t_0 - \delta), \dot{Z}(t_0 - \delta) \rangle = 0 \) as \( X \) was already shown to be normal. Likewise \( \langle Z(t_0 + \delta), \dot{Z}(t_0 + \delta) \rangle = (0, \dot{Z}(t_0 + \delta)) = 0 \).

By Proposition 4.3 \( Z \) must be minimal. \( \Box \)

**Exercise 6** Prove that the normal Jacobi field \( Z \) from the proof of Theorem 4.1 from the note on Lectures 1-4 is minimal.

Before beginning the proof of the minimality of \( Z \) I will state and prove a lemma.

**Lemma 0.1.** For any smooth normal vector field \( W = f^i W_i \) written in terms of an orthonormal basis \( \{W_1, \ldots, W_{n-1}\} \) for the space of normal Jacobi fields with a prescribed final value. If \( W(1) = 0 \) we can write \( I(W) \) as
\[ I(W) = \int_0^1 \left| \dot{f} \right| W_i^2 \, dt - f^i(0) f^j(0) \langle \nabla_{\dot{f} W_i, W_j} \rangle_{t=0} \]
Proof. The formula for $I$ of a normal vector field is as shown in the notes

$$I(W) = \int_{0}^{1} \left[ \langle \nabla \dot{\gamma} W, \nabla \dot{\gamma} W \rangle - R(W, \dot{\gamma}, W, \dot{\gamma}) \right] dt$$

Computing the first term

$$\langle \nabla \dot{\gamma} W, \nabla \dot{\gamma} W \rangle = \left\langle \dot{f}^i W_i + f^i \nabla \dot{\gamma} W_i, \dot{f}^j W_j + f^j \nabla \dot{\gamma} W_j \right\rangle$$

$$= \left| \dot{f}^i W_i \right|^2 + 2 \dot{f}^i f^j \langle W_i, \nabla \dot{\gamma} W_j \rangle + f^i f^j \langle \nabla \dot{\gamma} W_i, \nabla \dot{\gamma} W_j \rangle$$

Computing the second term

$$R(W, \dot{\gamma}, W, \dot{\gamma}) = f^i f^j \langle R(W_i, \dot{\gamma}), \dot{\gamma} W_j \rangle$$

$$= -f^i f^j \langle \nabla \dot{\gamma} \nabla \dot{\gamma} W_i, W_j \rangle$$

Where the last equality is because of the Jacobi equation applied to $W_i$. Noting

$$\frac{d}{dt} \langle \nabla \dot{\gamma} W_i, W_j \rangle = \langle \nabla \dot{\gamma} \nabla \dot{\gamma} W_i, W_j \rangle + \langle \nabla \dot{\gamma} W_i, \nabla \dot{\gamma} W_j \rangle$$

We see that

$$\langle \nabla \dot{\gamma} W, \nabla \dot{\gamma} W \rangle - R(W, \dot{\gamma}, W, \dot{\gamma}) = \left| \dot{f}^i W_i \right|^2 + 2 \dot{f}^i f^j \langle W_i, \nabla \dot{\gamma} W_j \rangle + f^i f^j \frac{d}{dt} \langle \nabla \dot{\gamma} W_i, W_j \rangle$$

Also we have

$$\frac{d}{dt} \langle f^i \nabla \dot{\gamma} W_i, f^j W_j \rangle = \dot{f}^i f^j \langle \nabla \dot{\gamma} W_i, W_j \rangle + f^j \dot{f}^i \langle \nabla \dot{\gamma} W_i, W_j \rangle + f^i f^j \frac{d}{dt} \langle \nabla \dot{\gamma} W_i, W_j \rangle$$

So we see that,

$$\langle \nabla \dot{\gamma} W, \nabla \dot{\gamma} W \rangle - R(W, \dot{\gamma}, W, \dot{\gamma}) = \left| \dot{f}^i W_i \right|^2 + f^i f^j [(W_i, \nabla \dot{\gamma} W_j) + \langle \nabla \dot{\gamma} W_i, W_j \rangle] + \frac{d}{dt} \langle f^i \nabla \dot{\gamma} W_i, f^j W_j \rangle$$

Now to show that the middle term of the right hand side is zero

$$\frac{d}{dt} \langle \nabla \dot{\gamma} W_i, W_j \rangle = \langle \nabla \dot{\gamma} \nabla \dot{\gamma} W_i, W_j \rangle + \langle \nabla \dot{\gamma} W_i, \nabla \dot{\gamma} W_j \rangle$$

$$= \langle \nabla \dot{\gamma} W_i, \nabla \dot{\gamma} W_j \rangle - R(W_j, \dot{\gamma}, W_i, \dot{\gamma})$$

And likewise

$$\frac{d}{dt} \langle W_i, \nabla \dot{\gamma} W_j \rangle = \langle \nabla \dot{\gamma} W_i, \nabla \dot{\gamma} W_j \rangle - R(W_j, \dot{\gamma}, W_i, \dot{\gamma})$$

And examining the middle term of the previous expansion

$$\frac{d}{dt} \left[ \langle W_i, \nabla \dot{\gamma} W_j \rangle - \langle W_i, W_j \rangle \right] = 0$$
So it is equal to a constant, but since $W_i(1) = W_j(1) = 0$ the whole term is equal to 0. And we can simplify

$$\langle \nabla \dot{\gamma} W, \nabla \dot{\gamma} W \rangle - R(W, \dot{\gamma}, W \dot{\gamma}) = |\dot{f}^i W_i|^2 + \frac{d}{dt} \langle f^i \nabla \dot{\gamma} W_i, f^j W_j \rangle$$

And integrating from 0 to 1 gives the desired result

Now we return to the proof of the minimality of $Z$.

**Proof.** $Z$ is in particular a smooth normal field, so it can be written in terms of an orthonormal basis $\{W_1, \ldots, W_{n-1}\}$ for the space of normal Jacobi fields with prescribed initial value $Z(t_0 - \delta) = X(t_0 - \delta)$. Let $Z(t) = f^i W_i$. Since $Z$ is a Jacobi field, it is a solution of the Jacobi equation. Evaluating the Jacobi equation for $f^i W_i$ and using the fact that each $W_i$ satisfies the Jacobi equation gives

$$0 = \nabla \dot{\gamma} \nabla \dot{\gamma} (f^i W_i) + R(\dot{\gamma}, f^i W_i) \dot{\gamma} = \nabla \dot{\gamma} \left( \dot{f}^i W_i + f^i \nabla \dot{\gamma} W_i \right) + f^i R(\dot{\gamma}, W_i) \dot{\gamma}$$

$$= \dot{f}^i W_i + 2 \dot{f}^i \nabla \dot{\gamma} W_i + f^i \nabla \dot{\gamma} \nabla \dot{\gamma} W_i + f^i R(\dot{\gamma}, W_i) \dot{\gamma}$$

$$= \ddot{f}^i W_i + 2 \dot{f}^i \nabla \dot{\gamma} W_i$$

This implies that $\dot{f}^i = 0 \ \forall i$. If not, then $\nabla \dot{\gamma} W_i$ would be some multiple of $W_i$ which would contradict the fact that $Z(t_0 + \delta) = 0$. Thus $f^i$ is a constant for each $i$.

By the lemma, $I(W)$ can be written

$$I(W) = \int_0^1 |\dot{f}^i W_i|^2 \ dt - f^i(0) f^j(0) \langle \nabla \dot{\gamma} W_i, W_j \rangle_{t=0}$$

And for $Z = f^i W_i$ where the coefficients are constant the formula reduces to

$$I(Z) = -f^i(t_0 - \delta) f^j(t_0 - \delta) \langle \nabla \dot{\gamma} W_i, W_j \rangle_{t=t_0-\delta}$$

And for any smooth field $W = g^i W_i$ which agrees with $Z$ on $t_0 - \delta$ and $t_0 + \delta$ we have

$$I(W) - I(Z) = \int_{t_0 - \delta}^{t_0 + \delta} |g^i W_i|^2 \ dt \geq 0$$

So $Z$ is strictly minimal amongst smooth fields with the same endpoints.