Stochastic Analysis on Manifolds

Prakash Balachandran
Department of Mathematics
Duke University

September 21, 2008

These notes are based on Hsu’s *Stochastic Analysis on Manifolds*, Kobayahi and Nomizu’s *Foundations of Differential Geometry Volume I*, and Lee’s *Introduction to Smooth Manifolds* and *Riemannian Manifolds*.

1 Chapter 1

1.1 SDE’s on Euclidean Space

1.1.1 A Review of Basic Results

This section reviews the theory of SDE’s in Euclidean space. All theorems are presented without proofs. See Hsu chapter 1.1 for details.

Throughout this section, we concern ourselves with the SDE:

\[ X_t = X_0 + \int_0^t \sigma(X_s) dZ_t \]  \hspace{1cm} (1)

where \( Z_t \) is an \( \mathbb{R}^n \) valued \( \mathcal{F}_\tau \)-semimartingale [henceforth, this always means adapted to the filtration \( \mathcal{F}_\tau = \{ \mathcal{F}_t \}_{t \geq 0} \)].

A classical result of the existence and uniqueness of a solution to this SDE is given by

**Theorem 1** Suppose that \( \sigma \) is globally Lipschitz and \( X_0 \) is square integrable. Then, the SDE (1) has a unique solution \( X = \{ X_t : t \geq 0 \} \)

For a proof, see Hsu, page 9.
In general, when $\sigma$ is locally Lipschitz, we have to allow the possibility of explosion.

**Example 1:** Consider the SDE:

$$dX_t = X_t^2 dt, \quad X_0 = 1.$$ 

Clearly, this has the solution $X_t = \frac{1}{1-t}$ which explodes at time $t = 1$.

To take care of this case and transfer the setting to manifolds, we establish some notation and introduce some definitions.

**Definition:** Let $M$ be a locally compact metric space, and let $\widehat{M} = M \cup \{\partial M\}$ be the one-point compactification of $M$.

An $M$-valued path $X_t$ with exploding time $e = e > 0$ is a continuous map $X : [0, \infty) \to \widehat{M}$ such that $X_t \in M$ for $0 \leq t < e$ and $X_t = \partial M$ for all $t \geq e$ if $e < \infty$.

The space of $M$-valued paths with explosion time is called the path space of $M$ and is denoted by $W(M)$.

When we consider processes $X_t(\omega)$, we can define $e(\omega)$ pathwise, but it’s not immediately clear that $e$ is a stopping time.

For our purposes, recall that an exhaustion of a locally compact metric space $M$ is a sequence of relatively compact open sets $\{U_N\}$ such that $U_N \subseteq U_{N+1}$ and $M = \bigcup_{N=1}^{\infty} U_N$.

**Theorem 2**

1. Let $\{U_N\}$ be an exhaustion, and $\tau_{U_N}$ be the first exit time from $U_N$. Then, $\tau_{U_N} \uparrow e$ as $N \uparrow \infty$.

2. Suppose that $d : M \times M \to \mathbb{R}^+$ is a metric on $M$ with the property that every bounded closed set is compact. Fix a point $o \in M$. Let $\tau_R$ be the first exit time from the ball $B(R) = \{x \in M : d(x, o) \leq R\}$. Then, $\tau_R \uparrow e$ as $R \uparrow \infty$.

Part 1 of theorem 2 guarantees us that the explosion time of a process taking values in a Riemannian manifold is indeed a stopping time (use the distance function on the manifold to define the exhaustion $\{U_N\}$).

We will use part 2 of theorem 2 in two situations:

1. $M$ is a complete Riemannian manifold and $d$ is the Riemannian distance function

2. $M$ is embedded as a closed submanifold in another complete Riemannian manifold, and $d$ is the Riemannian metric of the ambient space.
Finally, we define what it means to be a semimartingale up to a stopping time, what it means for a process to solve an SDE up to a stopping time, and give another existence and uniqueness result.

**Definition:** Let \((\Omega, \mathcal{F}, P)\) be a filtered probability space and \(\tau\) be an \(\mathcal{F}\)-stopping time. A continuous process \(X\) defined on the stochastic time interval \([0, \tau]\) is called an **\(\mathcal{F}\)**-semimartingale up to \(\tau\) if there exists a sequence of \(\mathcal{F}\)-stopping times \(\tau_n \uparrow \tau\) such that for each \(n\), the stopped processes \(X_{t \wedge \tau_n}\) are semimartingales.

**Definition:** A semimartingale \(X\) up to a stopping time \(\tau\) is a solution of the SDE (1) if there is a sequence of stopping times \(\tau_n \uparrow \tau\) such that for each \(n\) the stopped process \(X_{\tau_n} = X_{t \wedge \tau_n}\) is a semimartingale, and

\[
X_{t \wedge \tau_n} = X_0 + \int_0^{t \wedge \tau_n} \sigma(X_s) s Z_s, \quad ll t \geq 0.
\]

**Theorem 3** Suppose that we are given:

1. A locally Lipschitz coefficient matrix \(\sigma : \mathbb{R}^N \to \mathcal{M}(N, n)\)
2. An \(\mathbb{R}^n\) valued \(\mathcal{F}\)-semimartingale \(Z = \{Z_t : t \geq 0\}\) on a filtered probability space \((\Omega, \mathcal{F}, P)\)
3. An \(\mathbb{R}^n\) valued \(\mathcal{F}_0\)-measurable random variable \(X_0\).

Then, there is a unique \(W(\mathbb{R}^N)\)-valued random variable \(X\) which is a solution to (1) up to its explosion time \(e\).

**Theorem 4** Suppose that \(\sigma\) is locally Lipschitz. Let \(X\) and \(Y\) be two solutions of the SDE (1) up to stopping times \(\tau\) and \(\eta\) respectively. Then, \(X_t = Y_t\) for \(0 \leq t < \tau \wedge \eta\). In particular, if \(X\) is a solution up to its explosion time \(e\), then \(\eta \leq e\) and \(X_t = Y_t\) for \(0 \leq t < \eta\).

**Theorem 5** Suppose that \(\sigma\) and \(b\) are locally Lipschitz. Then the weak uniqueness property holds for the SDE

\[
X_t = X_0 + \int_0^t \sigma(X_s) dW_s + \int_0^t b(X_s) ds;
\]

That is, if \(\tilde{X}\) is the solution of

\[
\tilde{X}_t = \tilde{X}_0 + \int_0^t \sigma(\tilde{X}_s) d\tilde{W}_s + \int_0^t b(\tilde{X}_s) ds,
\]

where \(\tilde{W}\) is a Brownian motion defined on another filtered probability space \((\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})\) and \(\tilde{X}_0 \in \hat{\mathcal{F}}_0\) has the same law as \(X_0\), then \(\tilde{X}\) and \(X\) have the same law.

Computationally, the following result is useful:

**Proposition 1** Suppose that \(Z\) is defined on \([0, \infty)\). If \(\sigma\) is locally Lipschitz, and there is a constant \(C\) such that

\[
|\sigma(x)| \leq C(1 + |x|)
\]
then the solution of (1) does not explode.

Finally, we give the existence and uniqueness result of a solution to (1) when $Z$ is a semimartingale
defined up to a stopping time $\tau$:

**Theorem 6** Let $Z$ be a semimartingale defined up to a stopping time $\tau$. Then, there is a unique solution
$X$ to the SDE (1) up to the stopping time $e_X \wedge \tau$. If $Y$ is another solution up to a stopping time $\eta \leq \tau$
then $\eta \leq e_X \wedge \tau$ and $X_t = Y_t$ for $0 \leq t < e_X \wedge \eta$.

1.1.2 Stratonovich vs. Ito

In this section, we briefly review the difference between the Ito formalism and the Stratonovich formalism
of the stochastic integral, and proceed to determine how they are equivalent. Throughout, we take as
given the Ito machinery, and proceed to derive the equivalent Stratonovich machinery.

The Ito integral of a left continuous predictable function $H$ w.r.t. a continuous semimartingale
$Z = M + A$ is roughly defined to be a left-endpointed integral of the paths of $H$ over those of $Z$ (of course care must be taken when considering the stopping times defining the continuous local martingale $M$).

The Stratonovich integral is like the Ito integral except that the integral is taken w.r.t. midpoints. More
precisely, the two are related by:

$$ \int_0^t H_s(\omega) \circ dZ_s(\omega) = \int_0^t H_s(\omega) dZ_s(\omega) + \frac{1}{2} \langle H, Z \rangle_s $$

where $\langle H, Z \rangle_s$ denotes the covariance process (otherwise known as the bracket of $H$ and $Z$).

Now, consider the Stratonovich SDE (Note: we employ Einstein summation notation throughout; that is, repeated indices are always summed over):

$$ X_t = X_0 + \int_0^t V_\alpha(X_s) \circ dZ^\alpha_s $$

where $X_0$ is an $\mathbb{R}^N$-valued random variable, $Z_s = (Z^\alpha_s)_{\alpha=1}^n$ is an $\mathbb{R}^n$-valued semimartingale, and $V = \{V_\alpha\}_{\alpha=1}^n$ are $n$ smooth vector fields on $\mathbb{R}^N$.

How do we recast this SDE in terms of the Ito integral? Well, in components:

$$ X^i_t = X^i_0 + \int_0^t V^i_\alpha(X_s) \circ dZ^\alpha_s = X^i_0 + \int_0^t V^i_\alpha(X_s) dZ^\alpha_s + \frac{1}{2} \langle V^i_\alpha(X), Z^\alpha \rangle_s. $$

Using Ito’s formula:

$$ V^i_\alpha(X_t) = V^i_\alpha(X_0) + \int_0^t \frac{\partial V^i_\alpha}{\partial x_k}(X_s) dX^k_s + \int_0^t \frac{\partial^2 V^i_\alpha}{\partial x_k \partial x_l}(X_s) d\left\langle X^k, X^l \right\rangle_s. $$
so that:

\[ X^i_t = X^i_0 + \int_0^t V^i_\alpha(X_s)dZ_s^\alpha + \frac{1}{2} \left\langle V^i_\alpha(X_0) + \int_0^t \frac{\partial V^i_\alpha}{\partial x_k}(X_s)dX^k_s + \int_0^t \frac{\partial^2 V^i_\alpha}{\partial x_k \partial x_t}(X_s)d\left\langle X^k_s, X^t_s \right\rangle, Z^\alpha \right\rangle_s \]

\[ = X^i_0 + \int_0^t V^i_\alpha(X_s)dZ_s^\alpha + \frac{1}{2} \left\langle \int_0^t \frac{\partial V^i_\alpha}{\partial x_k}(X_s)dX^k_s, Z^\alpha \right\rangle_s \]

\[ = X^i_0 + \int_0^t V^i_\alpha(X_s)dZ_s^\alpha + \frac{1}{2} \left\langle \int_0^t \frac{\partial V^i_\alpha}{\partial x_k}(X_s)V^k_\beta(X_s)dZ^\beta_s, \int_0^t dZ^\alpha \right\rangle_s \]

\[ = X^i_0 + \int_0^t V^i_\alpha(X_s)dZ_s^\alpha + \frac{1}{2} \int_0^t \frac{\partial V^i_\alpha}{\partial x_k}(X_s)V^k_\beta(X_s)d\left\langle Z^\beta, Z^\alpha \right\rangle_s \]

Recalling the definition of the Euclidean connection: \( (\nabla_{V^\beta} V^\alpha)_i = V^\beta_i \frac{\partial V^\alpha_i}{\partial x_k} \), we have finally:

\[ X^i_t = X^i_0 + \int_0^t V^i_\alpha(X_s)dZ_s^\alpha + \frac{1}{2} \int_0^t \nabla_{V^\alpha} V^\alpha d\left\langle Z^\alpha, Z^\beta \right\rangle_s \]

(3)

\[ \Leftrightarrow X_t = X_0 + \int_0^t V_\alpha(X_s)dZ^\alpha + \frac{1}{2} \int_0^t \nabla_{V^\alpha} V^\alpha d\left\langle Z^\alpha, Z^\beta \right\rangle_s \]

(4)

Thus, the equations (2) and (4) are equivalent.

**Proposition 2** Let \( X \) be a solution to the equation (2), and \( f \in C^2(\mathbb{R}^d) \). Then:

\[ f(X_t) = f(X_0) + \int_0^t V_\alpha f(X_s)dZ^\alpha, \quad 0 \leq t < \epsilon_X. \]

**Proof of Proposition 2**: From the preceding argument, we may assume the equivalence of the SDE’s (2) and (4), and hence may assume (3).

Ito’s formula implies:

\[ df(X_s) = \frac{\partial f}{\partial x_i}(X_s)dX^i_s + \frac{1}{2} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s)d\left\langle X^i_s, X^j_s \right\rangle_s \]

\[ = \frac{\partial f}{\partial x_i}(X_s) \left[ V^i_\alpha(X_s)dZ^\alpha_s + \frac{1}{2} \left\langle \nabla_{V^\beta} V^\alpha \right\rangle_i(X_s)d\left\langle Z^\alpha, Z^\beta \right\rangle_s \right] + \frac{1}{2} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s)V^i_\alpha(X_s)V^j_\beta(X_s)d\left\langle Z^\alpha, Z^\beta \right\rangle_s \]

\[ = V^i_\alpha(X_s) \frac{\partial f}{\partial x_i}(X_s)dZ^\alpha_s + \frac{1}{2} \frac{\partial f}{\partial x_i}(X_s)\frac{\partial V^i_\alpha}{\partial x_k}(X_s)d\left\langle Z^\alpha, Z^\beta \right\rangle_s + \frac{1}{2} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s)V^i_\alpha(X_s)V^j_\beta(X_s)d\left\langle Z^\alpha, Z^\beta \right\rangle_s \]

\[ = V^i_\alpha(X_s) \frac{\partial f}{\partial x_i}(X_s)dZ^\alpha_s + \frac{1}{2} \frac{\partial f}{\partial x_j}(X_s) \left[ V^j_\beta(X_s) \frac{\partial f}{\partial x_i}(X_s) - V^i_\alpha(X_s) \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s) \right] d\left\langle Z^\alpha, Z^\beta \right\rangle_s \]

\[ + \frac{1}{2} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s)V^i_\alpha(X_s)V^j_\beta(X_s)d\left\langle Z^\alpha, Z^\beta \right\rangle_s \]
\[ = V^i_\alpha(X_s) \frac{\partial f}{\partial x_i}(X_s) dZ^\alpha_s + \frac{1}{2} V^j_\beta(X_s) \frac{\partial}{\partial x_j} \left( V^i_\alpha(X_s) \frac{\partial f}{\partial x_i}(X_s) \right) d\langle Z^\alpha, Z^\beta \rangle_s \]

where in the fourth line, we have used the product rule on \( V^i_\alpha \frac{\partial f}{\partial x_i} \). Comparing this final expression with (4), and backtracking through the proof of it, we find that upon integrating from 0 to \( t \):

\[ f(X_t) = f(X_0) + \int_0^t V_\alpha f(X_s) \circ dZ^\alpha_s, \quad 0 \leq t < e_X. \]
1.2 SDE’s on manifolds

Recall that in geometry, we always define objects on a manifold through coordinate charts.

1. Let $\mathcal{M}$ be a smooth, connected, manifold of dimension $m$. We define a function $f : \mathcal{M} \to \mathbb{R}$ to be smooth at $p \in \mathcal{M}$ if there exists a coordinate chart $(U_\alpha, \phi_\alpha)$ such that $p \in U_\alpha$ and $f \circ \phi_\alpha^{-1} : \mathbb{R}^m \to \mathbb{R}$ is smooth at $\phi(p)$.

$f$ is smooth on an open set $U \subseteq \mathcal{M}$ if $f$ is smooth at every point $p \in U$.

Clearly, if this definition holds in any one coordinate chart, it holds in all coordinate charts within the given differentiable structure of the manifold.

2. In addition to the notation above, let $\mathcal{N}$ be an $n$ dimensional, connected, smooth manifold. We say that a map $\Phi : \mathcal{M} \to \mathcal{N}$ is smooth at $p \in \mathcal{M}$ is there exists coordinate charts $(U_\alpha, \phi_\alpha)$, $(V_\beta, \psi_\beta)$ of $\mathcal{M}$ and $\mathcal{N}$ respectively, such that $\Phi(p) \in V_\beta$ and the map: $\psi_\beta \circ f \circ \phi_\alpha^{-1} : \mathbb{R}^m \to \mathbb{R}^n$ is smooth at $\phi_\alpha(p)$.

Again, it’s clear that this definition is invariant of the charts chosen [within respective differentiable structures].

Thus, it’s clear that we should make the definition:

**Definition 1** Let $\mathcal{M}$ be a differentiable manifold and $(\Omega, \mathcal{F}, P)$ be a filtered probability space with filtration $\mathcal{F}$. Let $\tau$ be a $\mathcal{F}$-stopping time. A continuous, $\mathcal{M}$-valued process $X$ defined on $[0, \tau)$ is called an $\mathcal{M}$-valued semimartingale if $f(X)$ is a real-valued semimartingale on $[0, \tau)$ for all $f \in C^\infty(\mathcal{M})$.

It’s easy to see that when $\mathcal{M} = \mathbb{R}^n$, this definition gives the usual definition of an $\mathbb{R}^m$ valued semimartingale.

A stochastic differential equation on a manifold $\mathcal{M}$ is defined by $n$ vector fields $V_1, \ldots, V_n$ on $\mathcal{M}$, a $\mathbb{R}^n$-valued semimartingale $Z$, and an $\mathcal{M}$-valued random variable $X_0 \in \mathcal{F}_0$ serving as an initial condition. We write the equation symbolically as:

$$dX_t = V_\alpha(X_t) \circ dZ^\alpha_t. \tag{5}$$

In light of Proposition 2, we make the following definition:

**Definition:** An $\mathcal{M}$-valued semimartingale $X$ defined up to a stopping time $\tau$ is a solution of the SDE (5) up to $\tau$ if for all $f \in C^\infty(\mathcal{M})$,

$$f(X_t) = f(X_0) + \int_0^t V_\alpha f(X_s) \circ dZ^\alpha_s, \quad 0 \leq t < \tau.$$
The advantage of the Stratonovich formulation is that stochastic differential equations on manifolds in this formalism transform consistently under diffeomorphisms between manifolds.

**Proposition 3** Suppose that \( \Phi : \mathcal{M} \to \mathcal{N} \) is a diffeomorphism and \( X \) is a solution of the SDE (5). Then, \( \Phi(X) \) is a solution of the SDE:

\[
dY_t = \Phi_\ast V_\alpha \circ dZ^\alpha_t, \quad Y_0 = \Phi(X_0).
\]  

**Proof:** Let \( Y = \Phi(X) \), and \( f \in C^\infty(\mathcal{N}) \). Then, \( \Phi^\ast f = f \circ \Phi \in C^\infty(\mathcal{M}) \). Since \( X \) solves (5), we have that

\[
(f \circ \Phi)(X_t) = (f \circ \Phi)(X_0) + \int_0^t V_\alpha(f \circ \Phi)(X_s) \circ dZ^\alpha_s
\]

\[
\Rightarrow f(Y_t) = f(Y_0) + \int_0^t (\Phi_\ast V_\alpha)(f(Y_s)) \circ dZ^\alpha_s.
\]

Hence, since \( f \) was arbitrary, \( Y = \Phi(X) \) is a solution to the SDE (6). \( \Box \)

[Insert remark about Ito integral here]

Now, we want a unique solution to exist for the SDE (5). To see that this is in fact true, we reduce the equation to an SDE in Euclidean space by embedding the manifold into \( \mathbb{R}^N \) for some \( N \). Then, we will use the results stated in section 1 about the existence and uniqueness of SDE’s in Euclidean space.

Towards this end, we have:

**Theorem 7** (Whitney’s embedding theorem) Suppose that \( \mathcal{M} \) is a differentiable manifold. Then, there exists an embedding \( i : \mathcal{M} \to \mathbb{R}^N \) for some \( N \) such that the image \( i(\mathcal{M}) \) is a closed subset of \( \mathbb{R}^N \).

It is well-known that \( N = 2 \dim \mathcal{M} + 1 \) will suffice.

Now, fix an embedding of \( \mathcal{M} \) into \( \mathbb{R}^N \) and regard \( \mathcal{M} \) as a closed submanifold of \( \mathbb{R}^N \). By the smooth Urysohn lemma:

**Proposition 4** Let \( U \subseteq \mathcal{M} \) be an open subset of a smooth manifold and \( K \subseteq U \) a subset that is closed in \( \mathcal{M} \). Then, there is a smooth function \( f : \mathcal{M} \to \mathbb{R} \) such that \( f|_K \equiv 1 \) and \( \text{supp}(f) \subseteq U \).

We identify \( \mathcal{M} \) with its embedding \( i(\mathcal{M}) \subseteq \mathbb{R}^N \).

Now, we can “fatten” the manifold \( \mathcal{M} \) slightly so that each \( V_\alpha \) is smoothly defined on the same open subset \( U \) containing \( \mathcal{M} \). Applying proposition 4, we obtain a smooth function \( f : \mathbb{R}^N \to \mathbb{R} \) such that \( f|_{\mathcal{M}} \equiv 1 \) and \( \text{supp}(f) \subseteq U \), and hence obtain vector fields \( \tilde{V}_\alpha = fV_\alpha \) defined on all of \( \mathbb{R}^N \).
From section 1, we know that the equation

$$X_t = X_0 + \int_0^t \tilde{V}_\alpha(X_s) \circ dZ_\alpha^s$$  \hfill (7)

on $\mathbb{R}^N$ has a unique solution $X$ up to its exploding time $e_X$.

We need to know that if $X$ starts on $M$ then it never leaves $M$, which makes sense since the vector fields $\tilde{V}_\alpha$ are tangent to $M$ on $M$. This is given by:

**Proposition 5** Let $X$ be the solution of the extended equation up to its explosion time $e_X$, and let $X_0 \in M$. Then, $X_t \in M$ for $0 \leq t < e_X$.

See Hsu page 22-23 for the proof.

Finally, we proceed to prove the existence and uniqueness result:

**Theorem 8** There is a unique solution of the SDE (5) up to its explosion time.

**Proof of theorem 8:** By proposition 5, the solution to the extended SDE (7) stays in $M$ up to its exploding time.

To show that this solves SDE (5), we prove the following

**Lemma 1** Suppose that $M$ is a closed submanifold of $\mathbb{R}^N$. Let $f^1, \ldots, f^N$ be the coordinate functions. Let $X$ be an $M$-valued continuous process.

1. $X$ is a semimartingale on $M$ iff it is an $\mathbb{R}^N$-valued semimartingale, or equivalently, iff $f^i(X)$ is a real-valued semimartingale for each $i = 1, \ldots, N$.

2. $X$ is a solution of the SDE (5) up to a stopping time $\sigma$ iff for each $i = 1, \ldots, N$,

$$f^i(X_t) = f^i(X_0) + \int_0^t V^i(X_s) \circ dZ^\alpha_s, \quad 0 \leq t < \sigma$$  \hfill (8)

**Proof of Lemma 1:**

1) Suppose that $X$ is an $M$-valued semimartingale. Each $f^i$ is a smooth function on $M$, so that $f^i(X) = X^i$ is a real-valued semimartingale. Thus, $X$ is a $\mathbb{R}^N$-valued semimartingale.

Conversely, suppose that $X$ lives on $M$ and is a $\mathbb{R}^N$-valued semimartingale. Since $M \subseteq \mathbb{R}^N$ is closed, we may use Urysohn’s lemma to extend $f \in C^\infty(M)$ to $\tilde{f} \in C^\infty(\mathbb{R}^N)$. Hence, $f(X) = \tilde{f}(X)$ is a real-value semimartingale, and so by definition, $X$ is an $M$-valued semimartingale.

2) If $X$ is a solution to SDE (5), then (8) must hold since $f^i \in C^\infty(M)$.

If (8) holds, and $f \in C^\infty(M)$, take an extension $\tilde{f} \in C^\infty(\mathbb{R}^N)$ of $f$. Then,

$$f(X_t) = \tilde{f}\left(f^1(X_t), \ldots, f^N(X_t)\right).$$
Thus, by Ito’s formula:

\[
    df(X_t) = \frac{\partial f}{\partial x_i}(f^1(X_t), \ldots, f^N(X_t)) \circ df^i(X_t)
\]

\[
    = \frac{\partial f}{\partial x_i}(f^1(X_t), \ldots, f^N(X_t)) \circ V_i f^i(X_t) \circ dZ_t^\alpha
\]

\[
    = \frac{\partial f}{\partial x_i}(X^1_t, \ldots, X^N_t) V_i f^i(X_t) \circ dZ_t^\alpha
\]

\[
    = V_i f(X_t) \circ dZ^\alpha
\]

where in the last step, we have used the chain rule for differentiating composite functions. △

Since (7) is a rewriting of (8), lemma 1 implies that \( X \) is a solution of SDE (5).

If \( Y \) is another solution up to a stopping time \( \tau \), then as a \( \mathbb{R}^N \)-valued semimartingale, it also solves SDE (7) up to \( \tau \). By the uniqueness statement of theorem 4, \( Y \) must coincide with \( X \) on \([0, \tau)\).

□

Existence and uniqueness of solutions to SDE (5) will be very important in the stochastic lifting process in Chapter 2 of Hsu.
1.3 Diffusion Processes

This section briefly discusses diffusion processes, which will take a prominent role in chapter 3 of Hsu when we discuss Brownian motion on manifolds.

Throughout this section $L$ denotes a smooth second order elliptic, but not necessarily nondegenerate, differential operator on a differentiable manifold $\mathcal{M}$. If $f \in C^2(\mathcal{M})$ and $\omega \in W(M)$, we let:

$$M^f(\omega)_t = f(\omega_t) - f(\omega_0) - \int_0^t Lf(\omega_s)ds, \quad 0 \leq t < e_\omega.$$

**Definition:**

1. An $\mathcal{F}_t$-adapted stochastic process $X : \Omega \rightarrow W(\mathcal{M})$ defined on a filtered probability space $(\Omega, \mathcal{F}_t, P)$ is called a **diffusion process generated by** $L$ (or simply an $L$-diffusion) if $X$ is an $\mathcal{M}$-valued $\mathcal{F}_t$-semimartingale up to $e_X$ and

$$M^X_t(X)_t = f(X_t) - f(X_0) - \int_0^t Lf(X_s)ds, \quad 0 \leq t < e_X$$

is a local $\mathcal{F}_t$-martingale for all $f \in C^\infty(\mathcal{M})$.

2. A probability measure $\mu$ on the standard filtered space $(W(\mathcal{M}), B(W(\mathcal{M}))_t)$ is called a **diffusion measure generated by** $L$ (or simply an $L$-diffusion measure) if

$$M^f(\omega)_t = f(\omega_t) - f(\omega_0) - \int_0^t Lf(\omega_s)ds, \quad 0 \leq t < e_\omega$$

is a local $B(W(\mathcal{M}))_t$-martingale for all $f \in C^\infty(\mathcal{M})$.

**Remark:** While at first it appears that the measure $\mu$ doesn’t appear anywhere in this definition, notice that the property of being a $B(W(\mathcal{M}))_t$-martingale depends on the measure, since different measure might be defined on different $\sigma$-algebras.

For a given $L$, the relation between $L$-diffusion measures and $L$-diffusion processes is as follows. If $X$ is an $L$-diffusion, then its law $\mu^X =) \circ X^{-1}$ is an $L$-diffusion measure. Conversely, if $\mu$ is an $L$-diffusion measure on $W(M)$, then the coordinate process $X(\omega)_t = \omega_t$ on $(W(M), B(W(M)))_t, \mu)$ is an $L$-diffusion process.

The main result of use is that given a smooth second order elliptic operator and a probability distribution $\mu_0$ on $\mathcal{M}$, there exists a unique $L$-diffusion measure whose initial distribution if $\mu_0$.

As a preview of chapter 3, we note that this will be used in the following case:

Let $\mathcal{M}$ be a Riemannian manifold equipped with the Levi-Civita connection $\nabla$ and Laplace-Beltrami operator $\triangle M$ on $\mathcal{M}$ (which is a second order elliptic operator). Given a probability measure $\mu$ on $\mathcal{M}$, the above implies that there exists a unique $\triangle M^2$-diffusion measure $P_\mu$ on the filtered measure space.
\((W(\mathcal{M}), \mathcal{B}(W(\mathcal{M})))\) (the path space over \(\mathcal{M}\)). Any \(\Delta_\frac{\lambda}{2}\)-diffusion measure on \(W(\mathcal{M})\) is called a Wiener measure on \(W(\mathcal{M})\).

If we define an \(\mathcal{M}\)-valued stochastic process to be a measurable map \(X : \Omega \to W(\mathcal{M})\) where \((\Omega, \mathcal{F})\) is some probability space, Brownian motion on \(\mathcal{M}\) is defined to be any \(\mathcal{M}\)-valued stochastic process \(X\) whose law is a Wiener measure on the space space \(W(\mathcal{M})\). In particular, it can be shown that any \(\Delta_\frac{\lambda}{2}\)-diffusion process is a Brownian motion.
2 Chapter 2

2.1 Frame Bundles and Connections

First, recall the definition of a vector bundle:

**Definition:** Let $M$ be a smooth $m$-manifold, $E$ a smooth manifold of dimension $m + n$, and $\pi : E \to M$ a smooth map. This will be called an $n$-plane bundle over $M$ (or a vector bundle over $M$ of fiber dimension $n$) if the following properties hold:

1) For each $x \in M$, $E_x = \pi^{-1}(x)$ has the structure of a real, $n$-dimensional vector space.

2) There is an open cover $\{W_j\}_{j \in J}$ of $M$ together with the commutative diagram

$$
\begin{array}{ccc}
\pi^{-1}(W_j) & \xrightarrow{\psi_j} & W_j \times \mathbb{R}^n \\
\downarrow \pi & & \downarrow p_1 \\
W_j & \xrightarrow{Id_M} & W_j
\end{array}
$$

such that $\psi_j$ is a diffeomorphism $\forall j \in J$, where $p_1$ is the projection onto the first factor.

3) For each $j \in J$ and $x \in W_j$, $\psi_jx = \psi_j|_{E_x}$ maps the vector space $E_x$ isomorphically onto the vector space $\{x\} \times \mathbb{R}^n$.

Now consider the definition:

**Definition:** Let $M$ be a manifold, and $G$ be a Lie group. A (differentiable) principal fiber bundle over $M$ with group $G$ consists of a manifold $P$ and an action $G$ on $P$ satisfying the following conditions:

1) $G$ acts freely on $P$ on the right: $(u, a) \in P \times G \to ua = R_a u \in P$. (freely just means that $ua = u$ implies that $a = I$; i.e. only the identity element fixes any $u \in P$)

2) $M$ is the quotient space of $P$ be the equivalence relation induced by $G$, $M = P/G$, and the canonical projection $\pi : P \to M$ is differentiable.

3) $P$ is locally trivial. That is, every point $x \in M$ has a neighborhood $U$ such that $\pi^{-1}(U)$ is isomorphic with $U \times G$ in the sense that there is a diffeomorphism $\psi : \pi^{-1}(U) \to U \times G$ such that $\psi(u) = (\pi(u), \phi(u))$ where $\phi$ is a mapping of $\pi^{-1}(U)$ into $G$ satisfying $\phi(ua) = \phi(u)a$ for all $u \in \pi^{-1}(U)$ and $a \in G$; i.e. the following diagram commutes:

$$
\begin{array}{ccc}
\pi^{-1}(U) & \xrightarrow{\psi} & U \times G \\
\downarrow \pi & & \downarrow p_1 \\
U & \xrightarrow{Id_M} & U
\end{array}
$$
(1) and (2) imply that the fibers are just the orbits of the action of $G$ on $P$. (3) implies that the fiber over a point $x \in M$ is identified with $\{x\} \times G$ through a local trivialization, and more importantly, says that the action of $G$ on $P$ maps the orbits onto themselves. It also provides a differentiable structure for $P$.

Thus, a principal fiber bundle is just a generalization of a vector bundle.

We call a principal fiber bundle over $M$ with group $G$ a $G$-bundle.

Now, let $M$ be an $m$-dimensional manifold, and consider the $GL(m)$-bundle over $M$.

**Definition:** A frame at $x \in M$ is an $\mathbb{R}$-linear isomorphism $u : \mathbb{R}^m \to T_xM$. We denote the space of all frames at $x \in M$ by $\mathcal{F}(M)_x$.

Since every $M \in GL(m)$ gives such an isomorphism to the tangent space $T_xM$ through its columns, we find that a $GL(m)$-bundle over $M$ is actually a frame bundle. That is, if $P$ denotes the $GL(m)$-bundle,

$$P = \mathcal{F}(M) = \bigcup_{x \in M} \mathcal{F}(M)_x.$$  

Now, the determinant function is a continuous function from $M(m) \to \mathbb{R}$ being a polynomial in the entries of a matrix. Since $GL(m) = \det^{-1}(\mathbb{R} - \{0\})$, we have that $GL(m) \subseteq M(m)$ is open, and hence is an open submanifold of $M(m)$.

Since this space has dimension $m^2$, we find that the local trivializing property of the frame bundle makes the frame bundle into a manifold of dimension $m + m^2$.

Before we can proceed further, we recall the definition of a connection $\nabla$ on a vector bundle $E$. Let $\Gamma(E)$ denote smooth sections of the vector bundle $E$; that is, all smooth maps $f : M \to E$ such that $\pi \circ f = id|_M$.

**Definition:** A connection $\nabla$ is a map

$$\nabla : \Gamma(E) \times \Gamma(E) \to \Gamma(E)$$

such that for all $X, Y, Z \in \Gamma(E)$ and $f, g \in C^\infty(M)$:

1. $\nabla_{fX + gY}Z = f\nabla_XZ + g\nabla_YZ$
2. $\nabla_X(Y + Z) = \nabla_XY + \nabla_XZ$.
3. $\nabla_X(fY) = X(f)Y + f\nabla_XY$.

In particular, when $E = TM$, $\nabla_XY$ is called the covariant differentiation of $Y$ along $X$.

In local coordinates, we may write $X = X^i\partial_i, Y = Y^i\partial_i$. We write

$$\nabla_{\partial_i}\partial_j = \Gamma^k_{ij}\partial_k$$
and refer to $\Gamma^k_{ij}$ as the **Christoffel symbols**.

Furthermore, it can be shown that $(\nabla_X Y)_p$ depends only on the value $X(p)$ and the values of $Y$ along a curve $\gamma$ through $p$ tangent to $X(p)$.

**Proposition 6** Let $\nabla$ be a linear connection on $\mathcal{M}$. For each curve $\gamma : I \to \mathcal{M}$, $\nabla$ determined a unique operator

$$D_t : \Gamma(\gamma(I)) \to \Gamma(\gamma(I))$$

satisfying:

1) $D_t(aV + bW) = aD_tV + bD_tW$.

2) $D_t(fV) = \dot{f}V + f D_tV \forall f \in C^\infty(I)$.

3) If $V$ is extendible, then for any extension $\tilde{V}$ of $V$:

$$D_tV = \nabla_{\dot{\gamma}(t)} \tilde{V}.$$ 

For any $V \in \Gamma(\gamma(I))$, $D_t V$ is called the **covariant derivative of $V$ along $\gamma$**.

The unique vector field guaranteed by the operator in a coordinate chart is specified by

$$D_t V(t_0) = V^j(t_0)\partial_j + V^j(t_0)\nabla_{\dot{\gamma}(t_0)}\partial_j = (V^k(t_0) + V^j(t_0)\dot{\gamma}^i(t_0)\Gamma^k_{ij}(\gamma(t_0)))\partial_k.$$ 

**Definition:** A vector field $V$ along $\gamma$ is said to be **parallel** if $D_t V = 0$. A curve $\gamma$ on $\mathcal{M}$ is said to be a **geodesic** if $D_t \dot{\gamma} = 0$.

Now, let $\mathcal{M}$ be a manifold equipped with a connection $\nabla$ on the tangent bundle $T\mathcal{M}$.

**Definition:** A vector $X \in T_u \mathcal{F}(\mathcal{M})$ is called **vertical** if $X$ is tangent to the fiber $\mathcal{F}(\mathcal{M})_{\pi(u)}$. The space of all vertical vectors at $u$ is denoted by $V_u \mathcal{F}(\mathcal{M})$.

The local trivializing property of the frame bundle implies that for any trivializing neighborhood $U$ of $x \in \mathcal{M}$:

$$\{x\} \times GL(m) \subseteq U \times GL(m) \cong \pi^{-1}(U).$$

The latter is a submanifold of $\mathcal{F}(\mathcal{M})$ since we use $U \times GL(m)$ to give $\mathcal{F}(\mathcal{M})$ a differentiable structure. Thus, the above inclusion is in fact an inclusion of a submanifold, so that $V_u \mathcal{F}(\mathcal{M}) \subseteq T_u \mathcal{F}(\mathcal{M})$ is a subspace of dimension $m^2$, and we have the direct sum decomposition:

$$T_u \mathcal{F}(\mathcal{M}) = V_u \mathcal{F}(\mathcal{M}) \oplus H_u \mathcal{F}(\mathcal{M})$$

with the space $H_u \mathcal{F}(\mathcal{M})$ yet to be determined.
To determine what $H_u\mathcal{F}(M)$ is, note that a curve $u(t)$ in $\mathcal{F}(M)$ is just a smooth choice of frames at each point of the curve $\pi(u(t))$ on $M$.

**Definition:** The curve $u(t)$ above is called **horizontal** if for each $r \in \mathbb{R}^m$, the vector field $u(t)e$ is parallel along $\pi(u(t))$. A tangent vector $X \in T_u\mathcal{F}(M)$ is called **horizontal** if it is the tangent vector of a horizontal curve.

It can be shown that the space of horizontal vectors at $u$ is precisely $H_u\mathcal{F}(M)$, which is isomorphic to $T_{\pi(u)}M$:

The smooth maps: $p_1 : U \times GL(m) \rightarrow U$, $\psi : \pi^{-1}(U) \rightarrow U \times GL(m)$ given by a locally trivializing neighborhood compose to give the linear surjection $(p_1 \circ \psi)|_u : T_u\mathcal{F}(M) \rightarrow T_{\pi(u)}U$. Since the kernel of this map is clearly $V_u\mathcal{F}(M)$ the first isomorphism theorem and the direct sum decomposition $T_u\mathcal{F}(M) = V_u\mathcal{F}(M) \oplus H_u\mathcal{F}(M)$ give $H_u\mathcal{F}(M) \cong T_{\pi(u)}M$.

Via this isomorphism, for each $X \in T_x\mathcal{M}$ and from $u$ at $x$, there is a unique horizontal vector $X^*$, the horizontal lift of $X$ to $u$, such that $\pi_*X^* = X$. It’s easily shown that horizontal vectors are contained in $H_u\mathcal{F}(M)$, and hence form a subspace of dimension less than or equal to $m$, so that he above linear injection of $T_x\mathcal{M}$ into the space of horizontal vectors is in fact an isomorphism.

Given a curve $\gamma(t)$ and a frame $u_0$ at $\gamma(0)$, there is a unique horizontal curve $u(t)$ such that $\pi(u(t)) = \gamma(t)$. One obtains this curve by parallel translating the frame at $\gamma(0)$ along $\gamma$. This is called the **horizontal lift** of $\gamma(t)$ from $u_0$.

Now, for each $e \in \mathbb{R}^m$, $ue \in T_{\pi(u)}M$. Thus, we may take the geodesic $\gamma(t)$ through $\pi(u)$ determined by $ue$ and obtain the horizontal lift $u(t)$ of $\gamma(t)$ from $u$. Defining:

$$H_e(u) = \dot{u}(0)$$

we find that $H_e$ is a horizontal vector field on $\mathcal{F}(M)$. Let $e_1, \ldots, e_m$ be the coordinate unit vectors in $\mathbb{R}^m$. Then, $H_i = H_{e_i}$ $i = 1, \ldots, m$ are the fundamental horizontal fields of $\mathcal{F}(M)$. They span $H_u\mathcal{F}(M)$ at each $u \in \mathcal{F}(M)$.

The action of $GL(m)$ on $\mathcal{F}(M)$ preserves the fundamental horizontal fields as determined by the

**Proposition 7** Let $e \in \mathbb{R}^m$ and $g \in GL(m)$. Then,

$$g_* H_e(u) = H_{ge}(u), \ u \in \mathcal{F}(M)$$

where $g_* : T_u\mathcal{F}(M) \rightarrow T_{ug}\mathcal{F}(M)$ is the action of $g$ on the tangent bundle $T\mathcal{F}(M)$ induced by the canonical action of $g$ on $\mathcal{F}(M)_x$ defined by $u \mapsto ug$. 

16
Now, a local chart $x = \{x^i\}$ on a neighborhood $O \subseteq \mathcal{M}$ induces a local chart on $\tilde{O} = \pi^{-1}(O)$ in $\mathcal{F}(\mathcal{M})$ as follows:

Let $X_i = \frac{\partial}{\partial x_i}$, $1 \leq i \leq m$ be the moving frame on $\mathcal{M}$ determined by the local chart. For a frame $u \in \tilde{O}$ we have that $ue_i = e_i^j X_j$ for some matrix $e = (e_i^j) \in GL(m)$. Then, $(x, e) = (x^i, e_i^j) \in \mathbb{R}^{m+m^2}$ is a local chart for $\tilde{O}$. In terms of this chart, the vertical subspace $V_u \mathcal{F}(\mathcal{M})$ is spanned by $X_{kj} = \frac{\partial}{\partial e_k^j}$ for some vector fields $\{X_i, X_{ij} : 1 \leq i, j \leq d\}$.

Proposition 8 In terms of the local coordinate chart on $\mathcal{F}(\mathcal{M})$ described above, at $u = (x, e) = (x^i, e_i^j) \in \mathcal{F}(\mathcal{M})$ we have

$$H_i(u) = e_i^j X_j - e_i^j e_m^l \Gamma_{jl}^k(x) X_{km}$$

where

$X_i = \frac{\partial}{\partial x_i}$, $X_{km} = \frac{\partial}{\partial e_m^j}$.

Definition: let $u(t)$ be a horizontal lift of a differentiable curve $\gamma(t)$ on $\mathcal{M}$. Since $\dot{\gamma}(t) \in T_{\gamma(t)}\mathcal{M}$, we have that $u_t^{-1}(\dot{\gamma}(t)) \in \mathbb{R}^m$.

The anti-development of $\gamma(t)$ (or of the horizontal curve $u(t)$) is a curve $w(t)$ in $\mathbb{R}^m$ determined by

$$w(t) = \int_0^t u_s^{-1}\dot{\gamma}(t)ds.$$ 

Note that $w$ in the above definition depends on the choice of the initial frame $u_0$ at $x_0$ but in a simple way: if $v(t)$ is another horizontal lift of $\gamma(t)$ and $u_0 = v_0 g$ for some $g \in GL(m)$, then the anti-development of $v(t)$ is $gw(t)$. Since $u(t)\dot{w}(t) = \dot{x}(t)$ we have

$$H_{\dot{w}(t)}(u(t)) = \dot{u}(t)\dot{w}(t) = \dot{x}(t) = \dot{u}(t).$$

Thus, the anti-development $w(t)$ and the horizontal lift $u(t)$ of a curve $\gamma(t)$ on $\mathcal{M}$ are connected by the ODE:

$$\dot{u}(t) = H_i(u(t))\dot{w}^i(t)$$

on $\mathcal{F}(\mathcal{M})$.

If $\mathcal{M}$ is a Riemannian manifold with a Riemannian metric, then we can restrict ourselves to the $O(m)$-bundle consisting of orthonormal frames. In the case we choose the Levi-Civita connection, or any connection that’s compatible with the metric, proposition 8 still holds except that the coordinates $(x, e)$ discussed prior to proposition 8 now give local coordinates of the $O(m)$ bundle.
2.2 Tensor Fields

We can realize covariant differentiation in the frame bundle $\mathcal{F}M$. At each frame $u$, the vectors $X_i = ue_i$, $i = 1, \ldots, m$ form a basis at $T_xM$ with $x = \pi(u)$. Let $X^i$ denote the dual frame on $T^*xM$. Then, an $(r,s)$-tensor $\theta$ can be expressed uniquely as:

$$\theta = \theta^{i_1, \ldots, i_r}_{j_1, \ldots, j_s} X_{i_1} \otimes \cdots \otimes X_{i_r} \otimes X^{j_1} \otimes \cdots \otimes X^{j_s}.$$ 

**Definition:** The scalarization of $\theta$ at $u$ is defined by

$$\tilde{\theta}(u) = \theta^{i_1, \ldots, i_r}_{j_1, \ldots, j_s} e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s}$$

where $\{e_i\}$ is the canonical basis for $\mathbb{R}^m$ and $\{e^i\}$ is the corresponding dual basis.

Thus, if $\theta$ is an $(r,s)$-tensor field on $M$, then its scalarization

$$\tilde{\theta} : \mathcal{F}(M) \to \mathbb{R}^{r} \otimes \mathbb{R}^{*s}$$

is a vector space-valued function on $\mathcal{F}(M)$.

**Definition:** The function $\tilde{\theta}$ defined above is $O(m)$-equivariant, in the sense that $\tilde{\theta}(ug) = g\tilde{\theta}(u)$ where the right action of $g$ is the action of $O(m)$ on the tensor space $\mathbb{R}^{r} \otimes \mathbb{R}^{*s}$.

The following theorem shows covariant differentiation can be realized on the frame bundle:

**Theorem 9** Let $X \in \Gamma(TM)$ and $\theta \in \Gamma(T^{r,s}M)$. Then, the scalarization of the covariant derivative $\nabla_X \theta$ is given by

$$\nabla_X \theta = X^*\tilde{\theta}$$

where $X^*$ is the horizontal lift of $X$.

**Proof of Theorem:** We begin with the case where $\theta = Y$ is a vector field.

Let $\gamma(t)$ be a smooth curve in $M$ such that $\gamma(0) = X_{\gamma(0)}$ and $u(t)$ a horizontal lift of $\gamma(t)$. Let $\tau_t = u_tu_0^{-1}$ be the parallel transport along the curve (that is to say, given a vector $Y \in T_{\gamma(0)}M$, $\tau_t(Y)$ is the parallel transport of $Y$ along $\gamma$, which is a vector field along $\gamma$).

**Lemma 2** $\nabla_X Y|_{\gamma(0)} = \frac{d}{dt}\tau_t^{-1}Y_{\gamma(t)}|_{t=0}$.

**Proof of Lemma:** Let $e_i(t) = u_te_i$ where $e_i$ is the $i$th coordinate unit vector of $\mathbb{R}^m$. Then, $\{e_i(t)\}$ is parallel along $\gamma(t)$.
Now let $Y_{\gamma(t)} = a^i(t)e_i(t)$, which is just the coordinate representation of $Y$ along $\gamma$ in the basis $\{e_i(t)\}$.

Then:

$$\nabla_X Y|_{\gamma(0)} = \nabla_{\gamma(t)} Y|_{\gamma(0)} = D_t Y_{\gamma(t)}|_{t=0} = \dot{a}^i(t)e_i(t)|_{t=0} = \dot{a}^i(0)e_i(0)$$

since $\{e_i(t)\}$ are parallel along $\gamma$, and $D_t$ has the Leibnitz property.

Now, $Y_{\gamma(t)} = \tau_t(a^i(t)e_i(0)) \Rightarrow \tau_t^{-1}Y_{\gamma(t)} = a^i(t)e_i(0)$. Upon differentiation:

$$\frac{d}{dt}\tau_t^{-1}Y_{\gamma(t)}|_{t=0} = \dot{a}^i(0)e_i(0).$$

Thus:

$$\nabla_X Y|_{\gamma(0)} = \frac{d}{dt}\tau_t^{-1}Y_{\gamma(t)}|_{t=0}. \triangle$$

By definition of our horizontal lift, $X^* = \dot{u}(0)$. Now, as in Lemma 2, $Y_{\gamma(t)} = a^i(t)e_i(t)$ implies that $\tilde{Y}_{u(t)} = a^i(t)e_i$. Thus, $u(t)\tilde{Y}_{u(t)} = Y_{\gamma(t)}$, so that $\tilde{Y}(u_t) = u_t^{-1}Y_{\gamma(t)}$.

Thus,

$$\tilde{\nabla}_X \tilde{Y}|_{u(0)} = u^{-1}(0)\nabla_X Y|_{\gamma(0)} = u_0^{-1}\frac{d}{dt}\tau_t^{-1}Y_{\gamma(t)}|_{t=0} = u^{-1}(0)\dot{a}^i(0)e_i(0) = \frac{d}{dt}\tilde{Y}_{u(t)} = X^*\tilde{Y}|_{u(0)}.$$
3 Chapter 3

3.4: The Distance Function

Using exponential coordinates at a point \( o \in M \), we can introduce polar coordinates \((r, \theta)\) in a neighborhood of \( o \).

3.4.1: A review of exponential coordinates (Taken from Peterson, 5.4)

For a tangent vector \( v \in T_pM \), let \( \gamma_v \) denote the unique geodesic with \( \gamma_v(0) = p \) and \( \dot{\gamma}_v(0) = v \), and let \([0, L_v]\) be the nonnegative part of the maximal interval containing 0 on which \( \gamma \) is defined. (recall that this is a result from second order ODE theory: A curve \( \gamma \) is called a geodesic if \( D^2_t \dot{\gamma}_v = 0 \), where \( D^t \) is the covariant derivative along \( \gamma \).)

In local coordinates at \( p \), letting \( \gamma(t) = (\gamma^1(t), \ldots, \gamma^m(t)) \), with \( \gamma(0) = p \), \( \dot{\gamma}(t) = v \), we have:

\[
0 = D^t \dot{\gamma} = D^t \dot{\gamma}^j(t) \partial_j = \dot{\gamma}^j(t) \partial_j + \dot{\gamma}^j(t) \dot{\gamma}^k(t) \Gamma^j_{jk} \partial_i \Rightarrow \dot{\gamma}^j(t) + \dot{\gamma}^l(t) \dot{\gamma}^k(t) \Gamma^j_{lk} \partial_j = 0, \ \forall j
\]

The ”patching” necessary to obtain the coordinate-free existence is straightforward).

Noticing that \( \gamma_{\alpha v}(t) = \gamma_v(\alpha t) \) for all \( \alpha > 0 \) and \( t < L_{\alpha v} \), let \( O_p \subset T_pM \) be the set of vectors \( v \) such that \( 1 < L_v \), so that \( \gamma_v(t) \) is defined on \([0, 1]\). Then the exponential map at \( p \), \( \exp : T_pM \to M \) is defined by

\[
\exp_p(v) = \gamma_v(1) = \gamma_{\frac{|v|}{|v|}}(1), \quad v \in O_p.
\]

A basic result from differential geometry is that the exponential map is locally a diffeomorphism, so that via the above structure, we obtain a coordinate chart on \( M \) at \( p \). It’s clear from the above then, how we obtain polar coordinates \((r, \theta)\) at \( p \in M \) via the exponential map: pick an orthonormal frame \( E_{\alpha} \) on the unit sphere \( S^{m-1} \) and define \( E_1 = \partial_r \) where \( \partial_r \) is the unit vector field such that

\[
\nabla f = \partial_r
\]

where \( f \) is the Riemannian distance function \( f(x) = d(p, x) \). It can be shown that integral curves to \( \partial_r \) are precisely geodesics. Furthermore, we have the important result that for any \( x \in M \) such that \( \exp_p(v) = x \) and \( x \) is in the domain for which \( \exp_p \) is invertible,

\[
x = \gamma_e(d(x, p))
\]

for some unit vector \( e \in T_pM \). To see this, note that since \( x \) is within the domain for which \( \exp_p \) is invertible, \( \gamma_v \) is the unique geodesic such that \( ||\gamma_v|| = d(x, p) \). Thus:

\[
d(x, p) = d(\gamma_v(1), p) = \int_0^1 \sqrt{\langle \gamma_v(s), \gamma_v'(s) \rangle} ds.
\]
Now, since $\gamma_v$ is a geodesic:

$$\frac{d}{dt} \langle \dot{\gamma}_v(t), \dot{\gamma}_v(t) \rangle = \left\langle \dot{\gamma}_v(t), \frac{D}{Dt} \dot{\gamma}_v(t) \right\rangle = 0 \Rightarrow \langle \dot{\gamma}_v(t), \dot{\gamma}_v(t) \rangle = \langle \dot{\gamma}_v(0), \dot{\gamma}_v(0) \rangle = |v|^2.$$  

Thus, $d(x, p) = \int_0^1 |v| ds = |v|$ so that

$$x = \gamma_v(1) = \gamma_v\left(\frac{|v|}{|v|}\right) = \gamma_e(d(x, p))$$

where $e = \frac{v}{|v|}$. Thus, the coordinates determined by the exponential map are isometric ones.

Now, let $t(e)$ be the largest $t$ such that the geodesic $\gamma_e([0, t])$ is distance minimizing from $o$ to $\gamma_e(t)$.

Define

$$\tilde{C}_o = \{ t(e) : e \in T_o M, |e| = 1 \}.$$  

**Definition:** The **cutlocus** of $o$ is the set $C_o = \exp_o \tilde{C}_o$.

The set within the cutlocus is the star-shaped domain

$$\tilde{E}_o = \{ te \in T_o M : e \in T_o M, 0 \leq t < t(e), |e| = 1 \}.$$  

On $\mathcal{M}$ the set within the cutlocus is $E_o = \exp_o \tilde{E}_o$.

We have the following results concerning the cutlocus [See Peterson chapter 5].

**Theorem 3.4.1**

1. The map $\exp : \tilde{E}_o \rightarrow E_o$ is a diffeomorphism.

2. If $x \in C_y$ then $y \in C_x$.

3. $E_o$ and $C_o$ are disjoint and $\mathcal{M} = E_o \cup C_o$.

**Definition:** $i_o = \inf \{ t(e) : e \in T_o M, |e| = 1 \}$ is called the **injectivity radius** at $o$.

We have that $\exp_o B_0(i_o)$ is the largest geodesic ball on which the exponential map is a diffeomorphism onto its image.

Now, this is why we need all this geometry.

Last time, we showed that if $X$ is Brownian motion on $\mathcal{M}$ starting within $E_o$, then before it hits the cutlocus $C_o$,

$$r(X_t) = r(X_0) + \beta_t + \frac{1}{2} \int_0^t \triangle_M r(X_s) ds, \quad t < T_{C_o}$$

where $\beta_t$ is one dimensional Brownian motion.
Thus, if we can control $\triangle_{\mathcal{M}} r$, then we can control the radial process $r(X_t)$ within the cutlocus by comparing it to a one-dimensional diffusion process.

It turns out that the curvature of the manifold $\mathcal{M}$ influences $\triangle_{\mathcal{M}} r$ within the cutlocus. Thus, we need to review the notions of curvature on a manifold $\mathcal{M}$.

3.4.2: A review of curvature on a Riemannian manifold

Given a Riemannian manifold $\mathcal{M}$, let $\nabla$ denote the Levi-Civita connection on $\mathcal{M}$ (A connection "connects" different tangent spaces in the sense that it enables one to compare vectors in different tangent spaces. The Levi-Civita connection is a special kind of connection in that parallel transport along a curve doesn’t change lengths or angles (metric compatible), and doesn’t “rotate” vectors (torsion-free)).

**Definition:** The Riemann curvature tensor is defined by

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

where $X, Y, Z$ are vector fields on $\mathcal{M}$ (while this is a $(1,3)$ tensor, we can define a $(0,4)$ tensor by

$$R(X,Y,Z,W) = \langle R(X,Y)Z, W \rangle,$$

which we shall also refer to as the Riemann curvature tensor).

**Definition:** The sectional curvature is the quadratic form

$$K(X,Y) = \langle R(X,Y)Y, X \rangle$$

where $X, Y$ are vector fields on $\mathcal{M}$. If $X$ and $Y$ are orthogonal unit vectors spanning the two dimensional plane $\sigma$ in the tangent space, then we define the sectional curvature of the plane $\sigma$ to be $K(\sigma) = K(X,Y)$.

**Definition:** Let $X, Y \in T_p\mathcal{M}$. Then, the Ricci curvature tensor is defined by

$$\text{Ric}(X,Y) = \sum_{j=1}^m \langle R(X,X_j)X_j, Y \rangle$$

where $\{X_i\}_{i=1}^m$ is an orthonormal basis of $T_p\mathcal{M}$. The Ricci curvature along $X$ is defined by $\text{Ric}(X,X)$.

**Definition:** The scalar curvature is defined by

$$S = 2 \sum_{i<j} K(X_i,X_j) = \sum_{i=1}^m \text{Ric}(X_i,X_i).$$

**Definition:** We denote the set of sectional curvatures at $x$ by

$$K_\mathcal{M}(x) = \{K(\sigma) : \sigma \subseteq T_x\mathcal{M} is a 2 plane\}.$$

**Definition:** We denote the set of Ricci curvatures at $x$ by

$$\text{Ric}_\mathcal{M}(x) = \{\text{Ric}(X,X) : X \in T_x\mathcal{M}, |X| = 1\}.$$
Now, let
\[ \kappa_1(r) \geq \sup \{ K_M(x) : r(x) = r \} \]
\[ \kappa_2(r) \leq \inf \{ \text{Ric}_M(x) : r(x) = r \} \frac{d-1}{d-1} . \]

and let \( G_i \) be solutions to
\[ G''_i(r) + \kappa_i(r) G_i(r) = 0, \quad G(0) = 0, \quad G'(0) = 1. \]

**Theorem 3.4.2** (Laplacian comparison theorem) With the notation established above, the following inequalities hold for \( x \) within the cut locus:

\[ (d-1) \frac{G'_1(r(x))}{G_1(r(x))} \leq \Delta_M r(x) \leq (d-1) \frac{G'_2(r(x))}{G_2(r(x))} \]

**Proof:** Let \( x \in M \) be within the cut locus of \( o \) and let \( \gamma \) be the unique geodesic from \( o \) to \( x \). Let \( \{X_i(x)\} \) be an orthonormal frame at \( x \) such that \( X_1(x) = X_r(x) \). Let \( \{X_i\} \) be the orthonormal frame along \( \gamma \) obtained by parallel translation of \( \{X_i(x)\} \). Then, \( X_1 \) is the tangent vector field of the geodesic, which will also be denoted by \( T \).

**Claim 1** For any orthonormal basis \( \{X_i\} \) of \( T_x M \) we have
\[ \Delta_M f = \sum_{i=1}^{m} \nabla^2 f(X_i, X_i) \]
where \( \nabla^2 f \) is the Hessian defined by
\[ \nabla^2 f(X, Y) = X(Y f) - (\nabla_X Y) f. \]

**Proof of Claim 1:** By definition
\[ \text{div} X = \sum_{i=1}^{m} \langle \nabla_X X_i, X_i \rangle . \]

So, letting \( X = \text{grad} f \) and using metric compatibility:
\[ \Delta_M f = \text{div} (\text{grad} f) = \sum_{i=1}^{m} \langle \nabla_X \text{grad} f, X_i \rangle = \sum_{i=1}^{m} [\nabla_X \langle \text{grad} f, X_i \rangle - \langle \text{grad} f, \nabla_X X_i \rangle] \]
\[ = \sum_{i=1}^{m} [X_i(X_i f) - \nabla_X X_i(f)] = \sum_{i=1}^{m} \nabla^2 f(X_i, X_i) \cdot \Delta. \]

By Claim 1,
\[ \Delta_M r(\gamma(t)) = \sum_{i=1}^{m} \nabla^2 r(X_i(\gamma(t)), X_i(\gamma(t))). \]
Claim 2  $\nabla^2 r(X_i(\gamma(t)), X_i(\gamma(t))) = \frac{d^2}{ds^2} r(\eta^i(s))|_{s=0}$ where $\eta^i(s)$ is the geodesic such that $\eta^i(0) = \gamma(t)$ and $\dot{\eta}^i(0) = X_i(\gamma(t))$.

Proof of Claim 1: