Notes for Math 250A

Bennett Chow
## Contents

<table>
<thead>
<tr>
<th>Chapter 1. Differentiable Manifolds</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>§1. Basics of differentiable manifolds</td>
<td>1</td>
</tr>
<tr>
<td>§2. Tangent spaces</td>
<td>5</td>
</tr>
<tr>
<td>§3. Differential of a map and tangent bundle</td>
<td>11</td>
</tr>
<tr>
<td>§4. Critical and regular points of functions and level surfaces</td>
<td>15</td>
</tr>
<tr>
<td>§5. Vector fields, Lie brackets and Lie derivatives</td>
<td>15</td>
</tr>
<tr>
<td>§6. The Lie derivative on vector fields</td>
<td>21</td>
</tr>
<tr>
<td>§7. Cotangent spaces, cotangent bundles and 1-forms</td>
<td>26</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 2. Riemannian Manifolds</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>§1. Introduction to Riemannian metrics</td>
<td>29</td>
</tr>
<tr>
<td>§2. Properties of Riemannian metrics</td>
<td>34</td>
</tr>
<tr>
<td>§3. Hypersurfaces of Euclidean space</td>
<td>36</td>
</tr>
<tr>
<td>§4. Lie groups</td>
<td>39</td>
</tr>
<tr>
<td>§5. Levi-Civita connection</td>
<td>41</td>
</tr>
<tr>
<td>§6. Covariant differentiation along a path and parallel translation</td>
<td>47</td>
</tr>
<tr>
<td>§7. The Riemann curvature tensor</td>
<td>52</td>
</tr>
<tr>
<td>§8. The sectional curvature</td>
<td>58</td>
</tr>
<tr>
<td>§9. The Ricci tensor</td>
<td>59</td>
</tr>
<tr>
<td>§10. Pull backs and Lie derivatives of tensors</td>
<td>59</td>
</tr>
<tr>
<td>§11. Covariant derivative of a tensor</td>
<td>61</td>
</tr>
</tbody>
</table>
§12. Connection 1-forms, curvature 2-forms and the Cartan structure equations 62
§13. The Gauss-Bonnet formula 67
§14. Index of a vector field 71
§15. Stokes’s Theorem 73
§16. Exercises on generalized geometry 74

Chapter 3. Geodesics 77
§1. The first variation of arc length 77
§2. Geodesics and the exponential map 82

Chapter 4. Introduction to Geometric Analysis 91
§1. The Hessian and Laplacian of a function 91
§2. Hypersurfaces in Riemannian manifolds 96
§3. Distance functions 99
§4. Volume 100
§5. Covariant tensors and differentiable forms 101

Bibliography 107
Preface

These are notes for my Fall 2011 Math 250A class, which uses do Carmo’s *Riemannian Geometry* book. Thanks to the students for helpful suggestions.

Initially I have strived for quantity instead of quality. Hopefully the notes will continue to be edited and corrected in the future. So for the moment, errors, like pages, abound. As the chapters move forward, the quality of the notes follows the second law of thermal dynamics, namely disorder increases.

**Chapter 1.** In §1 we define differentiable manifold and give the simplest examples of Euclidean spaces and spheres.

In §2 we linearize a manifold at a point. The tangent space comprises equivalence classes of curves in terms of how they act on functions as directional derivative. Equivalently, tangent vectors are derivations, acting on functions linearly and satisfying the product rule.

In §3, given a map between manifolds, we can linearize it at a point to get the differential, which maps between tangent spaces. We introduce diffeomorphisms and tangent bundle.

In §4 we state the result that the inverse image of a regular value is a submanifold.

In §5 we consider vector fields, their Lie brackets, and 1-parameter groups of diffeomorphisms.

In §6 we define the Lie derivative and show that, acting on vector fields, it is the same as the Lie bracket.

In §7 we consider cotangent bundles and 1-forms.
Chapter 2. In §1 we define Riemannian manifolds and give basic examples such as Euclidean space, the sphere, and hyperbolic space.

In §2 we discuss the existence of Riemannian metrics and introduce length and distance.

In §3 we discuss the geometry of Euclidean hypersurfaces, introducing the second fundamental form.

In §4 we discuss Lie groups and left-invariant metrics.

In §5 we prove the existence of the Levi-Civita connection. This includes a prelude to tensors.

In §6 we discuss covariant differentiation along a curve.

In §7 we define the Riemann curvature tensor and discuss its symmetries.

In §8 we define the Ricci tensor.

In §9 we discuss pull backs and Lie derivatives of tensors.

In §10 we discuss the covariant derivative of a tensor.

In §11 define the connection 1-forms and curvature 2-forms and derive the Cartan structure equations.

In §12 we present some exercises on generalized geometry.

Chapter 3. In §1 we discuss the first variation of arc length; vector fields along a map.

In §2 we discuss geodesics and the exponential map. Gauss lemma, Hopf–Rinow, conjugate points, cut points, injectivity radius definition. The sphere as an example.

Chapter 4. In §1 we discuss the Hessian and Laplacian of a function. Bochner formula for $\Delta |\nabla f|^2$. Divergence. Basic equations for Ricci solitons (quasi-Einstein metrics).

In §2 we discuss hypersurfaces in Riemannian manifolds, including level sets and parametrized hypersurfaces.

In §3 we discuss distance functions and derive the Ricatti equation.

In §4 we discuss volume.

In §5 we discuss tensors and differential forms; exterior derivative.
Chapter 1

Differentiable Manifolds

In this chapter we briefly discuss the theory of differentiable manifolds as a prerequisite to Riemannian geometry and geometric analysis.

1. Basics of differentiable manifolds

In this section we recall the basic notions related to the definition of differentiable manifold.

1.1. The differential of a function on Euclidean space.

Let $n$ be a positive integer. Euclidean space of dimension $n$ is defined by $\mathbb{R}^n \equiv \{(q^1, q^2, \ldots, q^n) : q^i \in \mathbb{R} \text{ for } i = 1, 2, \ldots, n\}$. Up to isomorphism, this is the unique real vector space of dimension $n$. An $\mathbb{R}^m$-valued map $\varphi$ defined in a neighborhood of a point $q \in \mathbb{R}^n$ is differentiable if there exists a linear map $d\varphi_q : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{x \to q} \frac{|\varphi(x) - \varphi(q) - d\varphi_q(x - q)|}{|x - q|} = 0.$$  

$d\varphi_q$ is called the differential of $\varphi$ at $q$. It is easy to see that $\varphi$ is differentiable at $q$ if and only if each component function $\varphi^1, \varphi^2, \ldots, \varphi^m$ is differentiable at $q$. Moreover, $d\varphi_q = (d\varphi^1_q, d\varphi^2_q, \ldots, d\varphi^m_q)$; as a matrix it is the corresponding column $m$-vector of differentials. One may express the differential as

$$d\varphi_q(v) = \left. \frac{d}{dt} \right|_{t=0} \varphi(q + tv) = \sum_{i=1}^n v^i \frac{\partial \varphi}{\partial x^i}(q)$$
for \( v = (v^1, v^2, \ldots, v^n) \in \mathbb{R}^n \). Note that for any \( q \) and \( v \) we have \( q + tv \in \text{dom}(\varphi) \) for \( t \) sufficiently small. The linear transformation \( d\varphi_q \) is represented by the matrix \( (\frac{\partial \varphi^i}{\partial x^j}(q)) \), where \( \varphi = (\varphi^1, \varphi^2, \ldots, \varphi^m) \) and where the entry \( \frac{\partial \varphi^i}{\partial x^j}(q) \) is in the \( i^{\text{th}} \) row and \( j^{\text{th}} \) column.

Given an open set \( V \subset \mathbb{R}^n \), a function \( f : V \to \mathbb{R} \) is differentiable if it is differentiable at each point of \( V \). Let \( k \) be a nonnegative integer. A function \( f : V \to \mathbb{R} \) is called \( C^k \) if it is continuously differentiable up to order \( k \), i.e., all of the partial derivatives of \( f \) of order \( \leq k \) exist as functions on \( V \) and they are all continuous on \( V \). Smooth or \( C^\infty \) means \( C^k \) for all \( k \geq 0 \).

There are many other notions of regularity of functions such as Lipschitz, real analytic, etc.

1.2. The definition of differentiable manifold.

A differentiable manifold is simply a locally Euclidean space in the ‘differentiable category’.

**Definition 1.1.** An \( n \)-dimensional differentiable manifold atlas (or simply atlas) on a set \( M \) is a family of injective maps

\[
x_\alpha : U_\alpha \to M
\]

of open sets \( U_\alpha \subset \mathbb{R}^n \) into \( M \) for \( \alpha \) in a set \( \Lambda \) such that

1. \( \bigcup_{\alpha \in \Lambda} x_\alpha(U_\alpha) = M \),
2. for each \( \alpha, \beta \in \Lambda \) with \( x_\alpha(U_\alpha) \cap x_\beta(U_\beta) \neq \emptyset \), \( x_\alpha^{-1}(W) \) and \( x_\beta^{-1}(W) \) are open subsets of \( \mathbb{R}^n \) and the (bijective) transition maps \( x_\beta^{-1} \circ x_\alpha : x_\alpha^{-1}(W) \to x_\beta^{-1}(W) \) and \( x_\alpha^{-1} \circ x_\beta : x_\beta^{-1}(W) \to x_\alpha^{-1}(W) \) are differentiable. (We say that \( (U_\alpha, x_\alpha) \) and \( (U_\beta, x_\beta) \) are (differentiably) compatible).

The map \( x_\alpha : U_\alpha \to M \) is called a parametrization or local coordinate chart.

The simplest example of an atlas on a set is \( \mathbb{R}^n \) with the single map \( \text{id}_{\mathbb{R}^n} : \mathbb{R}^n \to \mathbb{R}^n \).

**Exercise 1.2.** Given an atlas \( \{(U_\alpha, x_\alpha)\}_{\alpha \in \Lambda} \) on a set \( M \), show that there exists a unique atlas \( A \) on \( M \) such that

1. \( \{(U_\alpha, x_\alpha)\}_{\alpha \in \Lambda} \subset A \),
2. If \( B \) is any atlas such that \( \{(U_\alpha, x_\alpha)\}_{\alpha \in \Lambda} \subset B \), then \( B \subset A \).

We call \( A \) a maximal atlas or a differentiable manifold structure on \( M \).

**Hint:** Define \( A \) to be the set of all \( (U, x) \) such that \( x : U \subset \mathbb{R}^n \to M \) is compatible with each \( (U_\alpha, x_\alpha) \), i.e., for each \( \alpha \in \Lambda \) with \( x_\alpha(U_\alpha) \cap x(U) \neq \emptyset \).
1. Basics of differentiable manifolds

\( W_\alpha \neq \emptyset, x_\alpha^{-1}(W_\alpha) \) and \( x^{-1}(W_\alpha) \) are open subsets of \( \mathbb{R}^n \) and the maps \( x^{-1} \circ x_\alpha : x_\alpha^{-1}(W_\alpha) \to x^{-1}(W_\alpha) \) and \( x_\alpha^{-1} \circ x : x^{-1}(W_\alpha) \to x_\alpha^{-1}(W_\alpha) \) are differentiable.

**Definition 1.3.** An \( n \)-dimensional differentiable manifold is a set \( M \) with an \( n \)-dimensional maximal atlas \( A \).

The definition of \( C^k \) manifold, for \( 0 \leq k \leq \infty \), is same as for differentiable manifold except that we require that the transition maps \( x_\beta^{-1} \circ x_\alpha \) and \( x_\alpha^{-1} \circ x_\beta \) are \( C^k \).

**Remark 1.4.** The reader has probably noticed that given a ‘pseudogroup’ \( \Psi \) of local transformations of \( \mathbb{R}^n \), we may define a \( \Psi \) manifold by simply requiring that \( x_\beta^{-1} \circ x_\alpha, x_\alpha^{-1} \circ x_\beta \in \Psi \) (see Kobayashi and Nomizu [7]). In this way we may define the notions of Lipschitz manifold, real analytic manifold, etc.

Euclidean space, as an \( n \)-dimensional \( C^\infty \) manifold, is \( \mathbb{R}^n \) with the unique maximal \( C^\infty \) atlas containing the parametrization \( \text{id}_{\mathbb{R}^n} : \mathbb{R}^n \to \mathbb{R}^n \).

A differentiable manifold structure defines a topology on \( M \).

**Definition 1.5.** A set \( A \subset M \) is defined to be open iff \( x^{-1}(A \cap x(U)) \subset \mathbb{R}^n \) is open for all \( (U, x) \in A \). The topology on \( M \) is the collection of open sets.

**Exercise 1.6.** Show that

1. this definition of open set indeed defines a topology on \( M \),
2. for each \( (U, x) \in A \), \( x(U) \subset M \) is open with respect to this topology and \( x : U \to M \) is continuous.

**Solution:**

(i) Clearly \( \emptyset \subset \mathcal{M} \) is open since \( x^{-1}(\emptyset \cap x(U)) = \emptyset \subset \mathbb{R}^n \) is open. Also, \( \mathcal{M} \) is open since \( x^{-1}(\mathcal{M} \cap x(U)) = U \) is open.

(ii) If \( \{A_\kappa\}_{\kappa \in K} \) is a collection of open sets in \( \mathcal{M} \), then for any parametrization \( (U, x) \)

\[
x^{-1}\left( \bigcup_{\kappa \in K} A_\kappa \right) \cap x(U) = x^{-1}\left( \bigcup_{\kappa \in K} (A_\kappa \cap x(U)) \right) = \bigcup_{\kappa \in K} x^{-1}(A_\kappa \cap x(U))
\]

is a union of open sets in \( \mathbb{R}^n \) and hence open in \( \mathbb{R}^n \).

(iii) If \( \{A_\kappa\}_{\kappa \in K} \) is a finite collection of open sets, then for any parametrization \( (U, x) \)

\[
x^{-1}\left( \bigcap_{\kappa \in K} A_\kappa \right) \cap x(U) = x^{-1}\left( \bigcap_{\kappa \in K} (A_\kappa \cap x(U)) \right) = \bigcap_{\kappa \in K} x^{-1}(A_\kappa \cap x(U))
\]
is a finite intersection of open sets in \( \mathbb{R}^n \) and hence open in \( \mathbb{R}^n \).

(2) Given \((U,x) \in \mathcal{A} \), for each parametrization \((U',x') \in \mathcal{A} \) we have that \((x')^{-1}(x(U) \cap x'(U')) \) is open by the definition of atlas. Moreover, if \( A \subset \mathcal{M} \) is open, then \( x^{-1}(A) = x^{-1}(A \cap x(U)) \) is open. Thus \( x \) is continuous.

1.3. Examples of differentiable manifolds.

In the previous subsection we have seen the noncompact example of Euclidean space. Now we present one of the simplest compact examples.

On \( \mathbb{R}^m \), let \( \langle u,v \rangle = \sum_{i=1}^{m} u^i v^i \) for \( u = (u^1, \ldots, u^m) \), \( v = (v^1, \ldots, v^m) \) denote the standard Euclidean inner product. The Euclidean norm is defined by \( |u| = \sqrt{\langle u,u \rangle} \).

Let \( S^n = \{ x \in \mathbb{R}^{n+1} : |x| = 1 \} \) be the unit \( n \)-sphere. Given a point \( \nu \in S^n \), define the corresponding orthogonal hyperplane
\[
\nu^\perp = \{ x \in \mathbb{R}^{n+1} : \langle x,\nu \rangle = 0 \}
\]
and corresponding **stereographic projection**
\[
\sigma_\nu : S^n - \{ \nu \} \to \nu^\perp
\]
by
\[
\sigma_\nu (x) = \frac{x - \langle x,\nu \rangle \nu}{1 - \langle x,\nu \rangle}.
\]

**Exercise 1.7.**

(1) Show that \( \sigma_\nu \) is well defined.

(2) Define \( x_\nu : \nu^\perp \to S^n - \{ \nu \} \) by \( x_\nu (y) \triangleq \sigma_\nu^{-1} (y) \). Show that
\[
x_\nu (y) = \frac{2y + \left( |y|^2 - 1 \right) \nu}{|y|^2 + 1}.
\]

(3) Let \( N = (0, \ldots, 0, 1) \) and \( S = (0, \ldots, 0, -1) \). We naturally identify \( N^\perp \) with \( \mathbb{R}^n \times \{ 0 \} \) with \( \mathbb{R}^n \), so that \( x_N : \mathbb{R}^n \to S^n \) and \( x_S : \mathbb{R}^n \to S^n \).

(4) Show that the collection \( \{(\mathbb{R}^n,x_N), (\mathbb{R}^n,x_S)\} \) is a \( C^\infty \) atlas on \( S^n \).

The unit \( n \)-sphere \( S^n \), as a \( C^\infty \) manifold, is \( S^n \) with the unique maximal \( C^\infty \) atlas containing \( \{(\mathbb{R}^n,x_N), (\mathbb{R}^n,x_S)\} \).

**Exercise 1.8.** Given differentiable manifolds \( M^m \) and \( N^n \), show that their cartesian product \( M \times N \) is a differentiable manifold.

**HINT.** Let \( \{ (U_\alpha, x_\alpha) \}_{\alpha \in A} \) be an atlas for \( M \) and let \( \{ (V_\beta, y_\beta) \}_{\beta \in B} \) be an atlas for \( N \). Then \( \{ (U_\alpha \times V_\beta, x_\alpha \times y_\beta) \}_{\alpha \in A, \beta \in B} \) where \( x_\alpha \times y_\beta : U_\alpha \times V_\beta \to M \times N \) is defined by \( (x_\alpha \times y_\beta)(q,r) \triangleq (x_\alpha(q), y_\beta(r)) \), is an atlas for \( M \times N \).
2. Tangent spaces

Roughly speaking, the tangent space is the linearization of a manifold at a point.

2.1. The notion of tangent space.

Let $D_p$ be the set of functions on $M$ that are differentiable at $p$. Let $C^{k,\text{loc}}_p$ be the set of functions $f$ on $M$ such that there exists a neighborhood $V$ of $p$ on which $f$ is $C^k$.

Motivated by the notion of directional derivative, the following is a standard definition.

**Definition 1.9.** A **derivation at** $p$ is a linear map $X : D \rightarrow \mathbb{R}$ that satisfies the ‘product rule’:\footnote{The derivation $X$ being linear means that for any $c \in \mathbb{R}$ and $f, g \in D$, $X(cf + g) = cX(f) + X(g)$.}

$$X(fg) = f(p)X(g) + g(p)X(f).$$

The **tangent space** $T_pM$ at $p$ is the set of derivations at $p$. We also call $X(f)$ the **directional derivative** of $f$ in the direction $X$.

A variant of this definition, which we shall also use, is to replace $D$ by $C^{k,\text{loc}}_p$, where $2 \leq k \leq \infty$.

**Exercise 1.10.** Show that $T_pM$ is a real vector space.

Let $1 : M \rightarrow \mathbb{R}$ denote the constant function $1(x) \equiv 1$. Let $X$ be a tangent vector at some point $p \in M$, i.e., a derivation at $p$. Then by the product rule and since $1^2 = 1$ we have

$$X(1) = X(1^2) = 1 \cdot X(1) + 1 \cdot X(1) = 2X(1),$$

which implies $X(1) = 0$. Hence, for any constant function const, we have

$$X(\text{const}) = X(\text{const} \cdot 1) = \text{const} X(1) = 0.$$

Given a differentiable curve $\alpha : (-\epsilon, \epsilon) \rightarrow M$ with $\alpha(0) = p$, we define the corresponding derivation

$$\alpha'(0) : D \rightarrow \mathbb{R}$$

by

$$\alpha'(0)(f) = \frac{d}{dt} \bigg|_{t=0} (f \circ \alpha)(t).$$

**Exercise 1.11.** Show that $\alpha'(0)$ is a derivation at $p$, so that $\alpha'(0) \in T_pM$. 

$\alpha'(0)$ being linear means that for any $c \in \mathbb{R}$ and $f, g \in D$, $X(cf + g) = cX(f) + X(g)$. 

---

\text{(1.2) $X(\text{const}) = X(\text{const} \cdot 1) = \text{const} X(1) = 0$.} 

\text{(1.3) $\alpha'(0)(f) = \frac{d}{dt} \bigg|_{t=0} (f \circ \alpha)(t)$.} 

---
SOLUTION. The verification of linearity is elementary. As for the product rule, we compute

\[ \alpha' (0) (fg) = \left. \frac{d}{dt} \right|_{t=0} ((fg) \circ \alpha) (t) \]
\[ = \left. \frac{d}{dt} \right|_{t=0} ((f \circ \alpha) (t) \cdot (g \circ \alpha) (t)) \]
\[ = (f \circ \alpha) (0) \left. \frac{d}{dt} \right|_{t=0} ((g \circ \alpha) (t)) + (g \circ \alpha) (0) \left. \frac{d}{dt} \right|_{t=0} ((f \circ \alpha) (t)) \]
\[ = f (p) \alpha' (0) (g) + g (p) \alpha' (0) (f) . \]

**Definition 1.12.** We may also define the tangent space \( T_p M \) to be the set of \( \alpha' (0) \), where \( \alpha : (-\varepsilon, \varepsilon) \to M \) is a differentiable curve with \( \alpha (0) = p \).

The following, which we will prove later, implies that the two definitions of \( T_p M \) are the same.

**Theorem 1.13.** Let \( M \) be a \( C^2 \) manifold. If \( X \) is a derivation at \( p \), then there exists a differentiable curve \( \alpha \) such that \( \alpha' (0) = X \).

**2.2. The tangent space is \( n \)-dimensional.**

Manifolds are defined by parametrizations, so it is natural to understand the tangent space this way. Let \( x : U \to M \) be a parametrization with \( p \in x (U) \). Let \( e_i = (0, \ldots, 1, 0, \ldots, 0) \) for \( 1 \leq i \leq n \). \( \{ e_i \}_{i=1}^n \) is called the \( i \)th **standard basis** for \( \mathbb{R}^n \). Consider the Euclidean coordinate lines

\[ \alpha_i : t \mapsto x^{-1} (p) + te_i . \]  

(1.4)

Pushing these curves forward by the parametrization \( x \), we obtain the **coordinate curves**:

\[ c_i (t) = x (\alpha_i (t)) = x (x^{-1} (p) + te_i) . \]

(1.5)

Since \( c_i \) is differentiable and \( c_i (0) = p \), \( c_i' (0) : \mathcal{D} \to \mathbb{R} \) is a tangent vector at \( p \), called a **coordinate tangent vector**.

**Theorem 1.14.** Let \( M \) be a differentiable manifold. With Definition 1.12 of the tangent space, \( \{ c_i' (0) \}_{i=1}^n \) is a basis for \( T_p M \). In particular, \( T_p M \) is \( n \)-dimensional.
Proof. We use the parametrization $x$ to understand an arbitrary tangent vector $\alpha' (0)$. We compute

$$
\alpha' (0) (f) = \frac{d}{dt} \bigg|_{t=0} ((f \circ \alpha) (t))
$$

$$
= \frac{d}{dt} \bigg|_{t=0} \left( (f \circ x \circ (x^{-1} \circ \alpha)) (t) \right)
$$

$$
= d (f \circ x)_{x^{-1}(p)} \left( \frac{d}{dt} \bigg|_{t=0} ((x^{-1} \circ \alpha) (t)) \right),
$$

where the Euclidean differential $d (f \circ x)_{x^{-1}(p)}$ is defined as in (1.1). Note that $(x^{-1} \circ \alpha) (0) = x^{-1} (p)$.

Let

$$
\frac{d}{dt} \bigg|_{t=0} ((x^{-1} \circ \alpha) (t)) \div v = (v^1, \ldots, v^n) \in \mathbb{R}^n.
$$

Then we may write

$$
\alpha' (0) (f) = d (f \circ x)_{x^{-1}(p)} (v)
$$

for any $f \in \mathcal{D}$ and differentiable curve $\alpha$. So given a parametrization $x$, the derivation $\alpha' (0)$ depends only on $\frac{d}{dt} \bigg|_{t=0} ((x^{-1} \circ \alpha) (t)) = v \in \mathbb{R}^n$. This indicates that $T_p\mathcal{M}$ should be $n$-dimensional. We now prove this rigorously.

Now we show that $\alpha' (0)$ may be written as a linear combination of the $c'_i (0)$. Using $v = \sum_{i=1}^n v^i e_i$ and the linearity of $d (f \circ x)_{x^{-1}(p)}$, we compute that

$$
\alpha' (0) (f) = d (f \circ x)_{x^{-1}(p)} (v)
$$

$$
= \sum_{i=1}^n v^i d (f \circ x)_{x^{-1}(p)} (e_i)
$$

$$
= \sum_{i=1}^n v^i \left. \frac{d}{dt} \right|_{t=0} f \left( x \left( x^{-1} (p) + te_i \right) \right)
$$

$$
= \sum_{i=1}^n v^i \left. \frac{d}{dt} \right|_{t=0} (f \circ c_i) (t)
$$

$$
= \sum_{i=1}^n v^i c_i' (0) (f).
$$

Since this true for each $f \in \mathcal{D}$, we conclude that

$$
\alpha' (0) = \sum_{i=1}^n v^i c_i' (0),
$$
where \( v^i = \frac{d}{dt} \bigg|_{t=0} \left( (x^{-1})^i (\alpha (t)) \right) \). This shows that \( \{c_i'(0)\}_{i=1}^n \) spans the tangent space \( T_pM \) in Definition 1.12.

To see that the \( c_i'(0) \) are linearly independent, suppose that \( a^i \in \mathbb{R} \) are such that
\[
\sum_{i=1}^n a^i c_i'(0) = 0.
\]

Then choosing \( f = (x^{-1})^j \), we obtain
\[
0 = \sum_{i=1}^n a^i c_i'(0) \left( (x^{-1})^j \right).
\]

Now
\[
c_i'(0) \left( (x^{-1})^j \right) = \frac{d}{dt} \bigg|_{t=0} \left( (x^{-1})^j \circ c_i \right) (t)
= \frac{d}{dt} \bigg|_{t=0} \left( \delta_i^j t \right)
= \delta_i^j.
\]

Hence \( 0 = \sum_{i=1}^n a^i \delta_i^j = a^j \). This completes the proof of Theorem 1.14. □

**Exercise 1.15.** Let \( x : U \to M \) be a parametrization, \( p \in x(U) \), and \( c_i(t) = x \left( (x^{-1})(p) + te_i \right) \). Show that the derivation \( c_i'(0)_p \) at \( p \) satisfies
\[
c_i'(0)_p (f) = \frac{\partial (f \circ x)}{\partial q^i} (x^{-1}(p))
\]
for any \( C^\infty \) function \( f \) defined in a neighborhood of \( p \).

**Solution.** We compute
\[
c_i'(0)_p (f) = \frac{d}{dt} \bigg|_{t=0} (f \circ c_i) (t)
= \frac{d}{dt} \bigg|_{t=0} (f \circ x) \left( (x^{-1})(p) + te_i \right)
= \sum_{j=1}^n \frac{\partial (f \circ x)}{\partial q^j} \left( (x^{-1})(p) \right) \cdot \frac{d}{dt} \bigg|_{t=0} \left( (x^{-1})(p) + te_i \right)^j
= \frac{\partial (f \circ x)}{\partial q^i} \left( (x^{-1})(p) \right)
\]
since \( \frac{d}{dt} \bigg|_{t=0} \left( (x^{-1})(p) + te_i \right)^j = \delta_i^j \).

We may identify a vector \( v \in \mathbb{R}^n \) with the tangent vector at a point \( q \in \mathbb{R}^n \) defined by the derivation \( \alpha'(0) \), where \( \alpha(t) = q + tv \). In this way
we have a natural vector space isomorphism between $\mathbb{R}^n$ and $T_q\mathbb{R}^n$ for any $q \in \mathbb{R}^n$.

2.3. Equivalence of the two definitions of tangent space.

Now we return to Definition 1.9 of tangent space as the set of all derivations at a point. We prove that each derivation arises from a curve.

**Proof of Theorem 1.13.** Let $\mathcal{M}$ be a $C^2$ manifold. Let $X$ be a derivation at $p$, that is, $X : \mathcal{D}_p \rightarrow \mathbb{R}$ is a linear function such that

$$X (fg) = f (p) X (g) + g (p) X (f).$$

Let $x : \mathcal{U} \rightarrow \mathcal{M}$ be a parametrization with $x (0) = p$. We shall use the following fact (proved below):

If $f \in \mathcal{C}^{2, \text{loc}}_p$, then there exist constants $a_i \in \mathbb{R}$ and functions $g_i \in \mathcal{C}^{1, \text{loc}}_p$ with $g_i (p) = 0$ such that

$$f (q) = f (p) + \sum_{i=1}^n a_i (x^{-1})^i (q) + \sum_{i=1}^n g_i (q) (x^{-1})^i (q).$$

Let $X^i = X ((x^{-1})^i) \in \mathbb{R}$. Since $X$ is a derivation, we compute for $f \in \mathcal{C}^{2, \text{loc}}_p$ that

$$X (f) = \sum_{i=1}^n X \left( a_i (x^{-1})^i + g_i (x^{-1})^i \right)$$

$$= \sum_{i=1}^n (a_i + g_i (p)) X \left( (x^{-1})^i \right) + (x^{-1})^i (p) X (g_i)$$

$$= \sum_{i=1}^n a_i X \left( (x^{-1})^i \right)$$

$$= \sum_{i=1}^n a_i X^i,$$

where we used $g_i (p) = 0$ and $(x^{-1})^i (p) = 0.
On the other hand,
\[
\left( \sum_{i=1}^{n} X^i c'_i (0) \right) (f) = \sum_{i=1}^{n} X^i c'_i (0) (f) \\
= \sum_{i=1}^{n} X^i \sum_{j=1}^{n} a_j c'_i (0) (x^{-1})^j \\
= \sum_{i,j=1}^{n} X^i a_j \delta^j_i \\
= \sum_{i=1}^{n} X^i a_i
\]
since \( c'_i (0) (x^{-1})^j = \delta^j_i \). Since the above formulas are true for any \( f \in \mathcal{D}_p \), we conclude that \( X = \sum_{i=1}^{n} X^i c'_i (0) \) when acting on \( f \in C^2_p \), where \( X^i = X \left( (x^{-1})^i \right) \in \mathbb{R} \). By Lemma 1.16 below,
\[
X = \sum_{i=1}^{n} X^i c'_i (0)
\]
acting on \( f \in \mathcal{D}_p \). Note that \( X = \alpha' (0) \), where
\[
\alpha (t) = x \left( t \left( X^1, \ldots, X^n \right) \right).
\]

Proof of (1.6): By precomposing with the map \( x \), (1.6) is equivalent to:
\[
h (v) = h (0) + \sum_{i=1}^{n} a_i v^i + \sum_{i=1}^{n} k_i (v) v^i
\]
for \( v = (v^1, \ldots, v^n) \in U \subset \mathbb{R}^n \), where \( h = f \circ x \in C^1 \) and \( k_i = g_i \circ x \in C^1 \) and where \( k_i (0) = 0 \).

We shall prove that if \( h : U \rightarrow \mathbb{R} \) is \( C^2 \), then there exist a \( C^1 \) function \( k_i \) with \( k_i (0) = 0 \) such that
\[
(1.7) \quad h (v) = h (0) + \sum_{i=1}^{n} \frac{\partial h}{\partial x^i} (0) v^i + \sum_{i=1}^{n} k_i (v) v^i.
\]
3. Differential of a map and tangent bundle

Proof of (1.7): We have
\[
   h(v) - h(0) = \int_0^1 \frac{d}{dt} h(tv) dt = \int_0^1 \sum_{i=1}^n v^i \frac{\partial h}{\partial x^i}(tv) dt = \sum_{i=1}^n v^i \frac{\partial h}{\partial x^i}(0) + \sum_{i=1}^n v^i \int_0^1 \left( \frac{\partial h}{\partial x^i}(tv) - \frac{\partial h}{\partial x^i}(0) \right) dt.
\]
Define
\[
   g_i(v) = \int_0^1 \left( \frac{\partial h}{\partial x^i}(tv) - \frac{\partial h}{\partial x^i}(0) \right) dt.
\]
Since \( \varphi(v) = \frac{\partial h}{\partial x^i}(tv) - \frac{\partial h}{\partial x^i}(0) \) is \( C^1 \) in a neighborhood of 0, we have \( g_i \in C^1_{0,\text{loc}} \) and \( g_i(0) = 0 \).

The one loose end we have left is to prove the following.

**Lemma 1.16.** Let \( M \) be a \( C^2 \) manifold. If \( X \) and \( Y \) are derivations at \( p \) such that \( X \) equals \( Y \) when restricted to \( C^2_{p,\text{loc}} \), then \( X = Y \) (on all of \( D_p \)).

3. Differential of a map and tangent bundle

3.1. Differential of a map.

Let \( M \) and \( N \) be differentiable manifolds. A continuous map \( \varphi : M \to N \) is **differentiable** at \( p \in M \) if for any parametrizations \((U, x)\) of \( M \) and \((V, y)\) of \( N \) we have that \( y^{-1} \circ \varphi \circ x \) is (a vector-valued function of an open subset of Euclidean space) differentiable at \( x^{-1}(p) \).

Let \( \varphi : M \to N \) be a map which is differentiable at \( p \). Its **differential** at \( p \) is defined by
\[
   d\varphi_p : T_p M \to T_{\varphi(p)} N,
\]
where
\[
   (1.8) \quad d\varphi_p (\alpha' (0)) = \left( \varphi \circ \alpha \right)' (0)
\]
for any differentiable curve \( \alpha : (-\varepsilon, \varepsilon) \to M \) with \( \alpha (0) = p \).

**Exercise 1.17.** Show that the differential of \( \varphi \) is well defined, i.e., if \( \alpha \) and \( \beta \) are differentiable curves such that \( \alpha'(0) = \beta'(0) \), then \( (\varphi \circ \alpha)'(0) = (\varphi \circ \beta)'(0) \).

The differential satisfies (composition rule)
\[
   d(\varphi \circ \psi)_p = d\varphi_{\varphi(p)} \circ d\psi_p.
\]
1. Differentiable Manifolds

If $\phi$ and its inverse are differentiable, then

$$d (\phi^{-1})_{\phi(p)} = (d\phi_p)^{-1}.$$ 

We leave the verification of these properties, which follow from their Euclidean counterparts, to the reader.

3.2. Diffeomorphisms.

Let $M^n$ and $N^n$ be $C^\infty$ differentiable manifolds. A bijective map $\phi : M \to N$ is called a diffeomorphism if both $\phi$ and $\phi^{-1}$ are $C^\infty$. Given $p \in M$, a local diffeomorphism at $p$ is a diffeomorphism $\psi : U \to V$, where $U$ is a neighborhood of $p$ in $M$ and $U$ is a neighborhood of $\psi(p)$ in $N$.

**Proposition 1.18** (Inverse function theorem). If a $C^\infty$ map $\phi : M \to N$ and $p \in M$ are such that $d\phi_p$ is an isomorphism, then there exist neighborhoods $U$ of $p$ in $M$ and $V$ of $\phi(p)$ in $N$ such that $\phi|_U : U \to V$ is a diffeomorphism, i.e., $\phi$ is a local diffeomorphism at $p$.

3.3. Tangent bundle.

Given a differentiable manifold $M^n$, we can bundle the tangent spaces $T_pM$, $p \in M$, together to form the tangent bundle. As a point-set, it is defined by the disjoint union

$$TM = \bigcup_{p \in M} T_pM.$$ 

Equivalently, we may define it as the point-set

$$TM = \{(p, v) : p \in M, \ v \in T_pM\}.$$ 

We have the projection map $\pi : TM \to M$ defined by $\pi(p, v) = p$. Note that $\pi^{-1}(p) = T_pM$.

We shall now give the tangent bundle $TM$ a manifold structure. Given a parametrization $(U, x)$ of $M$, i.e., $x : U \subset \mathbb{R}^n \to x(U) \subset M$, at each point $p \in x(U)$ we have the basis of tangent vectors

$$\frac{\partial}{\partial x^i} p = c_i(0)_p,$$

where $c_i(t) = x(x^{-1}(p) + te_i)$. The subscript $p$ is used to denote that the tangent vector $\frac{\partial}{\partial x^i}$ is at $p$.

The coordinate tangent vectors are the push forwards by the parametrization of the standard basis vectors of $\mathbb{R}^n$. 
Lemma 1.19. Let $e_i \doteq (0, \ldots, 0, 1, 0, \ldots, 0)$ for $i = 1, \ldots, n$. Considering $\mathcal{U}$ as a differentiable manifold (it is an open subset of $\mathbb{R}^n$), we have

\begin{equation}
\frac{\partial}{\partial x^i} x(q) = (dx)_q (e_i),
\end{equation}

where $e_i \in T_q \mathbb{R}^n$, for each $q \in \mathcal{U}$ and $i$.

\textbf{Proof.} Given $q \in \mathcal{U}$ and $i$, define the Euclidean coordinate line $\alpha_i : (-\varepsilon, \varepsilon) \to \mathcal{U}$ by

$\alpha_i(t) \doteq q + te_i$.

Then $\alpha_i'(0) = e_i$ and

$c_i(t) = x(\alpha_i(t))$.

By the definition of differential (1.8),

$\left(dx\right)_q (e_i) = \left(dx\right)_q (\alpha_i'(0)) = (x \circ \alpha_i)'(0) = c_i'(0)_{x(q)} = \frac{\partial}{\partial x^i x(q)}$.

By taking linear combinations of the formula in the above lemma, we obtain an expression for $(dx)_q$.

Lemma 1.20.

$$(dx)_q(u) = \sum_{i=1}^{n} u^i \frac{\partial}{\partial x^i x(q)}.$$ 

\textbf{Proof.} Using the linearity of $(dx)_q$, we compute

$$(dx)_q(u) = (dx)_q \left( \sum_{i=1}^{n} u^i e_i \right)$$

$$= \sum_{i=1}^{n} u^i (dx)_q(e_i)$$

$$= \sum_{i=1}^{n} u^i \frac{\partial}{\partial x^i x(q)}$$

with the last equality by Lemma 1.19.  

Note that for each $q \in \mathcal{U}$, $(dx)_q : T_q \mathbb{R}^n = \mathbb{R}^n \to T_{x(q)} \mathcal{M}$ is a vector space isomorphism.

Corresponding to $(\mathcal{U}, x)$ define a parametrization $(\mathcal{V}, y)$ of the tangent bundle $T\mathcal{M}$ by

$y : \mathcal{V} \doteq \mathcal{U} \times \mathbb{R}^n \to T\mathcal{M},$
where
\[ y(q, u) \triangleq (x(q), (dx)_q(u)) \]
for \( q \in \mathcal{U} \) and \( u \in \mathbb{R}^n \). First note that since \( x : \mathcal{U} \to x(\mathcal{U}) \) is a bijection and each \((dx)_q\) is an isomorphism, \( y : \mathcal{U} \times \mathbb{R}^n \to \pi^{-1}(x(\mathcal{U})) \) is a bijection.

**Lemma 1.21.** Let \( \{(\mathcal{U}_\alpha, x_\alpha)\}_{\alpha \in \Lambda} \) be a \( C^{k+1} \) differentiable manifold structure on \( \mathcal{M} \). Then the maximal extension of the collection of corresponding parametrizations \( \{(\mathcal{V}_\alpha, y_\alpha)\}_{\alpha \in \Lambda} \) is a \( C^k \) differentiable manifold structure on \( T\mathcal{M} \).

**Proof.** Since \( \mathcal{M} = \bigcup_{\alpha \in \Lambda} x_\alpha(\mathcal{U}_\alpha) \) and since \( y_\alpha(\mathcal{V}_\alpha) = \pi^{-1}(x_\alpha(\mathcal{U}_\alpha)) \), we have \( T\mathcal{M} = \bigcup_{\alpha \in \Lambda} y_\alpha(\mathcal{V}_\alpha) \). Next, we need to show that the transition maps are differentiable. By definition, we have
\[
y_\alpha(q, u) = (x_\alpha(q), (dx)_q(u))
\]
for \( q \in \mathcal{U}_\alpha \) and \( u \in \mathbb{R}^n \). Thus for any \( \alpha, \beta \in \Lambda \) such that \( x_\alpha(\mathcal{U}_\alpha) \cap x_\beta(\mathcal{U}_\beta) \neq \emptyset \), the transition function \( y_\beta^{-1} \circ y_\alpha \) is given by
\[
y_\beta^{-1} \circ y_\alpha(q, u) = ((x_\beta^{-1} \circ x_\alpha)(q), (dx)_q^{-1} \circ (dx)_q(u)) = ((x_\beta^{-1} \circ x_\alpha)(q), d(x_\beta^{-1} \circ x_\alpha)_q(u))
\]
since \((dx_\beta)^{-1} = d(x_\beta^{-1})\) and \(d(x_\beta^{-1} \circ x_\alpha) = d(x_\beta^{-1}) \circ dx_\alpha\). Since \( \mathcal{M} \) is \( C^{k+1} \), we have that \( x_\beta^{-1} \circ x_\alpha \in C^{k+1} \) and hence \( y_\beta^{-1} \circ y_\alpha \in C^k \). \( \square \)

Denote \( \varphi = x_\beta^{-1} \circ x_\alpha \), which maps an open subset of \( \mathbb{R}^n \) diffeomorphically to an open subset of \( \mathbb{R}^n \). Then
\[
y_\beta^{-1} \circ y_\alpha(q, u) = (\varphi(q), d\varphi_q(u)).
\]
Then the differential of the transition function \( y_\beta^{-1} \circ y_\alpha \) is given by
\[
(1.10) \quad d \left( y_\beta^{-1} \circ y_\alpha \right) \begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} \frac{d\varphi_q(v)}{v} \\ d\varphi_q(w) + (\text{Hess } \varphi)_q(u, v) \end{pmatrix} \\
= \begin{pmatrix} \frac{d\varphi_q}{v} \\ (\text{Hess } \varphi)_q(u, v) \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix}.
\]
Here \((\text{Hess } \varphi)_q(u, v) \triangleq \left. \frac{d}{dt} \right|_{t=0} d\varphi_{q+tv}(u)\) is the **Hessian** (or second derivative) of \( \varphi \).

A differentiable manifold \( \mathcal{N} \) is **orientable** if \( \mathcal{N} \) has a (not maximal) differentiable structure \( \{(\mathcal{V}_\alpha, y_\alpha)\} \) such that the differentials of the transition maps \( y_\beta^{-1} \circ y_\alpha \) have positive determinant.

**Exercise 1.22.** Using formula (1.10), prove that the tangent bundle \( T\mathcal{M} \) is orientable.
Solution (sketch). We have
\[
\det \left( d \left( y^{-1} \circ y \right) \right) = \det \begin{pmatrix} d\varphi_q & 0 \\ (\text{Hess } \varphi)_q (u, \cdot) & d\varphi_q \end{pmatrix} = (\det (d\varphi_q))^2 > 0.
\]

4. Critical and regular points of functions and level surfaces

Let \( M^m \) and \( N^n \) be \( C^\infty \) manifolds and let \( \varphi : M \to N \) be a \( C^\infty \) map. We say that \( p \in M \) is a critical point of \( \varphi \) if the differential \( d\varphi_p : T_p M \to T_{\varphi(p)} N \) is not surjective. In this case, \( \varphi(p) \) is called a critical value of \( \varphi \). Note that if \( m < n \), then each point \( p \in M \) is a critical point of any map \( \varphi \).

We say that \( q \in N \) is a regular value of \( \varphi \) if it is not a critical value of \( \varphi \). In particular, if \( q \notin \varphi(M) \), then \( q \) is a regular value of \( \varphi \).

**Proposition 1.23.** Suppose \( m \geq n \). If \( q \in N \) is a regular value of \( \varphi \), then the set

\[ Q \triangleq \varphi^{-1}(q) = \{ p \in M : \varphi(p) = q \} \subset M \]

has the following property. For each \( p \in Q \) there exists a neighborhood \( V \) of \( p \) in \( M \) and a \( C^\infty \) homeomorphism

\[ x : U \subset \mathbb{R}^{m-n} \to Q \cap V \subset M \]

such that \( (dx)_u : T_u U = \mathbb{R}^{m-n} \to T_{x(u)} M \) is injective for each \( u \in U \). We call \( Q \) a regular submanifold of dimension \( m-n \).

**Exercise 1.24.** Prove that \( Q \) is an \((m-n)\)-dimensional manifold.

Define \( f : \mathbb{R}^{n+1} \to \mathbb{R} \) by \( f(x) = |x|^2 \). One sees that

\[ df_x = \nabla f_x = \left( \frac{\partial f}{\partial x^1}, \frac{\partial f}{\partial x^2}, \ldots, \frac{\partial f}{\partial x^{n+1}} \right) \]

is given by \( df_x = 2x \). As a map, \( df_x \) is

\[ df_x(v) = \sum_{i=1}^{n+1} \frac{\partial f}{\partial x^i} v^i \]

for \( v = (v^1, v^2, \ldots, v^{n+1}) \in \mathbb{R}^{n+1} \). So \( x \) is a critical point of \( f \) if and only if \( x = 0 \). We conclude that if \( r > 0 \), then \( f^{-1}(r^2) = S^n(r) \), i.e., the sphere of radius \( r \), is a a regular submanifold of dimension \( n \).

5. Vector fields, Lie brackets and Lie derivatives

5.1. Vector fields.

Much of the study of differential and Riemannian geometry is in the \( C^\infty \) category. For convenience, we now let \( M^n \) be a \( C^\infty \) differentiable manifold, so that the transition maps \( x_β^{-1} \circ x_α \) are \( C^\infty \).
Definition 1.25. A **vector field** \( X \) on \( \mathcal{M} \) is an assignment to each point \( p \in \mathcal{M} \) a tangent vector \( X_p \equiv X(p) \in T_p\mathcal{M} \).

Equivalently, if \( \pi : T\mathcal{M} \to \mathcal{M} \) denotes the projection map, then a vector field is a map \( X : \mathcal{M} \to T\mathcal{M} \) satisfying \( \pi \circ X = \text{id}_\mathcal{M} \), the identity map of \( \mathcal{M} \).

If \( \mathcal{M}^n \) is a \( C^\infty \) differentiable manifold and if \( \ell \geq 0 \), then we say that \( X \) is a \( C^\ell \) vector field if with respect to any \( C^\infty \) parametrization \((U, x)\) we have that each \( X^i : U \to \mathbb{R} \), defined by

\[
X^i(p) = X_p((x^{-1})^i),
\]
is a \( C^\ell \) function for \( i = 1, \ldots, n \).

**Exercise 1.26.** Show that for any \( p \in x(U) \),

\[
X(p) = \sum_{i=1}^{n} X^i(p) \frac{\partial}{\partial x^i_p}.
\]

Let \( f : U \subset \mathbb{R}^n \to \mathbb{R} \) be a \( C^\infty \) function on an open set. The **gradient vector field** of \( f \) is

\[
\nabla f = \text{grad } f = \left( \frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^n} \right).
\]

For example, if \( U = \mathbb{R}^n \) and \( f(x) = |x|^2 = (x^1)^2 + (x^2)^2 + \cdots + (x^n)^2 \), then

\[
\nabla f = 2x.
\]

If \( X = (X^1, X^2) \) is a vector field on \( \mathbb{R}^2 \), define the vector field \( J(X) \) by

\[
J(X) = (-X^2, X^1).
\]

The underlying transformation \( J : \mathbb{R}^2 \to \mathbb{R}^2 \) is called an **almost complex structure**. Note that \( \langle X, J(X) \rangle = 0 \). Also,

\[
J \left( \nabla |x|^2 \right) = J \left( 2 (x^1, x^2) \right) = 2 (-x^2, x^1).
\]

If \( X = (X^1, \ldots, X^n) \) is a \( C^\infty \) vector field on \( U \subset \mathbb{R}^n \), then its **divergence** is the function defined by

\[
\text{div } X = \sum_{i=1}^{n} \frac{\partial X^i}{\partial x^i}.
\]

Note that if \( f : U \subset \mathbb{R}^n \to \mathbb{R} \) is a \( C^\infty \) function, then its Laplacian is given by

\[
\Delta f = \sum_{i=1}^{n} \frac{\partial}{\partial x^i} \left( \nabla f \right)^i = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial (x^i)^2}.
\]
Exercise 1.27. Compute $\Delta |x|^r$ on $\mathbb{R}^n$ (possibly minus the origin) for $r \in \mathbb{R}$. Does $r = 2 - n$ for $n \geq 3$ give you a special answer.

5.2. Lie brackets of vector fields.

Definition 1.28. The Lie bracket of two $C^\infty$ vector fields $X$ and $Y$ is defined, as a derivation at each point in $\mathcal{M}$, by

\[(1.11) \quad [X,Y](f) = X(Y(f)) - Y(X(f))\]

for $C^\infty$ any function $f : \mathcal{M} \to \mathbb{R}$.

We need to show that the Lie bracket is well defined, namely that it is a derivation at each point.

Let $C^\infty(\mathcal{M})$ denote the set of $C^\infty$ functions on $\mathcal{M}$. Note that for a vector field $X$ and a function $f$, we have $X(f)$ is a function. At a point $p \in \mathcal{M}$ it is easy to check that (exercise)

\[ [X,Y]_p : C^\infty(\mathcal{M}) \to \mathbb{R} \]

is an $\mathbb{R}$-linear function, i.e.,

\[ [X,Y]_p(cf + g) = c[X,Y]_p(f) + [X,Y]_p(g). \]

Thus, to show that $[X,Y]_p$ is a derivation at $p$ it is enough to verify the product rule. We compute for any $f, g \in C^\infty(\mathcal{M})$ and $p \in \mathcal{M}$,

\[ [X,Y]_p(fg) = X_p(Y(fg)) - Y_p(X(fg)) \]

\[ = X_p( fY(g) + gY(f)) - Y_p(fX(g) + gX(f)) \]

\[ = f(p)X_p(Y(g)) + X_p(fY(g)) + X_p(Y(f)) + g(p)Y_p(f)X_p(g) + Y_p(f)X_p(g) + Y_p(g)X_p(f) \]

\[ - f(p)Y_p(X(g)) - Y_p(fX(g)) - Y_p(f)X_p(g) - Y_p(g)X_p(f) \]

\[ = f(p)(X_p(Y(g)) - Y_p(X(g))) + g(p)(X_p(Y(f)) - Y_p(X(f))) \]

\[ = f(p)[X,Y]_p(g) - g(p)[X,Y]_p(f). \]

Since $[X,Y]_p$ is $\mathbb{R}$-linear on $C^\infty(\mathcal{M})$ and since $[X,Y]_p(fg) = f(p)[X,Y]_p(g) - g(p)[X,Y]_p(f)$, we conclude that $[X,Y]_p$ is a derivation at $p$ and that the Lie bracket $[X,Y]$ is a vector field on $\mathcal{M}$.

It is clear that

\[ [Y,X] = -[X,Y]. \]

Exercise 1.29. Let $U$ be an open subset of $\mathbb{R}^n$. Show that for vector fields $V$ and $W$ on $U$ we have

\[ [V,W] = V(W) - W(V). \]

$^2$To simplify the way the expressions look, you may omit the parentheses and the subscripts $p$ in the notation.
Exercise 1.30. Show that if $V$ and $W$ are constant vector fields on $\mathbb{R}^n$, then their Lie bracket is zero: $[V, W] = 0$. In particular, $[e_i, e_j] = 0$ on $\mathbb{R}^n$.

Lemma 1.31. Let $(\mathcal{U}, x)$ be a parametrization of $\mathcal{M}$. Then the Lie bracket of any pair of coordinate vector fields is zero: $[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}] = 0$ on $x(\mathcal{U})$ for any $i, j = 1, \ldots, n$.

Proof. For any $p \in x(\mathcal{U})$ and $f \in C^\infty(\mathcal{M})$,
\[
c_i'(0)_p (f) = \frac{\partial (f \circ x)}{\partial q^i} (x^{-1}(p))
\]
(this is an exercise). Hence, as functions on $x(\mathcal{U})$,
\[
c_i'(0)(f) = \frac{\partial (f \circ x)}{\partial q^i} \circ x^{-1}.
\]
We then have
\[
c_i'(0)(c_j'(0)(f)) = \frac{\partial \left( \left( c_j'(0)(f) \right) \circ x \right)}{\partial q^i} \circ x^{-1}
\]
\[
= \frac{\partial (f \circ x)}{\partial q^i} \circ x^{-1}
\]
\[
= \frac{\partial^2 (f \circ x)}{\partial q^i \partial q^j} \circ x^{-1}.
\]
Hence
\[
[c_i'(0), c_j'(0)](f) = c_i'(0)(c_j'(0)(f)) - c_j'(0)(c_i'(0)(f))
\]
\[
= \frac{\partial^2 (f \circ x)}{\partial q^i \partial q^j} \circ x^{-1} - \frac{\partial^2 (f \circ x)}{\partial q^j \partial q^i} \circ x^{-1}
\]
\[
= 0
\]
by Clairaut’s theorem (mixed partial derivatives of a Euclidean function are equal), which holds since $f \circ x : \mathcal{U} \to \mathbb{R}$ is a function on an open subset of $\mathbb{R}^n$.

Remark 1.32. Later we shall prove a naturality property of the Lie bracket under the action by differentials, which as a special case, gives
\[
\left[ \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right]_p = \left[ (dx)_{x^{-1}(p)} (e_i), (dx)_{x^{-1}(p)} (e_j) \right]
\]
\[
\overset{\text{nat}}{=} (dx)_{x^{-1}(p)} ([e_i, e_j])
\]
\[
= (dx)_{x^{-1}(p)} (0)
\]
\[
= 0,
\]
where $\overset{\text{nat}}{=}$ denotes the naturality equality.
Exercise 1.33. Prove that for any $C^\infty$ vector fields $X$ and $Y$ and $C^\infty$ functions $g$ and $h$, 

\begin{equation}
[gX, hY] = g(Xh)Y - h(Yg)X + gh[X,Y].
\end{equation}

Solution. Acting as a derivation on a $C^\infty$ function $f : \mathcal{M} \to \mathbb{R}$, we have

\[
[gX, hY] f = g X (hYf) - h Y (gXf)
= g (XhYf + hX(Yf)) - hYgXf - hgY(Xf)
= g(Xh)Yf - h(Yg)Xf + gh[X,Y] f.
\]

Let $X = \sum_i X^i \frac{\partial}{\partial x^i}$ and $Y = \sum_j Y^j \frac{\partial}{\partial x^j}$. Using (1.12) and the fact that $[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}] = 0$, we compute

\[
[X, Y] = \left( \sum_i X^i \frac{\partial}{\partial x^i}, \sum_j Y^j \frac{\partial}{\partial x^j} \right)
= \sum_{i,j} \left( X^i \left( \frac{\partial}{\partial x^j} Y^j \right) - Y^j \left( \frac{\partial}{\partial x^i} X^i \right) \right)
= \sum_{i,j} \left( X^i \frac{\partial Y^j}{\partial x^i} \frac{\partial}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \frac{\partial}{\partial x^i} \right),
\]

that is,

\begin{equation}
[X, Y] = \sum_{i,j=1}^n \left( X^i \frac{\partial Y^j}{\partial x^i} - Y^j \frac{\partial X^i}{\partial x^j} \right) \frac{\partial}{\partial x^j}.
\end{equation}

The Jacobi identity says that for $C^\infty$ vector fields $X$, $Y$, and $Z$, we have

\begin{equation}
[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.
\end{equation}

We may see this as follows. Acting as derivations on a function $f : \mathcal{M} \to \mathbb{R}$, we have

\[
[X, [Y, Z]] (f) = X ((Y, Z)(f)) - [Y, Z](X(f))
= X (Y (Z(f)) - Z(Y(f))) - Y (Z(X(f))) + Z(Y(X(f)))
= XYZ f - XZ f - YZ f + YX f + ZY f + XY f,
\]

where in the last line we omitted the () in the notation for the sake of brevity.

By cyclically permuting the $X, Y, Z$, we obtain

\[
[Z, [X, Y]] f = ZXY f - ZYX f - XYZ f + YXZ f
\]

and

\[
[Y, [Z, X]] = YZ f - YZX f - ZYX f + XYZ f.
\]
Summing these three equations, we obtain 6 pairs of cancelling terms (each permutation of \(XYZ\) appears with both a plus and a minus. So we get the Jacobi identity.

A real (or complex) Lie algebra is a real (or complex) vector space \(g\) with a binary operation \([\ , \] : g \times g \to g\), called the Lie bracket, with the properties that for any \(c \in \mathbb{R}\) (or \(\mathbb{C}\)) and \(u, v, w \in g\),

1. (alternating) \([v, u] = -[u, v]\),
2. (bilinear) \([u + cv, w] = [u, w] + c[v, w]\),
3. (Jacobi identity) \([[[u, v], w] + [[[v, w], u] + [[[w, u], v] = 0.\)

The set \(\Xi(\mathcal{M})\) of \(C^\infty\) vector fields on \(\mathcal{M}\) with the Lie bracket defined by (1.11) is a Lie algebra.

For a differentiable curve \(\alpha : (-\varepsilon, \varepsilon) \to \mathcal{M}\), let \(\alpha'(t)\) be its corresponding derivation at \(\alpha(t)\), defined by

\[\alpha'(t)(f) = (f \circ \alpha)'(t)\]

for \(f \in C^\infty(\mathcal{M})\).

Lemma 1.34. Let \(\mathcal{M}\) be a differentiable manifold and let \(X\) be a differentiable vector field on \(\mathcal{M}\).

Then for each \(p \in \mathcal{M}\) there exists \(\delta > 0\), a neighborhood \(U\) of \(p\), and a differentiable map

\[\varphi : (-\delta, \delta) \times U \to \mathcal{M}\]

such that for each \(q \in U\) the differentiable curve \(\gamma_q : (-\delta, \delta) \to \mathcal{M}\) defined by

\[\gamma_q(t) = \varphi(t, q)\]

satisfies \(\gamma_q(0) = q\) and

\[\gamma_q'(t) = X(\gamma_q(t))\]

for \(t \in (-\delta, \delta)\).

Moreover \(\gamma_q\) is unique in the sense that if \(\alpha : (-\delta, \delta) \to \mathcal{M}\) is a differentiable curve such that \(\alpha(0) = q\) and

\[\alpha'(t) = X(\alpha(t))\]

for \(t \in (-\delta, \delta)\), then \(\alpha = \gamma_q\). We call \(\gamma_q\) an integral curve (or trajectory) to the vector field \(X\).

The essence of the proof of this, which we omit, is the existence and uniqueness theorem for systems of ordinary differential equations (ODE). By pulling back the problem by \(x : U \to \mathcal{M}\) from \(\mathcal{M}\) to \(U \subset \mathbb{R}^n\), we may assume that \(\mathcal{M}\) is an open subset of \(\mathbb{R}^n\). Given a differentiable vector field
6. The Lie derivative on vector fields

$X$ on an open set $U \subset \mathbb{R}^n$ and given $q \in U$, there exists a unique solution $\gamma_q$ to the system of ODE:

\[
\frac{d\gamma_q}{dt}(t) = X(\gamma_q(t)),
\]

$\gamma_q(0) = q$.

Moreover, this solution depends differentiably on $q$.

For each $t \in (-\delta, \delta)$, define the map $\varphi_t : U \rightarrow M$ by $\varphi_t(q) = \varphi(t, q)$. It turns out that for each $t \in (-\delta, \delta)$, $\varphi_t(U)$ is open and $\varphi_t : U \rightarrow \varphi_t(U)$ is a diffeomorphism. $\{\varphi_t\}$ is called a **local flow of** $X$. We also say that $\{\varphi_t\}$ is the 1-parameter (local) group of diffeomorphisms generated by $X$.

**Exercise 1.35.** Show that $\varphi_{t+s} = \varphi_t \circ \varphi_s$.

If $\varphi : M \rightarrow N$ is a $C^\infty$ diffeomorphism and if $X$ is a $C^\infty$ vector field on $M$, then $d\varphi(X)$, defined by

\[
d\varphi(X)(q) = d\varphi_{\varphi^{-1}(q)}(X_{\varphi^{-1}(q)}),
\]

is a $C^\infty$ vector field on $N$. We call $d\varphi(X)$ the push forward of $X$ by $\varphi$. Some books denote this by $\varphi_* (X)$.

When $M$ is compact we obtain the following.

**Lemma 1.36.** Let $M$ be a compact differentiable manifold and let $X$ be a differentiable vector field on $M$. Then there exists a differentiable map

$\varphi : \mathbb{R} \times M \rightarrow M$

such that for each $q \in M$ the differentiable curve $\gamma_q : \mathbb{R} \rightarrow M$ defined by

$\gamma_q(t) = \varphi(t, q)$.

satisfies $\gamma_q(0) = q$ and

$\gamma_q'(t) = X(\gamma_q(t))$

for $t \in \mathbb{R}$.

Assuming $M$ is compact, for each $t \in \mathbb{R}$, we define the differentiable map $\varphi_t : M \rightarrow M$ by $\varphi_t(q) = \varphi(t, q)$ (the local flow is now a ‘global flow’). Since $\varphi_t \circ \varphi_{-t} = \varphi_{-t} \circ \varphi_t = \varphi_{t+(-t)} = \varphi_0 = \text{id}_M$, we have $(\varphi_t)^{-1} = \varphi_{-t}$. So each $\varphi_t$ is a diffeomorphism of $M$. So $\varphi_t$ is a 1-parameter group of diffeomorphisms.

6. The Lie derivative on vector fields

6.1. Lie derivative is the Lie bracket.

The Lie bracket is related the concept of Lie derivative (which we now consider for vector fields and later consider for more general tensors).
Let $X$ be a $C^\infty$ vector field on $\mathcal{M}$. Given $p \in \mathcal{M}$, let $\varphi_t : \mathcal{V} \to \mathcal{M}$ be a local flow of $X$, where $\mathcal{V}$ is a neighborhood of $p$ and $t \in (-\varepsilon, \varepsilon)$ for some $\varepsilon > 0$. For $q \in \mathcal{V}$, since $\varphi_t(q) = \varphi(t, q) = \gamma_q(t)$ and $\frac{d\gamma_q}{dt}(t) = X_{\gamma_q(t)}$, we have

$$\frac{d\varphi_t}{dt}(q) = X_{\varphi_t(q)}.$$

Let $X$ and $Y$ be $C^\infty$ vector fields on $\mathcal{M}$. The Lie derivative of $Y$ with respect to $X$ is defined by:

$$(L_X Y)_p \equiv \frac{d}{dt} \bigg|_{t=0} d\varphi_{-t}(Y_{\varphi_t(p)})$$

$$= \lim_{t \to 0} \frac{1}{t} \left( d\varphi_{-t} \left( Y_{\varphi_t(p)} \right) - d\varphi_0 \left( Y_{\varphi_0(p)} \right) \right)$$

$$= \lim_{t \to 0} \frac{1}{t} \left( d\varphi_{-t} \left( Y_{\varphi_t(p)} \right) - Y_p \right),$$

where $\varphi_t$ is a local flow of $X$. So the Lie derivative is a measure of how $Y$ changes under the action of the local flow of $X$.

On vector fields, the Lie derivative and Lie bracket are the same.

**Lemma 1.37.** $L_X Y = [X, Y]$; in other words, for each $p \in \mathcal{M}$,

$$[X, Y]_p = \lim_{t \to 0} \frac{1}{t} \left( d\varphi_{-t} \left( Y_{\varphi_t(p)} \right) - Y_p \right).$$

**Proof.** Step 1. $L_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = 0$. We first show that $L_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = 0$, which by Lemma 1.31 implies that $L_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = [\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}]$. Let $(\varphi_i)_t$ be the corresponding local flow of the vector field $\frac{\partial}{\partial x^i}$ on $x(\mathcal{U})$. For $p$ in its domain,

$$(1.15) \quad (\varphi_i)_t(p) = x \left( x^{-1}(p) + te_i \right),$$

which is the same as $c_i(t)$ in (1.5). Indeed, differentiating this in $t$, we have the defining equation for the $(\varphi_i)_t$:

$$\frac{d(\varphi_i)_t}{dt}(p) = dx_{x^{-1}(p)+te_i} \frac{\partial}{\partial x^i}(\varphi_i)_t(p).$$

Then, given $t$, the differential of $(\varphi_i)_t$ (as a function of $p$) is

$$\left( d(\varphi_i)_t \right)_p = dx_{x^{-1}(p)+te_i} \circ d \left( x^{-1} \right)_p.$$
So, regarding the expression in the definition of Lie derivative, we have

\[ \left( d(\varphi_i)_t \right)_t (\varphi_i)_t (p) \left( \frac{\partial}{\partial x^j} (\varphi_i)_t (p) \right) \]

\[ = \left( dx^{-1}((\varphi_i)_t (p)) - te_i \circ d (x^{-1}) (\varphi_i)_t (p) \right) \left( \frac{\partial}{\partial x^j} (\varphi_i)_t (p) \right) \]

\[ = dx^{-1}((\varphi_i)_t (p)) - te_i (e_j) \]

\[ = \frac{\partial}{\partial x^j} (p) \]

since \( x^{-1} ((\varphi_i)_t (p)) - te_i = p \) and since (1.9) implies \( d (x^{-1})_q \left( \frac{\partial}{\partial x^j} \right) = e_i \). Thus

\[ \left( L \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} \right)_p = \lim_{t \to 0} \frac{1}{t} \left( d (\varphi_i)_t (p) \left( \frac{\partial}{\partial x^j} (\varphi_i)_t (p) \right) - \frac{\partial}{\partial x^j} (p) \right) = 0. \]

**Remark.** So far we have only shown \( L \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} = 0 = [\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}] \). One of the difficulties of the proving the general case this way is that it is not obvious how to obtain a formula for the local flow \( \varphi_t \) for \( X \) in general. We can circumvent this difficulty.

**Step 2.** \( L \frac{\partial}{\partial x^i} Y \). Now consider \( L \frac{\partial}{\partial x^i} Y \) (we take \( i = 1 \) for convenience and we replace \( \frac{\partial}{\partial x^j} \) by any vector field \( Y \)). Recall that \( (\varphi_1)_t (p) = x (x^{-1} (p) + te_1) \). Let

\[ Y_q = \sum_{j=1}^n Y^i (q) \frac{\partial}{\partial x^j} q. \]

Note that from (1.13) we have

\[ \left[ \frac{\partial}{\partial x^1}, Y \right] = \sum_j \frac{\partial Y^j}{\partial x^1} \frac{\partial}{\partial x^j}. \]

Then

\[ (d (\varphi_1)_t)_t (\varphi_1)_t (p) \left( Y (\varphi_1)_t (p) \right) \]

\[ = \left( dx^{-1}((\varphi_1)_t (p)) - te_1 \circ d (x^{-1}) (\varphi_1)_t (p) \right) \left( Y (\varphi_1)_t (p) \right) \]

\[ = dx^{-1}((\varphi_1)_t (p)) - te_1 \left( \sum_{j=1}^n Y^j ((\varphi_1)_t (p)) e_j \right) \]

\[ = \sum_{j=1}^n Y^j ((\varphi_1)_t (p)) \frac{\partial}{\partial x^j} (p). \]
Thus

\[ (\mathcal{L}_{\frac{\partial}{\partial x^1}} Y)_p = \frac{d}{dt} \bigg|_{t=0} (d(\varphi_1)_t(p)) \left(Y(\varphi_1)_t(p)\right) \]

\[ = \frac{d}{dt} \bigg|_{t=0} \sum_{j=1}^n Y^j((\varphi_1)_t(p)) \frac{\partial}{\partial x^j} p \]

\[ = \sum_{j=1}^n \left( \frac{d}{dt} \bigg|_{t=0} (\varphi_1)_t(p) \right) (Y^j) \frac{\partial}{\partial x^j} p \]

\[ = \sum_{j=1}^n (\frac{\partial}{\partial x^1} p) (Y^j) \frac{\partial}{\partial x^j} p \]

\[ = \sum_{j=1}^n \frac{\partial Y^j}{\partial x^1} (p) \frac{\partial}{\partial x^j} p \]

since \( \frac{d(\varphi_1)_t}{dt} (p) = \frac{\partial}{\partial x^1}(\varphi_1)_t(p) \). We have shown

\[ (1.16) \quad \mathcal{L}_{\frac{\partial}{\partial x^1}} Y = \sum_{j=1}^n \frac{\partial Y^j}{\partial x^1} \frac{\partial}{\partial x^j} = \left[ \frac{\partial}{\partial x^1}, Y \right] \]

for any \( C^\infty \) vector field \( Y \).

**Step 3. The general case.** For a continuous vector field \( V \), let \( \mathcal{R}(V) = \{q \in \mathcal{M} : V(q) \neq 0\} \). Since \( V \) is continuous, \( \mathcal{R}(V) \) is an open set. Now \( \overline{\mathcal{R}(V)} = \text{supp}(V) \) is called the **support of** \( V \). Of course, \( V \equiv 0 \) on the open set \( \mathcal{M} - \text{supp}(V) \).

Now consider any vector fields \( X \) and \( Y \).

**Fact.** If \( p \in \mathcal{R}(X) \), then there exists a parametrization \( x' : U' \to \mathcal{M} \) with \( x'(0) = p \) such that

\[ (1.17) \quad X = \frac{\partial}{\partial(x')^1} \]

on \( x'(U') \).

With this fact and (1.16), we obtain

\[ \mathcal{L}_X Y = [X, Y] \]

on \( \mathcal{R}(X) \), and by continuity, the equality also on \( \overline{\mathcal{R}(X)} = \text{supp}(X) \).

On the other hand, on the open set \( \mathcal{M} - \text{supp}(X) \) we have \( X \equiv 0 \) and \( \varphi_t(p) \equiv p \) for any \( p \in \mathcal{M} - \text{supp}(X) \). From this we easily conclude that

\[ \mathcal{L}_X Y = 0 = [X, Y] \]

on \( \mathcal{M} - \text{supp}(X) \). We are done proving Lemma 1.37.
Proof of (1.17). Let \((\mathcal{U}, x)\) be a parametrization with \(p \in x(\mathcal{U})\), \(x^{-1}(p) = 0\), and \((dx)_0(e_1) = X_p\) (we may obtain this by precomposing an arbitrary parametrization containing \(p\) in its image with an appropriate affine transformation of \(\mathbb{R}^n\)). Let \(\{\varphi_t\}\) be the local flow of \(X\). Then there exists \(\delta > 0\) such that the map
\[
x' : \mathcal{U}' \ni (-\delta, \delta) \times \{ x = (x^2, \ldots, x^n) \in \mathbb{R}^{n-1} : |x| < \delta \} \to \mathcal{M}
\]
given by
\[
x'(t, x^2, \ldots, x^n) = \varphi_t(x(0, x^2, \ldots, x^n))
\]
is a well defined parametrization. Since \(dx'_0 = dx_0\), that \(x'\) is a parametrization for some \(\delta > 0\) sufficiently small follows from the inverse function theorem (Proposition 1.18). Note that \(x'(0, x^2, \ldots, x^n) = x(0, x^2, \ldots, x^n)\). Then for any \(q \in x'(\mathcal{U}')\),
\[
\frac{\partial}{\partial (x')^1} = \left. \frac{dx'}{(x')^{-1}(q)}(e_1) \right|_{x = x^2, \ldots, x^n} = \varphi_t(x(0, x^{-1}(q)^2, \ldots, x^{-1}(q)^n)) = X_q.
\]

6.2. Push forward of the Lie bracket.

Let \(\varphi : \mathcal{N} \to \mathcal{M}\) be a diffeomorphism and let \(X\) and \(Y\) be vector fields on \(\mathcal{N}\). Then, acting on any function \(f : \mathcal{M} \to \mathbb{R}\), we have
\[
d\varphi([X,Y])(f) \circ \varphi = (XY - YX)(f \circ \varphi)
\]
\[
= X((d\varphi(Y)(f)) \circ \varphi) - Y((d\varphi(X)(f)) \circ \varphi)
\]
\[
= d\varphi(X)(d\varphi(Y)(f)) \circ \varphi - d\varphi(Y)(d\varphi(X)(f)) \circ \varphi
\]
\[
= [d\varphi(X), d\varphi(Y)](f) \circ \varphi.
\]
Hence we have the ‘diffeomorphism invariance of the Lie bracket’:
\[
(1.18) \quad d\varphi([X,Y]) = [d\varphi(X), d\varphi(Y)].
\]

Using this, we now give another proof of the Jacobi identity (1.14). Given \(X, Y, Z \in \mathfrak{X}(\mathcal{M})\), let \(\{\varphi_t\}\) be the 1-parameter local group of diffeomorphisms
generated by $X$. Then, using Lemma 1.37, we compute that

$$\left[ X, [Y, Z] \right]_p = \left( \mathcal{L}_X [Y, Z] \right)_p$$

$$= d \left. \frac{d}{dt} \phi_t \right|_{t=0} [Y, Z]_{\phi_t(p)}$$

$$= \left. \frac{d}{dt} \left[ d\phi_t (Y), d\phi_t (Z) \right] \right|_{t=0}$$

$$= \left[ \left. \frac{d}{dt} d\phi_t (Y), Z \right|_p, \left. \frac{d}{dt} d\phi_t (Z) \right|_p \right]$$

$$= \left[ [X, Y], Z \right]_p + \left[ Y, \left[ X, Z \right] \right]_p.$$

Thus

$$\left[ X, [Y, Z] \right] + \left[ Y, [Z, X] \right] + \left[ Z, [X, Y] \right] = 0.$$

**Exercise 1.38.** Show that for 1-parameter family of vector fields $Y (t)$ and $Z (t)$,

$$\frac{d}{dt} [Y (t), Z (t)] = \left[ \frac{dY}{dt} (t), Z (t) \right] + \left[ Y (t), \frac{dZ}{dt} (t) \right].$$

This justifies the fourth equality in (1.19).

**Solution.** Given $f \in C^\infty (\mathcal{M})$, let $g (t) = Y (t) (f)$ and $h (t) = Z (t) (f)$. Then

$$\frac{d}{dt} [Y (t), Z (t)] (f) = \frac{d}{dt} (Y (t) (h (t)) - Z (t) (g (t)))$$

$$= Y' (t) (h (t)) + Y (t) (h' (t) - Z' (t) (g (t)) - Z (t) (g' (t)))$$

$$= Y' (t) (Z (t) (f)) + Y (t) (Z' (t) (f)) - Z' (t) (Y (t) (f)) - Z (t) (Y' (t) (f))$$

$$= \left[ Y' (t), Z (t) \right] (f) + \left[ Y (t), Z' (t) \right] (f).$$

7. Cotangent spaces, cotangent bundles and 1-forms

7.1. Cotangent space and cotangent bundle.

Let $\mathcal{M}$ be a $C^\infty$ manifold. Given $p \in \mathcal{M}$, the cotangent space at $p$ is

$$T^*_p \mathcal{M} \doteq (T_p \mathcal{M})^* = \{ \omega : T_p \mathcal{M} \to \mathbb{R} \text{ is linear} \}.$$

That is $T^*_p \mathcal{M}$ is the dual vector space of $T_p \mathcal{M}$. Analogous to the tangent bundle, we define the cotangent bundle to be the disjoint union

$$T^* \mathcal{M} = \bigcup_{p \in \mathcal{M}} T^*_p \mathcal{M},$$

which is equivalent to

$$T^* \mathcal{M} = \{ (p, v) : p \in \mathcal{M}, \ v \in T^*_p \mathcal{M} \}.$$
We have the projection map \( \pi : T^*M \to M \) defined by \( \pi(p, \omega) = p \). Note that \( \pi^{-1}(p) = T^*_pM \).

7. 1-forms and dual bases.

A \( C^\infty \) 1-form (or, less commonly, cotangent vector field) \( \alpha \) on \( M \) is a \( C^\infty \) section of \( T^*M \), that is, a \( C^\infty \) map \( \alpha : M \to T^*M \) satisfying \( \pi \circ \alpha = \text{id}_M \). The notion of 1-form is dual to the notion of vector field on \( T^*M \) is replaced by \( T^*M \). We may also consider a 1-form \( \alpha \) as a function \( \alpha : T^*_pM \to \mathbb{R} \) such that \( \alpha \vert_{T^*_pM} : T^*_pM \to \mathbb{R} \) is linear for each \( p \in M \). We denote the set of 1-forms by \( C^\infty(T^*M) \).

Given a parametrization \((U, x)\), define the 1-forms \( dx^j \), for \( j = 1, \ldots, n \), on \( x(U) \) by

\[
(1.20) \quad dx^j \left( \frac{\partial}{\partial x^i} p \right) = \delta^j_i
\]

at each point \( p \in x(U) \). Namely, \( \{dx^j\}_{j=1}^n \) is the basis of \( T^*_pM \) dual to the basis \( \{\frac{\partial}{\partial x^i}\}_{i=1}^n \) of \( T_pM \).

Exercise 1.39. Show, for each \( j = 1, \ldots, n \), that \( dx^j \in T^*_pM \) is the same as the differential \( d(x^{-1})^j : T_pM \to \mathbb{R} \), where the functions \( (x^{-1})^j : x(U) \to \mathbb{R} \) are defined by \( x^{-1} = ((x^{-1})^1, \ldots, (x^{-1})^n) \).

Exercise 1.40. Show that if \( \alpha \) is a \( C^\infty \) 1-form on \( M \) and \((U, x)\) is a parametrization of \( M \), then on \( x(U) \) we have

\[
\alpha = \sum_{i=1}^n \alpha_i dx^i,
\]

where \( \alpha_i = \alpha \left( \frac{\partial}{\partial x^i} \right) : x(U) \to \mathbb{R} \). The functions \( \alpha_i \) are called the components of \( \alpha \) with respect to the parametrization \((U, x)\).

Exercise 1.41. Show that the cotangent bundle \( T^*M \) is a 2n-dimensional \( C^\infty \) manifold.

7.3. Lie derivative of a 1-form.

Let \( M^n \) be a \( C^\infty \) manifold. Let \( X \) be a \( C^\infty \) vector field on \( M \) and let \( f : M \to \mathbb{R} \) be a \( C^\infty \) function. The Lie derivative of \( f \) with respect to \( X \) is defined by:

\[
(L_X f)(p) \doteq \frac{d}{dt} \bigg|_{t=0} (f \circ \varphi_t)(p),
\]
where \( \{ \varphi_t \} \) is the local flow of \( X \). We compute that

\[
(L_X f)(p) = \left. \frac{d}{dt} \right|_{t=0} (f \circ \gamma_p)(t) = \gamma'_p(0)(f) = X_p(f),
\]

where \( \gamma_p \) is the integral curve to \( X \) with \( \gamma_p(0) = p \). So, acting on functions, the Lie derivative is simply the directional derivative.

Let \( \varphi : \mathcal{N} \to \mathcal{M} \) be a \( C^\infty \) map. Given a \( C^\infty \) 1-form \( \alpha \) on \( \mathcal{M} \), its pull back by \( \varphi \) to \( \mathcal{N} \) is defined by

\[
(\varphi^* \alpha)(u) = \alpha(d\varphi(u))
\]

for \( u \in TN \). Then \( \varphi^* \alpha \) is a \( C^\infty \) 1-form \( \alpha \) on \( \mathcal{N} \).

Let \( X \) be a \( C^\infty \) vector field on \( \mathcal{M} \) and let \( \alpha \) be a \( C^\infty \) 1-form on \( \mathcal{M} \). The Lie derivative of \( \alpha \) with respect to \( X \) is defined by:

\[
(L_X \alpha)_p \doteq \left. \frac{d}{dt} \right|_{t=0} \varphi^*_t \left( \alpha_{\varphi_t(p)} \right) = \lim_{t \to 0} \frac{1}{t} \left( \varphi^*_t \left( \alpha_{\varphi_t(p)} \right) - \alpha_p \right).
\]

**Exercise 1.42.** Show that if \((U, x)\) is a parametrization of \( \mathcal{M} \), then for each \( i \) and \( j \),

\[
\mathcal{L}_{\frac{\partial}{\partial x^i}} dx^j = 0
\]

in \( x(U) \).

**Exercise 1.43** (Lie derivative product rule). Show that if \( X \) and \( Y \) are \( C^\infty \) vector fields on \( \mathcal{M} \) and \( \alpha \) is a \( C^\infty \) 1-form on \( \mathcal{M} \), then

\[
X(\alpha(Y)) = (L_X \alpha)(Y) + \alpha([X,Y]).
\]

For example, as components with respect to \((U, x)\),

\[
(L_{\frac{\partial}{\partial x^i}} \alpha)_j = \frac{\partial \alpha_j}{\partial x^i}.
\]

\(^{3}\text{Note that } X(\alpha(Y)) = L_X(\alpha(Y)) \text{ since } \alpha(Y) \text{ is a function.}\)
Chapter 2

Riemannian Manifolds

The concept of Riemannian manifold is a generalization of Euclidean space with the Euclidean geometry (inner product).

1. Introduction to Riemannian metrics

1.1. Euclidean geometry.

Euclidean geometry is determined by the standard Euclidean inner product \( \langle \cdot, \cdot \rangle_{\mathbb{R}^n} \) on \( \mathbb{R}^n \). The length of a vector \( X \in \mathbb{R}^n \) is defined by

\[
|X| = \sqrt{\langle X, X \rangle_{\mathbb{R}^n}}
\]

and the angle between two nonzero vectors \( X \) and \( Y \) is defined by

\[
\angle(X,Y) = \cos^{-1}\left(\frac{\langle X, Y \rangle_{\mathbb{R}^n}}{|X||Y|}\right).
\]

More generally, an inner product \( \langle \cdot, \cdot \rangle \) on a real vector space \( V \) is defined to satisfy:

\[
\langle v, u \rangle = \langle u, v \rangle,
\]

\[
\langle u_1 + cu_2, v \rangle = \langle u_1, v \rangle + c \langle u_2, v \rangle,
\]

\[
\langle u, u \rangle \geq 0,
\]

\[
\langle u, u \rangle = 0 \text{ if and only if } u = 0
\]

for any \( c \in \mathbb{R} \) and where the \( u \)'s and \( v \)'s are in \( V \). Of course, \( \langle u, v_1 + cv_2 \rangle = \langle u, v_1 \rangle + c \langle u, v_2 \rangle \) follows by symmetry. That is,

\[
\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}
\]

is a symmetric, bilinear, positive-definite form.
1.2. Definition of Riemannian metric and isometry.

Since each tangent space of a differentiable manifold is a (finite-dimensional) real vector space, we may extend the infinitesimal (at each point) notion of geometry to manifolds.

**Definition 2.1.** A *Riemannian metric* on a differentiable manifold $\mathcal{M}^n$ is an assignment to each $p \in \mathcal{M}$ an inner product $g_p \equiv \langle \cdot, \cdot \rangle_p$ on $T_p \mathcal{M}$. When $\mathcal{M}$ is a $C^\infty$ differentiable manifold, we say that $g$ is $C^\infty$ if it depends on $p$ in a $C^\infty$ way.

More precisely, let $(U, x)$ be a parametrization of $\mathcal{M}$ and define the corresponding coefficients $g_{ij} : x(U) \to \mathbb{R}$ by

$$g_{ij} \equiv g\left( \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right)$$

for $i, j = 1, \ldots, n$. We say that the Riemannian metric $g$ is $C^\infty$ if for any parametrization $(U, x)$ the corresponding coefficients $g_{ij}$ of the metric are $C^\infty$ functions.

Equivalently, $g$ is $C^\infty$ iff for any $C^\infty$ vector fields $X, Y$ on $\mathcal{M}$, the function $p \mapsto g_p(X(p), Y(p))$ is $C^\infty$.

**Definition 2.2.** A $C^\infty$ *Riemannian manifold* is a pair $(\mathcal{M}^n, g)$, where $\mathcal{M}^n$ is a $C^\infty$ differentiable manifold and $g$ is a $C^\infty$ Riemannian metric.

The induced Riemannian metric on the graph of a Euclidean function is given by the following.

**Exercise 2.3.** Let $U \subset \mathbb{R}^n$ be an open set and let $f : U \to \mathbb{R}$ be a $C^\infty$ function. Define $x : U \to \mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$ by

$$x(q) = (q, f(q)).$$

Compute $dx_q(e_i)$ for $q \in U$, and compute

$$g_{ij}(q) \equiv \langle dx_q(e_i), dx_q(e_j) \rangle,$$

where $\langle \cdot, \cdot \rangle$ is the standard Euclidean inner product on $\mathbb{R}^{n+1}$.

**Solution.** For $x(q^1, \ldots, q^n) = (q^1, \ldots, q^n, f(q^1, \ldots, q^n))$ we have

$$dx_q(e_i) = \frac{\partial x}{\partial q_i}(q) = \left( 0, \ldots, 1, \ldots, 0, \frac{\partial f}{\partial q_i}(q) \right) = e_i + \frac{\partial f}{\partial q_i}(q) e_{n+1},$$
where the 1 is in the \( i \)th slot. Thus
\[
g_{ij} (q) = \left\langle e_i + \frac{\partial f}{\partial q^i} (q) e_{n+1}, e_j + \frac{\partial f}{\partial q^j} (q) e_{n+1} \right\rangle_{\mathbb{R}^{n+1}}
= \delta_{ij} + \frac{\partial f}{\partial q^i} (q) \frac{\partial f}{\partial q^j} (q).
\]

The notion of equivalence of Riemannian manifolds is given by the following.

**Definition 2.4.** Let \((M^n, g)\) and \((N^n, h)\) be \(C^\infty\) Riemannian manifolds. A diffeomorphism \(\phi : M \to N\) is called an **isometry** if it preserves the inner products in the sense that

\[
g_p (u, v) = h_{\phi(p)} (d\phi_p (u), d\phi_p (v))
\]

for any \(p \in M\) and \(u, v \in T_p M\).

**Exercise 2.5.** Show that if a map is an isometry, then its inverse is also an isometry.

Suppose that \(\phi : (M^n, g) \to (N^n, h)\) is an isometry. Let \(p \in M\), \((U, x)\) be a parametrization of \(M\) with \(p \in x (U)\), and \((V, y)\) be a parametrization of \(N\) with \(\phi (p) \in y (V)\). Define \(h_{\alpha \beta} \equiv h (\frac{\partial}{\partial y^\alpha}, \frac{\partial}{\partial y^\beta})\) and \(\phi^\alpha \equiv (y^{-1})^\alpha \circ \phi\).

**Exercise 2.6.** Show that
\[
g_{ij} (p) = \sum_{\alpha, \beta = 1}^n h_{\alpha \beta} (\phi (p)) \frac{\partial \phi^\alpha}{\partial x^i} (p) \frac{\partial \phi^\beta}{\partial x^j} (p).
\]

The set of isometries from a Riemannian manifold to itself forms a group, called its **isometry group**, with multiplication given by composition.

### 1.3. Examples of Riemannian metrics and isometries.

Some of the simplest Riemannian manifolds are the following.

**Examples.**

(0) **Euclidean space.** \(\mathbb{R}^n\) with the Euclidean Riemannian metric \(\langle , \rangle_{\mathbb{R}^n}\) defined by
\[
\langle u, v \rangle_{\mathbb{R}^n} = u^1 v^1 + \cdots + u^n v^n.
\]

For the global parametrization \(\text{id}_{\mathbb{R}^n} : \mathbb{R}^n \to \mathbb{R}^n\), we have \(g_{ij} = \delta_{ij}\), where
\[
\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}
\]

The isometry group of \((\mathbb{R}^n, \langle , \rangle_{\mathbb{R}^n})\) is generated by:

(a) maps of the form \(u \mapsto u + c\), where \(c \in \mathbb{R}^n\), i.e., **translations**,
(b) linear maps $A$ satisfying $\langle Av, Aw \rangle = \langle v, w \rangle$ for all $v, w \in \mathbb{R}^n$, i.e., orthogonal transformations. The group of orthogonal transformations is denoted by $O(n, \mathbb{R})$.

(1) **Graph.** Let $\mathcal{U} \subset \mathbb{R}^n$ be an open set and let $f : \mathcal{U} \to \mathbb{R}$ be a $C^\infty$ function. Define the $C^\infty$ manifold

$$\mathcal{M}^n = \{(q, f(q)) : q \in \mathcal{U}\} \subset \mathbb{R}^{n+1}.$$ 

Recall that we have the global parametrization $x : \mathcal{U} \to \mathcal{M}$ defined by

$$x(q) = (q, f(q)).$$ 

The coordinate vectors may identified with

$$\frac{\partial}{\partial x^i} x(q) = e_i + \frac{\partial f}{\partial q^i}(q) \in \mathbb{R}^{n+1}.$$ 

The Euclidean inner product $\langle , \rangle_{\mathbb{R}^{n+1}}$ induces a $C^\infty$ Riemannian metric $g$ on $\mathcal{M}$ defined by

$$g(u, v) = \langle u, v \rangle_{\mathbb{R}^{n+1}}$$ 

for $u, v \in T_p \mathcal{M}$, where $p$ is any point in $\mathcal{M}$. Exercise 2.3 asks you to compute $g_{ij}$.

(2) **The unit sphere** $\mathcal{S}^n = \{x \in \mathbb{R}^{n+1} : |x| = 1\} \subset \mathbb{R}^{n+1}$. Recall that for $p \in \mathcal{S}^n$ we may identify its tangent space with

$$T_p \mathcal{S}^n = p^\perp \cong \{x \in \mathbb{R}^{n+1} : \langle x, p \rangle = 0\}.$$ 

As a hypersurface in $\mathbb{R}^{n+1}$, $\mathcal{S}^n$ has the induced Riemannian metric defined, like above, by

$$g_{\mathcal{S}^n}(u, v) = \langle u, v \rangle_{\mathbb{R}^{n+1}}$$ 

for $u, v \in T_p \mathcal{S}^n$, where $p \in \mathcal{S}^n$.

The isometry group of $\mathcal{S}^n$ is $O(n+1, \mathbb{R})$.

Recall that given $\nu \in \mathcal{S}^n$, we have the parametrization

$$x_\nu : \nu^\perp \cong \mathbb{R}^n \to \mathcal{S}^n - \{\nu\}$$ 

defined, as the inverse of stereographic projection, by

$$x_\nu(q) = \frac{2q + (|q|^2 - 1) \nu}{|q|^2 + 1}.$$ 

Let $\nu = N = (0, \ldots, 0, 1)$ be the north pole and identify $N^\perp = \mathbb{R}^n \times \{0\}$ with $\mathbb{R}^n$ by $(q, 0) \mapsto q$ for $q \in \mathbb{R}^n$.

**Exercise 2.7.** Compute the Riemannian metric $g$ on $\mathbb{R}^n$ defined by

$$g_q(u, v) = g_{\mathcal{S}^n}((dx_N)_q(u), (dx_N)_q(v)).$$
for \( u, v \in T_q \mathbb{R}^n \cong \mathbb{R}^n \). How does \( g(u, v) \) compare with the Euclidean metric \( \langle u, v \rangle_{\mathbb{R}^n} \)? We remark that, by definition, \((\mathbb{R}^n, g)\) is isometric to \((S^n - \{N\}, g_{S^n})\).

**2) Hyperbolic space.** There are a number of models of hyperbolic space. All of the following examples are isometric to each other.

(a) **Disk model.** Let

\[
\mathcal{D} = \{ x \in \mathbb{R}^n : |x| < 1 \}.
\]

The hyperbolic disk metric \( g_\mathcal{D} \) on \( \mathcal{D} \) is defined by: for \( x \in \mathcal{D} \),

\[
(g_\mathcal{D})_x (u, v) = \frac{4 \langle u, v \rangle_{\mathbb{R}^n}}{1 - |x|^2}.
\]

(b) **Half-space model.** Let

\[
\mathcal{H} = \{ x \in \mathbb{R}^n : x_n > 0 \}.
\]

The hyperbolic half-space metric \( g_\mathcal{H} \) on \( \mathcal{H} \) is defined by: for \( x \in \mathcal{H} \),

\[
(g_\mathcal{H})_x (u, v) = \frac{\langle u, v \rangle_{\mathbb{R}^n}}{x_n^2}.
\]

(c) **Hyperboloid model.** Let

\[
\mathcal{L} = \{ x \in \mathbb{R}^{n+1} : x_1^2 + \cdots + x_n^2 - x_{n+1}^2 = -1, \ x_{n+1} > 0 \}.
\]

The hyperbolic half-space metric \( g_\mathcal{H} \) on \( \mathcal{H} \) is defined by: for \( x \in \mathcal{L} \),

\[
(g_\mathcal{L})_x (u, v) = \langle u, v \rangle_{\mathbb{M}^{n+1}} = u_1^1 v_1^1 + \cdots + u^n v^n - u_{n+1} v_{n+1}.
\]

**Exercise 2.8.** Show that for each \( x \in \mathcal{L} \), \( (g_\mathcal{L})_x \) is positive definite on \( T_x \mathcal{L} \). So \((\mathcal{L}, g_\mathcal{L})\) is a Riemannian manifold.

**Exercise 2.9.** Show that the map \( F : (\mathcal{H}, g_\mathcal{H}) \to (\mathcal{D}, g_\mathcal{D}) \) given by

\[
F(x^1, \ldots, x^n) = \left( \frac{4 - |x|^2}{|x|^2 + 4 x^n + 4}, \frac{4 x^1}{|x|^2 + 4 x^n + 4}, \ldots, \frac{4 x^{n-1}}{|x|^2 + 4 x^n + 4} \right)
\]

is well defined and an isometry.
Remark. Note that
\[
|F(x^1, \ldots, x^n)| = \sqrt{\left(\frac{|x|^2 + 4}{|x|^2 + 4x^n + 4}\right)^2 - 16 (x^n)^2} = \frac{\sqrt{|x|^2 - 4x^n + 4}}{\sqrt{|x|^2 + 4x^n + 4}} = \frac{\sqrt{\sum_{i=1}^{n-1} (x_i)^2 + (x^n - 2)^2}}{\sqrt{\sum_{i=1}^{n-1} (x_i)^2 + (x^n + 2)^2}}
\]

Exercise 2.10. Following up on the above exercise, find explicit isometries between the different models of hyperbolic space. Hint: see the expository paper [2].

2. Properties of Riemannian metrics

Given a differentiable manifold, we would like to know that there actually exists a Riemannian metric on it. Fortunately this is generally true.

2.1. Existence of Riemannian metrics.

The existence of a Riemannian metric on a differentiable manifold is given by Proposition 2.10 on p. 43 of Do Carmo (see also Proposition 11.26 on p. 284 of Lee [8]). The idea of the proof is to paste together Euclidean metrics, defined via parametrizations, using a partition of unity subordinate to the covering by parametrizations.

Theorem 2.11 (Existence of Riemannian metrics). If \( M^n \) is a connected, Hausdorff \( C^\infty \) manifold admitting a countable number of parametrizations whose coordinate neighborhoods cover \( M \), then there exists a \( C^\infty \) Riemannian metric \( g \) on \( M \).

Fact. Under the above hypotheses, there exists a countable number of parametrizations \( \{(U_\alpha, x_\alpha)\}_{\alpha \in A} \) such that \( \{x_\alpha(U_\alpha)\}_{\alpha \in A} \) is a locally finite cover of \( M \) (i.e., for each \( p \in M \) there exists a neighborhood \( V \) of \( p \) such that only a finite number of the \( x_\alpha(U_\alpha) \) intersect \( V \)) and there exists a \( C^\infty \) partition of unity \( \{\psi_\alpha\}_{\alpha \in A} \) subordinate to the cover \( \{x_\alpha(U_\alpha)\}_{\alpha \in A} \) of \( M \), i.e.,

1. \( \psi_\alpha : M \to [0, 1] \) is a \( C^\infty \) function with support in \( x_\alpha(U_\alpha) \) for each \( \alpha \in A \),
2. \( \sum_{\alpha \in A} \psi_\alpha (p) = 1 \) (this is a finite sum) at each point \( p \in M \).
Proof of Theorem 2.11. Let \( \{ (U_\alpha, x_\alpha) \}_{\alpha \in A} \) and \( \{ \psi_\alpha \}_{\alpha \in A} \) be as above. For each \( \alpha \in A \), define the Riemannian metric \( g^\alpha \) on \( x_\alpha (U_\alpha) \) by
\[
g^\alpha (u, v) = \left( d\left( x_\alpha^{-1} \right)_p (u), d\left( x_\alpha^{-1} \right)_p (v) \right)_{\mathbb{R}^n}
\]
for \( u, v \in T_p M \), where \( p \in x_\alpha (U_\alpha) \). We extend \( g^\alpha \) to a symmetric bilinear form on each \( T_p M \) for \( p \in M \) by setting it to be zero outside of \( x_\alpha (U_\alpha) \). Then \( \psi_\alpha g^\alpha \) is a \( C^\infty \) symmetric bilinear form on each \( T_p M \) for \( p \in M \). Note that \( \psi_\alpha g^\alpha \) is positive definite wherever \( \psi_\alpha \neq 0 \) (recall that \( \text{supp} (\psi_\alpha) \subset x_\alpha (U_\alpha) \)). Define
\[
g = \sum_{\alpha \in A} \psi_\alpha g^\alpha,
\]
which at each point of \( M \) is a finite sum. \( g \) is a \( C^\infty \) symmetric bilinear form on each \( T_p M \) for \( p \in M \). Moreover, for any \( p \in M \) there exists \( \beta \in A \) such that \( \psi_\beta (p) > 0 \) and \( p \in x_\beta (U_\beta) \). Then for any \( u \in T_p M \setminus \{0\} \) we have
\[
g(u, u) = \sum_{\alpha \in A} \psi_\alpha g^\alpha (u, u) \geq \psi_\beta g^\beta (u, u) > 0.
\]
Of course, \( g_p (0, 0) = 0 \) for any \( p \in M \). Thus \( g \) is a \( C^\infty \) Riemannian metric on \( M \). \( \square \)

2.2. Lengths of paths and distance.

Let \( (M^n, g) \) be a Riemannian manifold. Given a piecewise \( C^1 \) path \( \gamma : [a, b] \to M \), its length is defined by
\[
L(\gamma) \doteq \int_a^b |\gamma'(u)| \, du.
\]
This length is invariant under reparametrization of the path \( \gamma \). That is, if \( \phi : [c, d] \to [a, b] \) is a \( C^1 \) diffeomorphism, then
\[
L(\gamma \circ \phi) = L(\gamma).
\]
Indeed, by the chain rule and the change of variables formula,
\[
L(\gamma \circ \phi) = \int_a^b \left| \frac{d}{du} (\gamma \circ \phi) (u) \right| \, du
= \int_a^b |\gamma' (\phi (u))| \, |\phi' (u)| \, du
= L(\gamma).
\]

\(^1\)Note that \( g^\alpha \left( \frac{\partial}{\partial x^i_\alpha}, \frac{\partial}{\partial x^j_\alpha} \right) = \delta_{ij} \).
We say that a \( C^1 \) path \( \gamma(u) \) is parametrized by arc length if \( \| \gamma'(u) \| \equiv 1 \). Assuming \( \| \gamma'(u) \| \neq 0 \) for all \( u \in [a, b] \), we may reparametrize \( \gamma \) by arc length by defining
\[
s = \phi^{-1} (\bar{u}) = \int_a^{\bar{u}} \| \gamma'(u) \| \, du,
\]
so that \( \frac{ds}{d\bar{u}} (\bar{u}) = \| \gamma'(\bar{u}) \| \). Then
\[
\left| \frac{d}{ds} (\gamma \circ \phi)(s) \right| = \| \gamma'(\phi(s)) \| \left| \frac{d\bar{u}}{ds} (s) \right| = \| \gamma'(\bar{u}) \| \cdot \frac{1}{\frac{ds}{d\bar{u}} (\bar{u})} = 1.
\]

The distance function \( d : \mathcal{M} \times \mathcal{M} \to [0, \infty) \) is defined by
\[
d(x, y) = \inf_{\gamma} L(\gamma),
\]
where the infimum is taken over all piecewise \( C^\infty \) paths \( \gamma : [0, 1] \to \mathcal{M}^n \) with \( \gamma(0) = x \) and \( \gamma(1) = y \).

3. Hypersurfaces of Euclidean space

Euclidean hypersurfaces provide basic and concrete examples of Riemannian manifolds.

Let \( \mathcal{M}^n \subset \mathbb{R}^{n+1} \) be a regular (Euclidean) hypersurface, i.e., for each \( p \in \mathcal{M} \) there exists a neighborhood \( \mathcal{V} \) of \( p \) in \( \mathcal{M} \) and a \( C^\infty \) homeomorphism \( x : \mathcal{U} \subset \mathbb{R}^n \to \mathcal{M} \cap \mathcal{V} \subset \mathbb{R}^{n+1} \) such that \( dx_u : T_u \mathcal{M} = \mathbb{R}^n \to T_{x(u)} \mathbb{R}^{n+1} = \mathbb{R}^{n+1} \) is injective for each \( u \in \mathcal{U} \).

**Exercise 2.12.** Show that there is a unique \( C^\infty \) manifold structure on \( \mathcal{M} \) that contains each \( (\mathcal{U}, x) \) such that \( x : \mathcal{U} \to \mathcal{M} \cap \mathcal{V} \) is a \( C^\infty \) homeomorphism and \( dx_u \) is injective for \( u \in \mathcal{U} \).

The induced Riemannian metric \( I \) on \( \mathcal{M} \), also called the first fundamental form, is defined by
\[
I(U, V) = \langle U, V \rangle_{\mathbb{R}^{n+1}}
\]
for \( U, V \in T_p \mathcal{M} \), where \( p \in \mathcal{M} \). This is simply the restriction of the Euclidean inner product to the tangent spaces \( T_p \mathcal{M} \). It is easy to see that \( I \) is a \( C^\infty \) Riemannian metric.

We may naturally identify \( T_p \mathcal{M} \) with an \( n \)-dimensional subspace of \( \mathbb{R}^{n+1} \). In particular, at each \( p \in \mathcal{M} \) there are exactly two unit vectors perpendicular to \( T_p \mathcal{M} \), called unit normals. If \( \mathcal{M} \) is orientable, then there exactly two
choices of \( C^\infty \) unit normal vector fields defined at all points of \( \mathcal{M} \). Let \( \nu : \mathcal{M} \to \mathbb{R}^{n+1} \) be such a choice. We have

1. \( |\nu| \equiv 1 \) on \( \mathcal{M} \),
2. \( \langle \nu_p, U \rangle = 0 \) for each \( U \in T_p \mathcal{M}, p \in \mathcal{M} \),
3. \( \nu^i : \mathcal{M} \to \mathbb{R} \) is \( C^\infty \) for each \( i = 1, \ldots, n + 1 \).

Note that if \( \alpha, \beta : (-\varepsilon, \varepsilon) \to \mathcal{M} \) are differentiable curves with \( \alpha'(0) = \beta'(0) \) in the Euclidean sense considering \( \alpha, \beta : (-\varepsilon, \varepsilon) \to \mathbb{R}^{n+1} \), then for any \( C^\infty \) function \( f : \mathcal{M} \to \mathbb{R} \) we have \( (f \circ \alpha)'(0) = (f \circ \beta)'(0) \), i.e., \( \alpha'(0) = \beta'(0) \) as derivations on \( \mathcal{M} \). Of course, the same holds true for \( \mathbb{R}^{n+1} \)-valued functions on \( \mathcal{M} \) such as the normal \( \nu : \mathcal{M} \to \mathbb{R}^{n+1} \).

Given \( p \in \mathcal{M} \), the second fundamental form at \( p \)

\[ \Pi : T_p \mathcal{M} \times T_p \mathcal{M} \to \mathbb{R} \]

is defined by

\[ \Pi(U, V) \overset{\text{def}}{=} \langle U(\nu), V \rangle \]

for \( U, V \in T_p \mathcal{M} \).

First we note the following elementary fact. If \( \alpha : (-\varepsilon, \varepsilon) \to \mathbb{R}^n \) is a differentiable curve and if \( V \) and \( W \) are vector fields defined on the image curve \( \alpha(-\varepsilon, \varepsilon) \) such that \( V(\alpha(t)) \) and \( W(\alpha(t)) \) are differentiable, then

(2.4) \[ \frac{d}{dt} \langle V(\alpha(t)), W(\alpha(t)) \rangle = \left\langle \frac{d}{dt} V(\alpha(t)), W(\alpha(t)) \right\rangle + \left\langle V(\alpha(t)), \frac{d}{dt} W(\alpha(t)) \right\rangle. \]

Now extend \( V \) to a vector field on \( \mathcal{M} \) defined in a neighborhood \( U \subset \mathcal{M} \) of \( p \). Since \( \langle \nu, V \rangle \equiv 0 \) in \( U \), at \( p \) we have by (2.4)

\[ 0 = U \langle \nu, V \rangle = \langle U(\nu), V \rangle + \langle \nu, U(V) \rangle, \]

i.e., the second fundamental form may be rewritten as

(2.5) \[ \Pi(U, V) = -\langle \nu, U(V) \rangle. \]

Next extend \( U \) on \( \mathcal{M} \) to a neighborhood of \( p \). We compute at \( p \) that

\[ \Pi(U, V) - \Pi(V, U) = -\langle \nu, U(V) \rangle + \langle \nu, U(V) \rangle \]

\[ = -\langle \nu, [U, V] \rangle \]

\[ = 0, \]

where we used Exercise 1.29 and \( [U, V]_p \in T_p \mathcal{M} \) is the Lie bracket of \( U \) and \( V \) as vector fields on \( \mathcal{M} \). Hence the second fundamental form is symmetric:

\[ \Pi(U, V) = \Pi(V, U). \]
We may also describe the second fundamental form in terms of the Weingarten map
\[ L : T_p \mathcal{M} \to T_p \mathcal{M} \]
defined by
\[ L(U) = U(\nu). \]
The reason \( L(T_p \mathcal{M}) \subset T_p \mathcal{M} \) is because
\[ \langle L(U), \nu \rangle = \langle U(\nu), \nu \rangle = \frac{1}{2} U(\langle \nu, \nu \rangle) = 0. \]
By definition,
\[ \II(U, V) = \langle L(U), V \rangle. \]
Since \( \II \) is symmetric, the linear map \( L \) is self-adjoint, i.e., \( \langle L(U), V \rangle = \langle U, L(V) \rangle \).

The eigenvalues of the Weingarten map, which we denote by
\[ \kappa_1, \ldots, \kappa_n, \]
are called the principal curvatures. Note that since \( L \) is self-adjoint, there exists an orthonormal basis of eigenvectors \( \{e_i\}_{i=1}^n \) for \( T_p \mathcal{M} \) such that
\[ L(e_i) = \kappa_i e_i, \]
where the \( \kappa_i \) are real. Note that the principal curvatures are associated to a choice of unit normal vector field \( \nu \). If we choose the opposite unit normal, then all the signs of the \( \kappa_i \) change.

The mean curvature (depends on the choice of \( \nu \)) at a point \( p \in \mathcal{M} \) is the sum of the principal curvatures (i.e., the trace of the Weingarten map)
\[ H(p) \triangleq \kappa_1 + \cdots + \kappa_n = \text{tr}(L) = \sum_{i=1}^n \II(\bar{e}_i, \bar{e}_i), \]
where \( \{\bar{e}_i\} \) is any orthonormal basis for \( T_p \mathcal{M} \).

The Gauss curvature is the product of the principal curvatures (i.e., the determinant of the Weingarten map)
\[ K \triangleq \kappa_1 \cdots \kappa_n = \det L = \frac{\det \II}{\det \II}. \]

We now calculate using a parametrization \( x : \mathcal{U} \to \mathcal{M} \cap \mathcal{V} \) of our hypersurface, which is a \( C^\infty \) homeomorphism such that each \( dx_u \) is injective. Now the coordinate tangent vectors are
\[ \frac{\partial}{\partial x^i} = e'_i(0) = e_i(x) \]
for \( i = 1, \ldots, n \). The components of the induced Riemannian metric (first fundamental form) on \( \mathcal{M} \) are given in \( \mathcal{M} \cap \mathcal{V} \) by
\[ I_{ij} = \langle e_i(x), e_j(x) \rangle_{\mathbb{R}^{n+1}}. \]
We may pull back the Riemannian metric $I$ on $\mathcal{M}$ by $x$ to a Riemannian metric $g$ on $U$, defined by
\begin{equation}
(2.8) \quad g(U, V) \triangleq I(dx(U), dx(V)) = \langle dx(U), dx(V) \rangle_{\mathbb{R}^{n+1}}
\end{equation}
for $U, V \in T_x U = \mathbb{R}^n$. Then
\[ g_{ij}(p) \triangleq g_p(e_i, e_j) = I_{ij}(x(p)) \]
for $p \in U$.

The coefficients of the second fundamental form are
\[ \Pi_{ij} = \Pi(e_i(x), e_j(x)) = \langle e_i(x)(\nu), e_j(x) \rangle_{\mathbb{R}^{n+1}} , \]
where $e_i(x)(\nu)$ is the tangent vector $e_i(x)$ to $\mathcal{M}$ acting on $\nu$ as an $\mathbb{R}^{n+1}$-valued function on $\mathcal{M}$. Note that $e_i(x)(\nu) = e_i(\nu \circ x)$. By (2.5) and $e_i(x)(e_j(x)) = e_i(e_j(x))$, we may rewrite this as
\[ \Pi_{ij} = -\langle \nu, e_i(e_j(x)) \rangle_{\mathbb{R}^{n+1}} . \]

More explicitly, i.e., keeping track at which point each quantity is, we have
\[
\Pi_{ij}(x(p)) = \left\langle \left( (e_i)_p(x) \right)_{x(p)}(\nu), (e_j)_p(x) \right\rangle_{\mathbb{R}^{n+1}} \\
= -\left\langle \nu(x(p)), (e_i)_p \left( (e_j)_p(x) \right) \right\rangle_{\mathbb{R}^{n+1}} .
\]

We may pull back the second fundamental form $\Pi$ to $U$ by defining
\[ h(U, V) \triangleq \Pi(dx(U), dx(V)) \]
for $U, V \in T_x U = \mathbb{R}^n$. Then
\[ h_{ij}(p) \triangleq h_p(e_i, e_j) = \Pi_{ij}(x(p)) \]
for $p \in U$.

**Exercise 2.13.** Compute the second fundamental form, Weingarten map, mean curvature, and Gauss curvature for the graph in Exercise 2.3.

### 4. Lie groups

A **Lie group** $G$ is a set $G$ with:

1. a binary operation, called multiplication, $\cdot : G \times G \to G$ such that $(G, \cdot)$ is a group,
2. a $C^\infty$ differentiable manifold structure on $G$

such that the map
\[ \phi : G \times G \to G \]
defined by
\[ \phi(x, y) = x \cdot y^{-1} \]
is $C^\infty$, where $G \times G$ has the product manifold differentiable structure.
Define the right and left multiplication maps
\[ R_x : G \to G \quad \text{and} \quad L_x : G \to G \]
by
\[ R_x (y) = xy \quad \text{and} \quad L_x (y) = yx. \]
It is easy to see that for any \( x \in G \), \( R_x \) is a diffeomorphism with inverse \( R_x^{-1} \) and that \( L_x \) is a diffeomorphism with inverse \( L_x^{-1} \).

A \( C^\infty \) vector field \( X \) on \( G \) is **left-invariant** if
\[ dL_x (X_y) = X_{xy} \]
for each \( x, y \in G \). **Right-invariant** means
\[ dR_x (X_y) = X_{yx} \]
for each \( x, y \in G \).

Let \( e \) denote the identity element of \( G \). Given a tangent vector \( X_e \in T_e G \), we define a corresponding left-invariant vector field (LIVF) on \( G \) by
\[ X_x = dL_x (X_e) \]
for each \( x \in G \) (exercise: check that \( X \) is left-invariant). In this way we may identify the set of left-invariant vector fields on \( G \) with \( T_e G \).

A subspace \( \mathfrak{h} \) of a Lie algebra \( \mathfrak{g} \) is called a **Lie subalgebra** if it is closed under the Lie bracket operation.

**Exercise.** Show that if \( X \) and \( Y \) are LIVF on a Lie group, then \([X, Y] \) is a LIVF.

Because of this, LIVF is a Lie subalgebra of the set of all vector fields on \( G \). In other words, \( T_e G \) is a Lie algebra with Lie bracket defined by
\[ [X_e, Y_e] \triangleq [X, Y]_e. \]

Let \( \langle \,, \, \rangle_e \) be an inner product on \( T_e G \). Define a corresponding Riemannian metric \( \langle \,, \, \rangle \) on \( G \) by
\[ (2.9) \quad \langle u, v \rangle_x = \langle (dL_{x^{-1}})_x (u), (dL_{x^{-1}})_x (v) \rangle_e \]
for any \( x \in G \) and \( u, v \in T_x G \). Note that if \( u, v \in T_e G \), then
\[ \langle (dL_x)_e (u), (dL_x)_e (v) \rangle_x = \langle u, v \rangle_e. \]
If a Riemannian metric \( g \) on \( G \) satisfies (2.9), then we say that \( g \) is a **left-invariant metric**.

One similarly defines **right-invariant metric** on a Lie group.

**Examples:**
Let \( \mathbb{F} \) be \( \mathbb{R} \) or \( \mathbb{C} \).

1. The general linear group \( \text{GL}(n, \mathbb{F}) \) is the set of invertible \( n \times n \) matrices with entries in \( \mathbb{F} \), i.e., \( \{ A : \det A \neq 0 \} \), with matrix multiplication.
2. The orthogonal group $O(n, \mathbb{F})$ is the set
\[ \{ A \in \text{GL}(n, \mathbb{F}) : A^T \cdot A = A \cdot A^T = I \} , \]
where $T$ denotes transpose and $I$ denotes identity. $O(n, \mathbb{F})$ is a subgroup of $\text{GL}(n, \mathbb{F})$. We have
\[ O(n, \mathbb{F}) = \{ A \in \text{GL}(n, \mathbb{F}) : \langle Av, Aw \rangle = \langle v, w \rangle \text{ for all } v, w \in \mathbb{F}^n \} . \]

3. The special orthogonal group $SO(n, \mathbb{F})$ is the set
\[ O(n, \mathbb{F}) \cap \{ A : \det A = 1 \} . \]

4. The unitary group $U(n)$ is the subgroup of $\text{GL}(n, \mathbb{C})$ given by
\[ \{ A : \bar{A}^T \cdot A = A \cdot \bar{A}^T = I \} . \]
Define the Hermitian inner product on $\mathbb{C}^n$ by $(v, w) = \sum_i v_i \bar{w}_i$. Then
\[ U(n) = \{ A \in \text{GL}(n, \mathbb{C}) : (Av, Aw) = (v, w) \text{ for all } v, w \in \mathbb{C}^n \} . \]
For further discussion of Lie groups, see Lee [8], pp. 37–40, pp. 93–100, pp. 194–199, and Chapter 9.

5. Levi-Civita connection

Tangent vectors act on $C^\infty$ functions and we call this action directional differentiation by analogy with Euclidean space. The Lie derivative is a derivative acting on vector fields and it is equal to the Lie bracket. However, it is not analogous to the directional derivative. On a Riemannian manifold we do naturally have such a notion, called covariant differentiation.


Let $(\mathcal{M}^n, g)$ be a Riemannian manifold. Let $\Xi(\mathcal{M})$ denote the set of $C^\infty$ vector fields. The Levi-Civita connection (or Riemannian covariant derivative) is the unique map
\[ \nabla : \Xi(\mathcal{M}) \times \Xi(\mathcal{M}) \to \Xi(\mathcal{M}) , \]
denoted by $\nabla (X, Y) = \nabla_X Y$, satisfying: for each $X, Y, Z \in \Xi(\mathcal{M})$ and $f \in C^\infty(\mathcal{M})$,
1. $\nabla_{fX+Y}Z = f\nabla_X Z + \nabla_Y Z,$
2. $\nabla_X (Y + Z) = \nabla_X Y + \nabla_X Z,$
3. $\nabla_X (fY) = f\nabla_X Y + X(f)Y,$
4. $\nabla_X Y - \nabla_Y X = [X, Y],$
5. $X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle.$
A map satisfying only (1)–(3) is called an **affine connection**. Property (4) is called **symmetric** or **torsion-free**; property (5) is called **compatible with the metric**.

We need to show:

**Theorem 2.14.** *The Levi-Civita connection exists, is unique, and satisfies*

\[
\langle \nabla_X Y, Z \rangle = \frac{1}{2} \{ X \langle Y, Z \rangle + Y \langle Z, X \rangle - Z \langle X, Y \rangle + \langle [X, Y] , Z \rangle + \langle [Z, X], Y \rangle - \langle [Y, Z], X \rangle \}.
\]

**Proof.** *(Uniqueness)* If a map \( \nabla \) satisfies properties (1)–(5), then by (5)

\[
X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle,
\]

\[
Y \langle Z, X \rangle = \langle \nabla_Y Z, X \rangle + \langle Z, \nabla_Y X \rangle,
\]

\[
Z \langle X, Y \rangle = \langle \nabla_Z X, Y \rangle + \langle X, \nabla_Z Y \rangle.
\]

Adding the first two equations and subtracting the third equation yields

\[
X \langle Y, Z \rangle + Y \langle Z, X \rangle - Z \langle X, Y \rangle = \langle \nabla_X Y, Z \rangle + \langle Z, \nabla_Y X \rangle + \langle Y, \nabla_X Z \rangle - \langle \nabla_Z X, Y \rangle - \langle X, \nabla_Z Y \rangle,
\]

which by (4) equals

\[
2 \langle \nabla_X Y, Z \rangle - \langle [X, Y] , Z \rangle - \langle [Z, X], Y \rangle + \langle [Y, Z], X \rangle.
\]

Hence we obtain

\[
2 \langle \nabla_X Y, Z \rangle = T(X, Y, Z),
\]

where the map \( T : \Xi(\mathcal{M}) \times \Xi(\mathcal{M}) \times \Xi(\mathcal{M}) \to C^\infty(\mathcal{M}) \) is defined by

\[
T(X, Y, Z) \triangleq X \langle Y, Z \rangle + Y \langle Z, X \rangle - Z \langle X, Y \rangle + \langle [X, Y] , Z \rangle + \langle [Z, X], Y \rangle - \langle [Y, Z], X \rangle.
\]

This is equivalent to formula (2.10) for \( \nabla \). Since this formula uniquely determines \( \nabla \), if the Levi-Civita connection exists, then it is unique.

*(Existence)* It is easy to see that \( T(X, Y, Z) \) is \( \mathbb{R} \)-linear in each component. We compute that (see Exercise 2.16 below)

\[
T(X, Y, fZ) = fT(X, Y, Z) = T(fX, Y, Z),
\]

i.e., \( T(X, Y, Z) \) is \( C^\infty(\mathcal{M}) \)-linear in the first and third components. Hence, by Lemma 2.18 below,

\[
T(X, Y, Z)(p) = M(X_p, Y, Z_p) \in \mathbb{R}
\]

depends only on \( X_p, Z_p \in T_p\mathcal{M} \) and \( Y \in \Xi(\mathcal{M}) \). Given \( X_p \in T_p\mathcal{M} \) and \( Y \in \Xi(\mathcal{M}) \), define

\[
L_{X,Y} : T_p\mathcal{M} \to \mathbb{R}
\]
by
\[(2.15)\]
\[L_{X,Y}(Z) \equiv M(X,Y,Z).\]

Then \(L_{X,Y}\) is an \(\mathbb{R}\)-linear function. By (2.69), this implies that there exists a unique tangent vector \(L^*_{X,Y} \in T_pM\) such that
\[(2.16)\]
\[L_{X,Y}(Z) = \left\langle L^*_{X,Y}, Z \right\rangle_{p}\]
for each \(Z \in T_pM\). Define \(\nabla : \Xi(M) \times \Xi(M) \rightarrow \Xi(M)\) by
\[(2.17)\]
\[\left(\nabla_{X}Y\right)(p) \equiv \frac{1}{2}L^*_{X,Y},\]
which is easily seen to be well-defined. Note that by (2.17), (2.16), (2.15) and (2.14), we have
\[
\left\langle \left(\nabla_{X}Y\right)(p), Z_{p} \right\rangle = \frac{1}{2}L_{X,Y}(Z)_{p} = \frac{1}{2}M(X,Y,Z)_{p} = \frac{1}{2}T(X,Y,Z)(p),
\]
which is (2.11), a formula we derived assuming the existence of an affine connection satisfying properties (1)–(5). We now show, in fact, that \(\nabla_{X}Y\) satisfies the five properties of the Levi-Civita connection.

Since \(T(X,Y,Z)\) is \(C^\infty(M)\)-linear in the first component, we have that \(L_{X,Y}\) is a linear in \(X\), which implies that \(\left(\nabla_{X}Y\right)(p)\) depends linearly on \(X\). I.e., property (1) holds. Similarly, we see that \(\left(\nabla_{X}Y\right)(p)\) is \(\mathbb{R}\)-linear in \(Y\), so that property (2) holds.

For each \(X,Y,Z \in \Xi(M)\), we have
\[
2 \left\langle \left(\nabla_{X}Y\right)(p) - \left(\nabla_{Y}X\right)(p), Z_{p} \right\rangle = L_{X,Y}(Z)_{p} - L_{Y,X}(Z)_{p} = T(X,Y,Z)(p) - T(Y,X,Z)(p) = \left\langle [X,Y], Z \right\rangle_{p} + \left\langle [Z,X], Y \right\rangle_{p} - \left\langle [Y,Z], X \right\rangle_{p} - \left\langle [Y,X], Z \right\rangle_{p} - \left\langle [Z,Y], X \right\rangle_{p} + \left\langle [X,Z], Y \right\rangle_{p} = 2 \left\langle [X,Y], Z \right\rangle\]
by (2.12). Since \(Z\) is arbitrary, this implies that for each \(X,Y \in \Xi(M)\),
\[
\left(\nabla_{X}Y\right)(p) - \left(\nabla_{Y}X\right)(p) = [X,Y](p),
\]
i.e., property (4) holds.
Finally,

\[ 2 (\langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle) = T(X,Y,Z) + T(X,Z,Y) = X \langle Y, Z \rangle + Y \langle Z, X \rangle - Z \langle X, Y \rangle + X \langle Z, Y \rangle + Z \langle Y, X \rangle - Y \langle X, Z \rangle + \langle [X,Y], Z \rangle + \langle [Z,X], Y \rangle - \langle [Y,Z], X \rangle + \langle [X,Z], Y \rangle + \langle [Y,X], Z \rangle - \langle [Z,Y], X \rangle = 2X \langle Y, Z \rangle, \]

which verifies property (5). Hence \( \nabla \) defined by (2.17) satisfies properties (1)--(5).

\[ \square \]

Remark 2.15. By (2.17), the Levi-Civita connection is equivalent to maps \( \nabla : T_pM \times \Xi (M) \to T_pM \) for all \( p \in M \).

Exercise 2.16. Verify (2.13).

Exercise 2.17. Let \( \varphi : (\mathcal{N}^n, h) \to (\mathcal{M}^n, g) \) be an isometry. Let \( \nabla^g \) and \( \nabla^h \) denote the Levi-Civita connections of \( g \) and \( h \), respectively. Prove that

\[ (2.18) \quad \nabla^h_X Y = (d\varphi)^{-1} \left( \nabla^g_{d\varphi(X)} d\varphi (Y) \right). \]

5.2. A prelude to tensors.

Now we prove a result we used above and which is related to the concept of tensor, to be discussed later in more detail.

Lemma 2.18. Suppose that \( T : \Xi (M) \to \Xi (M) \) satisfies

1. \( T(X + Y) = T(X) + T(Y), \)
2. \( T(fX) = fT(X), \)

Then for any \( X \in \Xi (M) \) we have \( T(X)_p \) depends only on \( X_p \). In other words, for any \( p \in M \) and any two \( C^\infty \) vector fields \( X_1, X_2 \in \Xi (M) \) with \( X_1(p) = X_2(p) \), we have \( T(X_1)_p = T(X_2)_p \).

When the map \( T \) satisfies (1) and (2), we call it a \textbf{tensor}.

We first need the following.

Lemma 2.19. If \( X_1, X_2 \in \Xi (M) \) are such that \( X_1 = X_2 \) in a neighborhood \( V \) of a point \( p \in M \), then \( T(X_1)_p = T(X_2)_p \).

Proof. Let \( f : M \to \mathbb{R} \) be a smooth function with \( f(p) \neq 0 \) and \( \operatorname{supp} (f) \subset V \). Then \( fX_1 = fX_2 \in \Xi (M) \). Hence \( T(fX_1)_p = T(fX_2)_p \in \Xi (M) \) and, in particular, \( f(p) T(X_1)_p = f(p) T(X_2)_p \in T_pM \). The lemma follows since \( f(p) \neq 0 \).

\[ \square \]
Proof of Lemma 2.18. Suppose \( p \in \mathcal{M} \) and \( X \in \mathfrak{X}(\mathcal{M}) \) are such that \( X_p = 0 \). The lemma will follow from showing that

\[
(2.19) \quad T(X)_p = 0.
\]

Let \((\mathcal{U}, x)\) be a parametrization of \( \mathcal{M} \) with \( p \in x(\mathcal{U}) \). Then the vector fields \( \left\{ \frac{\partial}{\partial x^i} \right\}_{i=1}^n \) are a basis for \( T_q \mathcal{M} \) for each \( q \in x(\mathcal{U}) \). Then in \( x(\mathcal{U}) \) we have

\[
X = \sum_{i=1}^n X^i \frac{\partial}{\partial x^i},
\]

where \( X^i \in C^\infty(x(\mathcal{U})) \) and \( X^i(p) = 0 \). Let \( \varphi : \mathcal{M} \to \mathbb{R} \) be a smooth function with \( \varphi(p) = 1 \) in a neighborhood of \( p \) and \( \text{supp}(\varphi) \subset x(\mathcal{U}) \). Then by Lemma 2.19,

\[
T(X)_p = T \left( \varphi \sum_{i=1}^n X^i \frac{\partial}{\partial x^i} \right)_p = \varphi(p) \sum_{i=1}^n X^i(p) T \left( \frac{\partial}{\partial x^i} \right)_p = 0.
\]

The lemma indeed follows since if \( X_1, X_2 \in \mathfrak{X}(\mathcal{M}) \) with \( X_1(p) = X_2(p) \), then \( X_1 - X_2 \in \mathfrak{X}(\mathcal{M}) \) satisfies \( (X_1 - X_2)_p = 0 \), which by (2.19) implies that \( T(X_1)_p = T(X_2)_p \). \( \square \)

Exercise 2.20. Let \( D, \bar{D} : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M}) \) be affine connections. Show that their difference \( D - \bar{D} : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M}) \) is a tensor.

The main point is to show that for \( X, Y, Z \in \mathfrak{X}(\mathcal{M}) \) and \( f \in C^\infty(\mathcal{M}) \),

\[
D_X(fY) - \bar{D}_X(fY) = f \left( D_X Y - \bar{D}_X Y \right).
\]

Since we also have

\[
D_{fX} Y - \bar{D}_{fX} Y = f \left( D_X Y - \bar{D}_X Y \right),
\]

we say that \( D - \bar{D} \) is a tensor.

Exercise 2.21. Let \((\mathcal{U}, x)\) be a parametrization of \( \mathcal{M} \) and let \( (g^{kl}) \) denote the inverse matrix of \( (g_{ij}) \), i.e.,

\[
(2.21) \quad \sum_{k=1}^n g^{ik} g_{jk} = \delta^i_j.
\]

Show that if \( V \) is a tangent vector, then

\[
(2.22) \quad V = \sum_{k,l=1}^n g^{kl} \left( V \frac{\partial}{\partial x^l} \right) \frac{\partial}{\partial x^k}.
\]

Hint. We just need to show that the inner products of the LHS and RHS of (2.22) with any \( \frac{\partial}{\partial x^l} \) are equal.
5.3. The Christoffel symbols.

For differential geometry calculations, it is useful to consider the Levi-Civita connection with respect to the coordinate vectors \( \frac{\partial}{\partial x^i} \). In particular, we define the **Christoffel symbols** \( \Gamma^k_{ij} \) by

\[
\nabla \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} = \sum_{k=1}^{n} \Gamma^k_{ij} \frac{\partial}{\partial x^k}.
\]

We think of the Christoffel symbols as being the components of the Levi-Civita connection with respect to a parametrization. From (2.10) and \([\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}] = 0\), we obtain

\[
\left< \nabla \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right> = \frac{1}{2} \left( \frac{\partial}{\partial x^i} g_{j\ell} + \frac{\partial}{\partial x^j} g_{i\ell} - \frac{\partial}{\partial x^\ell} g_{ij} \right).
\]

Hence, by (2.22) we have the fundamental formula for the Christoffel symbols:

**Lemma 2.22.**

\[
\Gamma^k_{ij} = \frac{1}{2} \sum_{\ell=1}^{n} g^{k\ell} \left( \frac{\partial}{\partial x^i} g_{j\ell} + \frac{\partial}{\partial x^j} g_{i\ell} - \frac{\partial}{\partial x^\ell} g_{ij} \right),
\]

where \((g^{k\ell})\) is the inverse of the metric component matrix as in (2.21).

Let \( U \subset \mathbb{R}^n \) be an open set and let \( f : U \to \mathbb{R}^+ \) be a \( C^\infty \) positive function. Define the Riemannian metric

\[
g = f^2 g_{\mathbb{R}^n}.
\]

With respect to the identity parametrization \( x = \text{id}_U \), we have

\[
g_{ij} = f^2 \delta_{ij}.
\]

Thus (2.24) yields

\[
\Gamma^k_{ij} = \frac{1}{f} \left( \frac{\partial f}{\partial x^i} \delta_{jk} + \frac{\partial f}{\partial x^j} \delta_{ik} - \frac{\partial f}{\partial x^k} \delta_{ij} \right).
\]

5.4. Covariant differentiation of 1-forms.

The covariant derivative of a 1-form \( \alpha \) on \( \mathcal{M} \) is defined by

\[
(\nabla_X \alpha)(Y) \doteq X(\alpha(Y)) - \alpha(\nabla_X Y)
\]

for all \( X, Y \in T_p \mathcal{M}, p \in \mathcal{M} \). The motivation for this definition is that is essentially the product rule, i.e.,

\[
X(\alpha(Y)) = (\nabla_X \alpha)(Y) + \alpha(\nabla_X Y).
\]

**Exercise 2.23.** Show that if \( \alpha \) is a 1-form, then \( \nabla_X (\alpha^*) = (\nabla_X \alpha)^* \).
Solution. We compute
\[ \langle \nabla_X (\alpha^*), Y \rangle = X \langle \alpha^*, Y \rangle - \langle \alpha^*, \nabla_X Y \rangle \]
\[ = X (\alpha (Y)) - \alpha (\nabla_X Y) \]
\[ = (\nabla_X \alpha) (Y) \]
\[ = \langle (\nabla_X \alpha)^*, Y \rangle. \]

Exercise 2.24. Show that if \( \alpha \) and \( \beta \) are 1-forms, then for any vector \( X \),
\[ (2.27) \quad X \langle \alpha, \beta \rangle = \langle \nabla_X \alpha, \beta \rangle + \langle \alpha, \nabla_X \beta \rangle. \]

Solution. Let \( \alpha^* \coloneqq g^{-1} (\alpha) \in \Xi (\mathcal{M}) \). We have
\[ X \langle \alpha, \beta \rangle = X \langle \alpha^*, \beta^* \rangle \]
\[ = \langle \nabla_X (\alpha^*), \beta^* \rangle + \langle \alpha^*, \nabla_X (\beta^*) \rangle \]
\[ = \langle (\nabla_X \alpha)^*, \beta^* \rangle + \langle \alpha^*, (\nabla_X \beta)^* \rangle \]
\[ = \langle \nabla_X \alpha, \beta \rangle + \langle \alpha, \nabla_X \beta \rangle, \]
where we used \( \nabla_X (\alpha^*) = (\nabla_X \alpha)^* \).

6. Covariant differentiation along a path and parallel translation

In this section we discuss covariant differentiation along a curve. Later we shall discuss the more general notion of covariant differentiation along a map, which, in the case of a map from a 2-dimensional rectangle, is useful for computing variation of arc length formulas for families of curves.

Let \( \mathcal{M}^n \) be a \( C^\infty \) manifold and let \( D : \Xi (\mathcal{M}) \times \Xi (\mathcal{M}) \to \Xi (\mathcal{M}) \) be an affine connection. Let \( c : I \to \mathcal{M} \) be a \( C^\infty \) curve, where \( I \subset \mathbb{R} \) is an interval. By a \( C^\infty \) vector field along \( c \) we mean a \( C^\infty \) curve \( V : I \to T \mathcal{M} \) with \( \pi \circ V = c \), i.e., \( V (t) \in T_{c(t)} \mathcal{M} \). Let \( \Xi (c) \) denote the set of \( C^\infty \) vector fields along \( c \).

Lemma 2.25. There exists a unique map \( \frac{D}{dt} : \Xi (c) \to \Xi (c) \), called covariant differentiation along \( c \), with the properties that: for any \( V, W \in \Xi (c) \) and \( C^\infty \) function \( f : I \to \mathbb{R} \),

(i) \[ \frac{D}{dt} (V + W) = \frac{D}{dt} V + \frac{D}{dt} W, \]
(ii) \[ \frac{D}{dt} (fV) = f \frac{D}{dt} V + \frac{df}{dt} V, \]
(iii) For any \( \bar{V} \in \Xi (\mathcal{M}) \) we have \( \frac{D}{dt} (\bar{V} \circ c) = D_c \bar{V} \).

Proof. (Uniqueness) Suppose that \( \frac{D}{dt} : \Xi (c) \to \Xi (c) \) satisfies (i)–(iii). Let \( t \in I \) and \( (U, x) \) be a parametrization with \( c (t) \in x (U) \). Given \( V \in \Xi (c) \),
we may write

\[(2.28) \quad V(t) = \sum_{i=1}^{n} V^i(t) \frac{\partial}{\partial x^i(c(t))},\]

where \(V^i: \mathcal{I} \to \mathbb{R}\) is a \(C^\infty\) function. By (i) and (ii), we have

\[(2.29) \quad \frac{dV}{dt} = \sum_{i=1}^{n} V^i(t) \frac{d}{dt} \left( \frac{\partial}{\partial x^i \circ c} \right) + \sum_{i=1}^{n} \frac{dV^i}{dt}(t) \frac{\partial}{\partial x^i \circ c}.
\]

Hence to prove uniqueness, we only need to show that \(\frac{d}{dt} \left( \frac{\partial}{\partial x^i \circ c} \right)\) is uniquely determined. By (ii), for any \(W \in \Xi(M)\) and \(p \in M\), \(\frac{dW}{dt}(p)\) depends only on \(W\) in any given neighborhood of \(p\). By (iii), we have

\[(2.30) \quad \frac{d}{dt} \left( \frac{\partial}{\partial x^i \circ c} \right)(t) = \frac{D}{dt} \frac{\partial}{\partial x^i}.
\]

By (2.29), (2.28) and (2.30), \(\frac{dV}{dt}\) is uniquely determined.

\((Existence)\) Given \(t \in \mathcal{I}\) and a parametrization \((\mathcal{U}, x)\), define

\[(2.31) \quad \left( \frac{d}{dt} V \right)(t) = \sum_{i=1}^{n} V^i(t) D_{c'(t)} \frac{\partial}{\partial x^i} + \sum_{i=1}^{n} \frac{dV^i}{dt}(t) \frac{\partial}{\partial x^i \circ c},\]

where \(V^i(t)\) is defined by (2.28). Clearly (i) holds. We have

\[
\frac{d}{dt} (fV) = \sum_{i=1}^{n} fV^i(t) D_{c'(t)} \frac{\partial}{\partial x^i} + \sum_{i=1}^{n} \frac{dfV^i}{dt}(t) \frac{\partial}{\partial x^i \circ c}
= f \frac{dV}{dt} + \sum_{i=1}^{n} \frac{df}{dt}(t) V^i(t) \frac{\partial}{\partial x^i \circ c}
= f \frac{dV}{dt} + \frac{df}{dt} V,
\]

which is (ii). Suppose \(\bar{V} \in \Xi(M)\) and define \(V = \bar{V} \circ c\). We compute

\[(2.32) \quad D_{c'(t)} \bar{V} = D_{c'(t)} \left( \sum_{i=1}^{n} \bar{V}^i \frac{\partial}{\partial x^i} \right)
= \sum_{i=1}^{n} \left( c'(t) (\bar{V}^i) \frac{\partial}{\partial x^i \circ c(t)} + \bar{V}^i(c(t)) D_{c'(t)} \frac{\partial}{\partial x^i} \right)
= \sum_{i=1}^{n} \left( \frac{dV^i}{dt}(t) \frac{\partial}{\partial x^i \circ c(t)} + V^i(t) D_{c'(t)} \frac{\partial}{\partial x^i} \right)
= \left( \frac{d}{dt} V \right)(t),
\]

so that (iii) holds. To prove existence, we just need to show that the definition (2.31) of \(\left( \frac{d}{dt} V \right)(t)\) is independent of the parametrization. With \((\mathcal{U}, x)\)
6. Covariant differentiation along a path

and corresponding \( \left( \frac{D}{dt}V \right)_{(t,x)} \) (as above (we added the subscript to emphasize dependence), given another parametrization \((V,y)\) with \(c(t) \in y(V)\), the corresponding \( \left( \frac{D}{dt}V \right)_{(V,y)}(t) \) satisfies (i)–(iii). Hence, by uniqueness, it must be equal to \( \left( \frac{D}{dt}V \right)_{(t,x)}(t) \). □

**Remark 2.26.** In the proof we expressed \( V \) locally in terms of coordinate vector fields. In place of \( \{ \frac{\partial}{\partial x_i} \}_{i=1}^n \) we may have used any \( C^\infty \) local vector fields \( \{ E_i \}_{i=1}^n \) which form a basis at each point.

Given an affine connection \( D \), the covariant derivative along a path defines the notion of parallelism. We say that a vector field \( V \) along a path \( c : I \to \mathcal{M}^n \) is parallel if

\[
\left( \frac{D}{dt}V \right)(t) = 0
\]

for \( t \in I \).

**Lemma 2.27.** Let \( \mathcal{M}^n \) be a differentiable manifold, let \( D \) be an affine connection, and let \( c : I \to \mathcal{M} \) be a \( C^\infty \) path. Then given \( t_0 \in I \) and \( V_0 \in T_{c(t_0)}\mathcal{M} \), there exists a unique parallel vector field \( V \) along \( c \) with \( V(t_0) = V_0 \).

**Proof.** Suppose that \( (a,b) \subset I \) and a parametrization \((U,x)\) are such that \( c(t) \in x(U) \) for \( t \in (a,b) \). Let \( V \) be a parallel vector field along \( c|_{(a,b)} \) and write

\[
V(t) = \sum_{i=1}^n V^i(t) \frac{\partial}{\partial x_i c(t)}
\]

for \( t \in (a,b) \). Then by (2.32) we have

\[
\sum_{i=1}^n \left( \frac{dV^i}{dt}(t) \frac{\partial}{\partial x^i c(t)} + V^i(t) D_{c(t)} \frac{\partial}{\partial x^i} \right) = \left( \frac{D}{dt}V \right)(t) = 0.
\]

Define \( x^j(t) \doteq ((x^{-1})^j \circ c)'(t) \), so that

\[
(2.33) \quad c'(t) = \sum_{j=1}^n x^j(t) \frac{\partial}{\partial x^j c(t)}.
\]

Define also \( \Delta^k_{ji} \) by

\[
D \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} = \sum_{k=1}^n \Delta^k_{ji} \frac{\partial}{\partial x^k}
\]

(if \( D = \nabla \) is the Levi-Civita connection, then \( \Delta^k_{ji} = \Gamma^k_{ji} \) are the Christoffel symbols). We have

\[
\sum_{i=1}^n \left( \frac{dV^i}{dt}(t) \frac{\partial}{\partial x^i c(t)} + V^i(t) \sum_{j,k=1}^n x^j(t) \Delta^k_{ji} (c(t)) \frac{\partial}{\partial x^k c(t)} \right) = 0,
\]
so that
\[ (2.34) \quad \frac{dV^k}{dt} (t) + \sum_{i,j=1}^n V^i(t) x^j(t) \Delta_{ji}^k (c(t)) = 0 \]
for each \( k = 1, \ldots, n \). Note that each \( x^j(t) \Delta_{ji}^k (c(t)) \) is a \( C^\infty \) function of \( t \). Define \( V_0 \equiv \sum_{k=1}^n V^k \frac{\partial}{\partial x^k} c(t_0) \). We leave it to the reader to check that we have actually shown that \( V \) is a parallel vector field along \( c \) with \( V(t_0) = V_0 \) if and only if the \( \{ V^k(t) \}^n_{k=1} \) satisfy (2.34) with \( V^k(0) = V^k_0 \).

Since (2.34) is a (linear) system of \( n \) ODE with \( C^\infty \) coefficients, there exists a unique such solution. Hence there does exist a unique parallel vector field along \( c \) with \( V(t_0) = V_0 \).

Now we prove that is true on all of \( c \). Let \( t_1 \in I \). For each \( t \in [t_0, t_1] \), let \( (U_t, x_t) \) be a parametrization with \( c(t) \in x_t(U_t) \) and let \( \epsilon_t > 0 \) be such that \( c([t - \epsilon_t, t + \epsilon_t] \cap I) \subset x_t(U_t) \). Since \([t_0, t_1]\) is compact, there exists a finite number of parametrizations \( \{(U_{t_\alpha}, x_{t_\alpha})\}_{\alpha=1}^m \) such that \([t_0, t_1] \subset \bigcup_{\alpha=1}^m (t_\alpha - \epsilon_{t_\alpha}, t_\alpha + \epsilon_{t_\alpha}) \). For each \( \alpha \) and any initial vector \( V_{0,\alpha} \in T_{c(t_\alpha)}M \) for some \( t_{0,\alpha} \in (t_\alpha - \epsilon_{t_\alpha}, t_\alpha + \epsilon_{t_\alpha}) \), there exists a unique parallel vector field \( V_{\alpha}(t) \) along \( c|_{[t_{0,\alpha}-\epsilon_{t_{0,\alpha}}, t_{0,\alpha}+\epsilon_{t_{0,\alpha}}]} \) with \( V_{\alpha}(t_{0,\alpha}) = V_{0,\alpha} \). From this fact we may easily derive the existence of a unique parallel vector field \( V \) along \( c|_{[t_0, t_1]} \) with \( V(t_0) = V_0 \). Since \( t_1 \in I \) is arbitrary and by uniqueness, the lemma follows.

Given an affine connection which is compatible with the metric, we now derive the corresponding property for the associated covariant differentiation along a curve.

**Lemma 2.28.** Let \((M^n, g)\) be a Riemannian manifold and let \( D \) be an affine connection which is compatible with the metric, i.e.,
\[ (2.35) \quad X \langle Y, Z \rangle = \langle D_X Y, Z \rangle + \langle Y, D_X Z \rangle \]
for \( X, Y, Z \in \mathfrak{X}(M) \). Let \( c : I \to M \) be a \( C^\infty \) curve and let \( V \) and \( W \) be \( C^\infty \) vector fields along \( c \). Then
\[ (2.36) \quad \frac{d}{dt} \langle V, W \rangle = \left\langle \frac{D}{dt} V, W \right\rangle + \left\langle V, \frac{D}{dt} W \right\rangle. \]

**Proof.** In essence, the idea of the proof is to reduce the formula to the case where \( V \) and \( W \) are coordinate vector fields. Let \( t \in I \) and let \((U, x)\) be a
parametrization with \( c(t) \in x(U) \). By (2.32), we have
\[
\left\langle \frac{D}{dt}V, W \right\rangle + \left\langle V, \frac{D}{dt}W \right\rangle = \sum_{i=1}^{n} \left( \frac{dV^i}{dt} \left\langle \frac{\partial}{\partial x^i}, W \right\rangle + V^i \left\langle D_{c'} \frac{\partial}{\partial x^i}, W \right\rangle \right) \\
+ \sum_{j=1}^{n} \left( \frac{dW^j}{dt} \left\langle \frac{\partial}{\partial x^j}, V \right\rangle + W^j \left\langle D_{c'} \frac{\partial}{\partial x^j}, V \right\rangle \right) \\
= \sum_{i,j=1}^{n} \left( \frac{dV^i}{dt} W^j g_{ij} + V^i W^j \left\langle D_{c'} \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle \right) \\
+ \sum_{i,j=1}^{n} \left( \frac{dW^j}{dt} V^i g_{ji} + W^j V^i \left\langle D_{c'} \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^i} \right\rangle \right) \\
= \sum_{i,j=1}^{n} \left( \frac{d}{dt} (V^i W^j) g_{ij} + V^i W^j c' (g_{ij}) \right) \\
= \frac{d}{dt} \langle V, W \rangle
\]
with the penultimate inequality using
\[
\left\langle D_{c'} \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle + \left\langle D_{c'} \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^i} \right\rangle = c' \left\langle \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^i} \right\rangle ,
\]
which in turn follows from (2.35).

As a consequence, along a curve the inner products of parallel vector fields are constant.

**Corollary 2.29.** Suppose that \( D \) is an affine connection which is compatible with the metric. If \( V \) and \( W \) are parallel vector fields along \( c \), then \(|V(t)|^2\) and \( \langle V(t), W(t) \rangle \) are constant. In particular, if \( \{e_{i,0}\}_{i=1}^{n} \) is an orthonormal frame at \( c(t_0) \), where \( t_0 \in \mathcal{I} \), then the unique parallel vector fields \( e_i \) along \( c \) with \( e_i (t_0) = e_{i,0} \), \( 1 \leq i \leq n \), satisfy
\[
\langle e_i (t), e_j (t) \rangle = \delta_{ij}
\]
for \( t \in \mathcal{I} \) and \( 1 \leq i, j \leq n \).

Let \((\mathcal{M}^n, g)\) be a Riemannian manifold and let \( \nabla \) be its Levi-Civita connection. We say that a path \( c : \mathcal{I} \to \mathcal{M} \) is a **geodesic** if its unit tangent vector field is parallel along \( C \), that is,
\[
\frac{\nabla}{dt} \left( \frac{c'}{|c'|} \right) = 0,
\]
where \( \nabla \) denotes the Levi-Civita covariant derivative along \( c \). The motivation for this definition will be apparent when we discuss the first variation of arc length of curves.
We say that a geodesic has **constant speed** if $|c'(t)|$ is constant for $t \in I$. We have that (2.37) and $|c'|$ being constant is equivalent to the **constant speed geodesic equation**:

\[ \nabla_{\dot{c}} \dot{c}' = 0. \]  

By (2.34) and (2.33), the constant speed geodesic equation, with respect to a parametrization, is:

\[ \frac{dx^k}{dt} (t) + \sum_{i,j=1}^{n} x^i(t) x^j(t) \Gamma^k_{ji} (c(t)) = 0, \]

where $x^j(t) = ((x^{-1})^j \circ c)'(t)$. If we define $c^j(t) = (x^{-1})^j (c(t))$, then this may be written as

\[ \frac{d^2 c^k}{dt^2} (t) + \sum_{i,j=1}^{n} \frac{dc^i}{dt} (t) \frac{dc^j}{dt} (t) (\Gamma^k_{ji} \circ x) (c_1, \ldots, c^n) (t) = 0. \]

Let $\bar{c}(t) \doteq (x^{-1} \circ c)(t) = (c_1(t), \ldots, c^n(t))$. Then this is a (nonlinear) system of ODE of the form

\[ \frac{d^2 \bar{c}}{dt^2} (t) = F \left( \frac{d\bar{c}}{dt} (t), \bar{c}(t) \right), \]

where $F : \mathbb{R}^n \times \mathcal{U} \to \mathbb{R}^n$ is a $C^\infty$ function. By the existence and uniqueness of solutions to systems of ODE, we have that given $q \in \mathcal{U}$ and $X \in \mathbb{R}^n$, there exists $\varepsilon > 0$ and a unique $C^\infty$ curve $\bar{c} : (-\varepsilon, \varepsilon) \to \mathcal{U}$ satisfying (2.40) and, equivalently, (2.39). Moreover, by the smooth dependence of solutions of ODE, $\bar{c}(t)$ is a $C^\infty$ function of $q$, $X$, and $t$. This translates to the same results for the solution $c(t) = x(\bar{c}(t))$ to (2.38).

### 7. The Riemann curvature tensor

The Lie bracket measures the noncommutativity of the directional derivative acting on functions: $[X,Y] f = X(Y(f)) - Y(X(f))$. By the torsion-free condition (4) in the definition of the Levi-Civita connection, the commutator of the first (Riemannian) covariant derivative is equal to the Lie bracket: $\nabla_X Y - \nabla_Y X = [X,Y]$. The Riemann curvature tensor is the commutator of the second covariant derivative acting on vector fields.

---

2Note that $F$ is quadratic in $\frac{d\bar{c}}{dt} (t)$. 

7. The Riemann curvature tensor

7.1. Tensors and $C^\infty(M)$-linearity.

We may extend Lemma 2.18 to the following. Let $\times^k \Xi(M) \cong \Xi(M) \times \cdots \times \Xi(M)$ ($k$-fold product) and let $T: \times^k \Xi(M) \rightarrow A$, where $A$ is $C^\infty(M)$ or $\Xi(M)$. We say that $T$ is a $C^\infty(M)$-multilinear map if it is $C^\infty(M)$-linear in each component, i.e., for each $1 \leq i \leq k,$

$$T(X_1, \ldots, X_{i-1}, X_i + Y_i, X_{i+1}, \ldots, X_k) = T(X_1, \ldots, X_{i-1}, X_i, X_{i+1}, \ldots, X_k) + T(X_1, \ldots, X_{i-1}, Y_i, X_{i+1}, \ldots, X_k)$$

and

$$T(X_1, \ldots, X_{i-1}, fX_i, X_{i+1}, \ldots, X_k) = fT(X_1, \ldots, X_{i-1}, X_i, X_{i+1}, \ldots, X_k)$$

for any $f \in C^\infty(M)$ and $C^\infty$ vector fields $X_*$ and $Y_*$. 

Remark 2.30. The map $T$ being $\mathbb{R}$-multilinear is the same definition with $f \in C^\infty(M)$ replaced by $f \in \mathbb{R}$.

Lemma 2.31. If $T$ is a $C^\infty(M)$-multilinear map, then for any $X_1, \ldots, X_k \in \Xi(M)$ we have that $T(X_1, \ldots, X_k)_p \in A$ depends only on $(X_1)_p, \ldots, (X_k)_p \in T_pM$.

Exercise 2.32. Prove this lemma.

If $T: \times^k \Xi(M) \rightarrow C^\infty(M)$ is a $C^\infty(M)$-multilinear map, then it is equivalent to the map

$$\bar{T}: \times^k T\mathcal{M} \rightarrow \mathbb{R}$$

defined by

$$\bar{T}\big((X_1)_p, \ldots, (X_k)_p\big) \doteq T(X_1, \ldots, X_k)_p$$

for arbitrary extensions of $(X_i)_p \in T_p\mathcal{M}$ to $X_i \in \Xi(M)$. Note that, by Lemma 2.31, we have that $\bar{T}$ is well defined and for each $p \in \mathcal{M}$, $\bar{T}|_{\times^k T_p\mathcal{M}}: \times^k T_p\mathcal{M} \rightarrow \mathbb{R}$ is $\mathbb{R}$-multilinear. For simplicity, we shall use the notation for both $T$ and $\bar{T}$. We call $T$ a covariant $k$-tensor.

Note that a 1-form is the same thing as a covariant 1-tensor. Indeed, recall that a 1-form $\alpha$ may be considered as a function $\alpha: T\mathcal{M} \rightarrow \mathbb{R}$ such that $\alpha|_{T_p\mathcal{M}}: T_p\mathcal{M} \rightarrow \mathbb{R}$ is linear for each $p \in \mathcal{M}$.

A Riemannian metric is the same as a covariant 2-tensor which is symmetric and positive definite.
7.2. Definition of $\text{R}m$ and its first properties.

More precisely, the **Riemann curvature** $(3,1)$-tensor

$$\text{R}m : \Xi (\mathcal{M}) \times \Xi (\mathcal{M}) \times \Xi (\mathcal{M}) \to \Xi (\mathcal{M})$$

is defined by

$$(2.41) \quad \text{R}m (X,Y) Z \triangleq \nabla_X (\nabla_Y Z) - \nabla_Y (\nabla_X Z) - \nabla_{\left[ X,Y \right]} Z$$

for $X,Y,Z \in \Xi (\mathcal{M})$. Note, for example, that $\nabla_Y Z \in \Xi (\mathcal{M})$, so that $\nabla_X (\nabla_Y Z)$ is well defined.

Clearly we have the antisymmetry:

$$(2.42) \quad \text{R}m (Y,X) Z = -\text{R}m (X,Y) Z.$$  

Furthermore, one may easily check that $\text{R}m$ is $\mathbb{R}$-multilinear. That is, for $c \in \mathbb{R}$,

$$(2.43) \quad \text{R}m (X_1 + cX_2, Y) Z = \text{R}m (X_1, Y) Z + c \text{R}m (X_2, Y) Z, \quad \text{R}m (X, Y_1 + cY_2) Z = \text{R}m (X, Y_1) Z + c \text{R}m (X, Y_2) Z, \quad \text{R}m (X, Y) (Z_1 + cZ_2) = \text{R}m (X, Y) Z_1 + c \text{R}m (X, Y) Z_2,$$

where the $X$’s, $Y$’s and $Z$’s are vector fields.

**Lemma 2.33.** For any $X,Y,Z \in \Xi (\mathcal{M})$ and any $C^\infty$ function $f : \mathcal{M} \to \mathbb{R}$, we have

$$(2.44) \quad \text{R}m (fX, Y) Z = \text{R}m (X, fY) Z = \text{R}m (X, Y) (fZ) = f \text{R}m (X, Y) Z.$$  

By Lemma 2.31, $(\text{R}m (X,Y) Z)_p \in T_p \mathcal{M}$ depends only on $X_p, Y_p, Z_p \in T_p \mathcal{M}$.

**Proof.** (1) We compute that

$$\text{R}m (fX, Y) Z = \nabla_{fX} (\nabla_Y Z) - \nabla_Y (\nabla_{fX} Z) - \nabla_{\left[ fX,Y \right]} Z$$

$$= f \nabla_X (\nabla_Y Z) - (f \nabla_Y (\nabla_X Z) + Y (f) \nabla_X Z) - \nabla_{f\left[ X,Y \right] - Y(f)X} Z$$

$$= f \text{R}m (X, Y) Z.$$

(2) Since $\text{R}m (X,Y) Z$ is antisymmetric in $X$ and $Y$, we have

$$\text{R}m (X, fY) Z = f \text{R}m (X, Y) Z.$$
(3) We also compute
\[
R_m (X,Y) fZ = \nabla_X (\nabla_Y (fZ)) - \nabla_Y (\nabla_X (fZ)) - \nabla_{[X,Y]} (fZ)
\]
\[
= \nabla_X ((Yf) Z + f\nabla_Y Z) - \nabla_Y ((Xf) Z + f\nabla_X Z)
\]
\[
- ([X,Y] f) Z - f\nabla_{[X,Y]} Z
\]
\[
= X (Yf) Z + (Yf) \nabla_X Z + (Xf) \nabla_Y Z + f\nabla_X (\nabla_Y Z)
\]
\[
- Y (Xf) Z - (Xf) \nabla_Y Z - (Yf) \nabla_X Z - f\nabla_Y (\nabla_X Z)
\]
\[
- ([X,Y] f) Z - f\nabla_{[X,Y]} Z
\]
\[
= f R_m (X,Y) Z,
\]
where we used \(X (Yf) - Y (Xf) - [X,Y] f = 0\) and cancelled terms to obtain the last equality. □

For each \(p \in M\), since \(R_m (X,Y) Z) \in T_p M\) depends only on \(X_p, Y_p, Z_p \in T_p M\), there exists a well-defined map
\[
R_m : T_p M \times T_p M \times T_p M \to T_p M
\]
given by
\[
R_m (X_p, Y_p, Z_p) = (R_m (X,Y) Z)_p,
\]
where \(X, Y, Z \in \Omega(M)\) are any vector fields satisfying \(X (p) = X_p, Y (p) = Y_p \) and \(Z (p) = Z_p\). Since \(R_m\) is \(C^\infty(M)\)-multilinear, we obtain that \(R_m\) is \(\mathbb{R}\)-multilinear. Since 3 copies of \(T_p M\) are mapped into 1 copy of \(T_p M\), we say that \(R_m\) is a **tensor of type** \((3,1)\) or that \(R_m\) is a \((3,1)\)-**tensor**.

**Exercise 2.34.** Let \(\varphi : (N^n, h) \to (M^n, g)\) be an isometry. Let \(R_m^g\) and \(R_m^h\) denote the Riemann curvature tensors of \(g\) and \(h\), respectively. Prove that
\[
(2.45) \quad R_m^h (X,Y) Z = (d\varphi)^{-1} (R_m^g (d\varphi (X), d\varphi (Y)) d\varphi (Z)).
\]

It is convenient to define the **second covariant derivative** of a vector field by
\[
\nabla^2 : \Omega(M) \times \Omega(M) \times \Omega(M) \to \Omega(M),
\]
where
\[
(2.46) \quad \nabla^2 (X,Y,Z) \doteq \nabla^2_{X,Y} Z \doteq \nabla_X (\nabla_Y Z) - \nabla_{\nabla_X Y} Z.
\]

Using \(\nabla_X Y - \nabla_Y X = [X,Y]\), we see that \(R_m\) is the commutator of the second covariant derivative:
\[
(2.47) \quad R_m (X,Y) Z = \nabla^2_{X,Y} Z - \nabla^2_{Y,X} Z.
\]

We shall later see why this definition is natural from the point of view of ‘tensors’. One computes that for \(f \in C^\infty(M)\),
\[
(2.48) \quad \nabla^2_{fX,Y} Z = \nabla^2_{X,fY} Z = f \nabla^2_{X,Y} Z.
\]
By Lemma 2.31, $\nabla^2 (X,Y,Z)_p$ depends only on $X_p, Y_p \in T_pM$ and $Z \in \Xi (M)$. We also compute that

$$\nabla^2_{X,Y} (fZ) = f\nabla^2_{X,Y} Z + Y (f) \nabla_X Z + X (f) \nabla_Y Z$$

Note that (2.44) also follows from (2.48), (2.49), and (2.47).

**Exercise 2.35.** Prove (2.48) and derive $\text{Rm} (X,Y) (fZ) = f \text{Rm} (X,Y) Z$ from (2.49).

### 7.3. Algebraic identities of the curvature and the first Bianchi identity.

The **Riemann curvature covariant 4-tensor** (or $(4,0)$-tensor) is defined by

$$\text{Rm} (X,Y,Z,W) \equiv \langle \text{Rm} (X,Y) Z,W \rangle .$$

Some basic symmetries of the Riemann curvature tensor are

$$\text{Rm} (X,Y,Z,W) = - \text{Rm} (Y,X,Z,W) = - \text{Rm} (X,Y,W,Z) = \text{Rm} (Z,W,X,Y) .$$

The first equality in (2.50) is (2.42). We prove the second and third equalities below (see (2.51) and (2.52)).

Besides (2.50) there are additional symmetries that the Riemann curvature tensor satisfies.

**Proposition 2.36.** The Riemann curvature tensor satisfies the **first Bianchi identity**, i.e.,

$$\text{Rm} (X,Y) Z + \text{Rm} (Y,Z) X + \text{Rm} (Z,X) Y = 0 .$$

**Proof.** To see the first Bianchi identity, we simply sum the following three formulas (by definition)

$$\begin{align*}
\text{Rm} (X,Y) Z &= \nabla_X (\nabla_Y Z) - \nabla_Y (\nabla_X Z) - \nabla_{[X,Y]} Z, \\
\text{Rm} (Y,Z) X &= \nabla_Y (\nabla_Z X) - \nabla_Z (\nabla_Y X) - \nabla_{[Y,Z]} X, \\
\text{Rm} (Z,X) Y &= \nabla_Z (\nabla_X Y) - \nabla_X (\nabla_Z Y) - \nabla_{[Z,X]} Y,
\end{align*}$$

while applying the torsion-free property $\nabla_{[V,W]} - \nabla_{W,V} = [V,W]$, to obtain

$$\begin{align*}
\text{Rm} (X,Y) Z + &\text{Rm} (Y,Z) X + \text{Rm} (Z,X) Y \\
= &\nabla_X [Y,Z] + \nabla_Y (Z,X) + \nabla_Z [X,Y] \\
- &\nabla_{[Y,Z]} X - \nabla_{[Z,X]} Y - \nabla_{[X,Y]} Z \\
= &\left[ X, [Y,Z] \right] + \left[ Y, [Z,X] \right] + \left[ Z, [X,Y] \right] \\
= &\left[ X, [Y,Z] \right] + \left[ Y, [Z,X] \right] + \left[ Z, [X,Y] \right] \\
= &0
\end{align*}$$

by the Jacobi identity (1.14).
We also have the following identities

**Proposition 2.37.** The Riemann curvature tensor is antisymmetric in the last two components, i.e.,

\[ \langle Rm (X,Y) Z, W \rangle = - \langle Rm (X,Y) W, Z \rangle \]

and is symmetric in the switch of the first pair with the second pair, i.e.,

\[ \langle Rm (X,Y) Z, W \rangle = \langle Rm (Z,W) X,Y \rangle . \]

**Proof.** First we prove (2.51). A straightforward computation, where we move covariant derivatives off of \( Z \) and onto \( W \), yields

\[
\langle Rm (X,Y) Z, W \rangle = \langle \nabla_X (\nabla_Y Z) - \nabla_Y (\nabla_X Z) - [X,Y] Z, W \rangle \\
= X \langle \nabla_Y Z, W \rangle - \langle \nabla_Y Z, \nabla_X W \rangle - Y \langle \nabla_X Z, W \rangle \\
+ \langle \nabla_X Z, \nabla_Y W \rangle - [X,Y] \langle Z, W \rangle + \langle Z, \nabla_{[X,Y]} W \rangle \\
= XY \langle Z, W \rangle - YX \langle Z, W \rangle - [X,Y] \langle Z, W \rangle \\
+ \langle \nabla_X Z, \nabla_Y W \rangle - X \langle Z, \nabla_Y W \rangle \\
- \langle \nabla_Y Z, \nabla_X W \rangle + Y \langle Z, \nabla_X W \rangle + \langle Z, \nabla_{[X,Y]} W \rangle \\
= - \langle Z, \nabla_X \nabla_Y W \rangle + \langle Z, \nabla_Y \nabla_X W \rangle + \langle Z, \nabla_{[X,Y]} W \rangle \\
= - \langle Rm (X,Y) W, Z \rangle ,
\]

which is (2.51).

Next we prove (2.52). By the first Bianchi identity we have

\[
\langle Rm (X,Y) Z, W \rangle^{[1]} + \langle Rm (Y,Z) X, W \rangle^{[2]} + \langle Rm (Z,X) Y, W \rangle = 0, \\
\langle Rm (Y,Z) W, X \rangle^{[2]} + \langle Rm (Z,W) Y, X \rangle^{[3]} + \langle Rm (W,Y) Z, X \rangle = 0, \\
\langle Rm (Z,W) X, Y \rangle^{[3]} + \langle Rm (W,X) Z, Y \rangle^{[4]} + \langle Rm (X,Z) W, Y \rangle = 0, \\
\langle Rm (W,X) Y, Z \rangle^{[4]} + \langle Rm (X,Y) W, Z \rangle^{[1]} + \langle Rm (Y,W) X, Z \rangle = 0,
\]

where the superscripts \( \Box \), \( \heartsuit = 1,2,3,4 \), denote pairs of terms which cancel by applying (2.51) when the four equations above are added together. Hence

\[ \langle Rm (Z,X) Y, W \rangle + \langle Rm (W,Y) Z, X \rangle = 0, \]

which, by one final application of (2.51), yields (2.52). \( \square \)

### 7.4. Riemann curvature tensor in local coordinates.

Given a parametrization \((U,x)\), the **components** of the \((3,1)\)-tensor \( Rm \) are defined by

\[
Rm \left( \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) \frac{\partial}{\partial x^k} = \sum_{\ell=1}^n R_{ijk}^\ell \frac{\partial}{\partial x^\ell}.
\]
for \( i, j, k = 1, \ldots, n \). The components of \( \text{Rm} \) as a \((4,0)\)-tensor are

\[
R_{ijkl} = \sum_{m=1}^{n} g_{m} R_{ijk}^{m}.
\]

In a parametrization, (2.50) may be written as

\[
R_{ijkl} = -R_{jikl} = -R_{ijlk} = R_{klij}.
\]

Lemma 2.38 (Components of the Riemann curvature \((3,1)\)-tensor). We have

\[
R_{ij}^{k} = \partial_{i} \Gamma_{jk}^{l} - \partial_{j} \Gamma_{ik}^{l} + \sum_{p=1}^{n} \left( \Gamma_{jk}^{p} \Gamma_{ip}^{l} - \Gamma_{ik}^{p} \Gamma_{jp}^{l} \right) \partial_{l},
\]

where \( \partial_{i} \equiv \frac{\partial}{\partial x^{i}} \).

Proof. Using (2.41), (2.23), and \( [\partial_{i}, \partial_{j}] = 0 \), we compute

\[
R_{ij}^{k} \partial_{l} = \text{Rm} \left( \partial_{i}, \partial_{j} \right) \partial_{k}
\]

\[
= \nabla_{\partial_{i}} \left( \nabla_{\partial_{j}} \partial_{k} \right) - \nabla_{\partial_{j}} \left( \nabla_{\partial_{i}} \partial_{k} \right)
\]

\[
= \nabla_{\partial_{i}} \left( \sum_{\ell=1}^{n} \Gamma_{jk}^{\ell} \partial_{\ell} \right) - \nabla_{\partial_{j}} \left( \sum_{\ell=1}^{n} \Gamma_{ik}^{\ell} \partial_{\ell} \right)
\]

\[
= \partial_{i} \Gamma_{jk}^{\ell} - \partial_{j} \Gamma_{ik}^{\ell} + \sum_{p=1}^{n} \left( \Gamma_{jk}^{p} \nabla_{\partial_{i}} \partial_{p} - \Gamma_{ik}^{p} \nabla_{\partial_{j}} \partial_{p} \right)
\]

\[
= \left( \partial_{i} \Gamma_{jk}^{\ell} - \partial_{j} \Gamma_{ik}^{\ell} + \sum_{p=1}^{n} \left( \Gamma_{jk}^{p} \Gamma_{ip}^{\ell} - \Gamma_{ik}^{p} \Gamma_{jp}^{\ell} \right) \right) \partial_{l}.
\]

In a parametrization, the first Bianchi identity is

\[
R_{ijk}^{l} + R_{jik}^{l} + R_{kij}^{l} = 0.
\]

8. The sectional curvature

If \( P \subset T_{x} \mathcal{M} \) is a 2-plane, then the sectional curvature of \( P \) is defined by

\[
K \left( P \right) \equiv \langle \text{Rm} \left( e_{1}, e_{2} \right) e_{2}, e_{1} \rangle,
\]

where \( \{e_{1}, e_{2}\} \) is an orthonormal basis of \( P \); this definition is independent of the choice of such a basis.
Exercise 2.39. Show that if $X$ and $Y$ are any two vectors spanning $P$, then

$$K(P) = \frac{\langle \text{Rm}(X,Y)Y,X \rangle}{|X|^2|Y|^2 - \langle X,Y \rangle^2}. \quad (2.56)$$

9. The Ricci tensor

The Ricci tensor is the covariant 2-tensor defined by

$$\text{Rc}(X,Y) = \sum_{i=1}^{n} \langle \text{Rm}(X,e_i)e_i,Y \rangle. \quad (2.57)$$

Exercise 2.40. Show that $\text{Rc}$ is symmetric.

We may also consider $\text{Rc}$ as the linear map

$$\text{Rc} : T_pM \to T_pM$$

defined by

$$\text{Rc}(X,Y) \triangleq \langle \text{Rc}(X),Y \rangle$$

for each $X,Y \in T_pM$.

Exercise 2.41 (Geometric interpretation of tracing). Show that the trace of a symmetric 2-tensor $\alpha$ is given by the following formula:

$$\text{Trace}_g(\alpha) = \frac{1}{\omega_n} \int_{S^{n-1}} \alpha(V,V) \, d\sigma(V),$$

where $S^{n-1}$ is the unit $(n-1)$-sphere, $\omega_n$ its volume, and $d\sigma$ its volume form. From this, show for any unit vector $U$ that $\frac{1}{n-1} \text{Rc}(U,U)$ is the average of the sectional curvatures of planes containing the vector $U$. Similarly, $\frac{1}{n} R(p)$ is the average of $\text{Rc}(U,U)$ over all unit vectors $U \in S^{n-1} \subset T_pM$.

Solution. There exists an orthonormal basis $\{e_i\}_{i=1}^{n}$ such that $\alpha = \sum_{i=1}^{n} \lambda_i e_i^* \otimes e_i^*$. Furthermore, $\text{Trace}_g(\alpha) = \sum_{i=1}^{n} \lambda_i$ and

$$\frac{1}{\omega_n} \int_{S^{n-1}} \langle V,e_i \rangle^2 \, d\sigma(V) = 1.$$

See also [5].

10. Pull backs and Lie derivatives of tensors

Let $T$ be a $C^\infty$ covariant $k$-tensor on a $C^\infty$ manifold $\mathcal{M}^n$. Generalizing the notion of the pull back of a 1-form, i.e., a covariant 1-tensor, we define the pull back of $T$ by a $C^\infty$ map $\varphi : \mathcal{N} \to \mathcal{M}$ as

$$(\varphi^*T)(X_1,\ldots,X_k) \triangleq T(d\varphi(X_1),\ldots,d\varphi(X_k)).$$

This generalizes definition (1.21), which is the case where $k = 1$. 
For example, for a Euclidean hypersurface $\mathcal{M}^n \subset \mathbb{R}^{n+1}$ parametrized by a map $x : \mathcal{U} \to \mathcal{M}$, definition (2.8) says that

$$g \doteq x^*(I),$$

i.e., the Riemannian metric $g$ on $\mathcal{U}$ is defined to be the pull back by $x$ of the Riemannian metric $I$ on $\mathcal{M}$.

More generally, given a Riemannian manifold $(\mathcal{M}^n, g)$ and a $C^\infty$ immersion $\varphi : \mathcal{N} \to \mathcal{M}$, i.e., for each $q \in \mathcal{N}$ we have $d\varphi_{q}$ is injective, the pull back metric on $\mathcal{N}$ is defined by

$$\left(\varphi^* g \right) (X,Y) = g \left( d\varphi \left( X \right), d\varphi \left( Y \right) \right).$$

Thus definition (2.1) of $\varphi : (\mathcal{N}^n, h) \to (\mathcal{M}^n, g)$ being an isometry is the same as $\varphi^* g = h$.

Directly generalizing (1.22), we may now define the Lie derivative of a covariant $k$-tensor $T$ with respect to a $C^\infty$ vector field $X$ as:

$$\left( L_X T \right)_p \doteq \lim_{t \to 0} \frac{1}{t} \left( \varphi_t^* \left( T_{\varphi_t(p)} \right) - T_p \right).$$

Note that if $\varphi_t^* \left( T_{\varphi_t(p)} \right) = T_p$ for $p \in \mathcal{M}$ and all $t$ such that $\varphi_t \left( p \right)$ is defined, then $L_X T = 0$. For example, if $X$ is such that $\varphi_t$ are (local) isometries, then $L_X g = 0$.

**Definition 2.42.** We say that $X$ is a Killing vector field if $L_X g = 0$.

The following is the product rule for the Lie derivative.

**Proposition 2.43.**

(2.58)

$$\left( L_X T \right) (Y_1, \ldots, Y_k) = X (T(Y_1, \ldots, Y_k)) - \sum_{i=1}^{k} T(Y_1, \ldots, [X,Y_i], \ldots, Y_k),$$

or equivalently,

$$L_X \left( T(Y_1, \ldots, Y_k) \right) = \left( L_X T \right) (Y_1, \ldots, Y_k) + \sum_{i=1}^{k} T(Y_1, \ldots, L_X Y_i, \ldots, Y_k).$$

10.1. Pulled back bundles and affine connections.

Let $\varphi : \mathcal{N}^n \to \mathcal{M}^m$ be a $C^\infty$ map. The pulled back tangent bundle $\varphi^* T\mathcal{M}$ is defined by

(2.59)

$$\varphi^* T\mathcal{M} = \bigsqcup_{x \in \mathcal{N}} T_{\varphi(x)}\mathcal{M}.$$

Equivalently, we may define it as the point-set

(2.60)

$$\varphi^* T\mathcal{M} = \{ (x, v) : x \in \mathcal{N}, \ v \in T_{\varphi(x)}\mathcal{M} \}.$$
We have the projection map $\pi : \varphi^*T_M \to N$ defined by $\pi(x,v) = x$. Note that $\pi^{-1}(x) = T_{\varphi(x)}M$.

**Exercise 2.44.** Show that $\varphi^*T_M$ is a $C^\infty$ manifold of dimension $n + m$.

11. Covariant derivative of a tensor

The **covariant derivative** of a covariant $k$-tensor $T$ is defined by the product rule:

\[ (\nabla_X T)(Y_1, \ldots, Y_k) = X(T(Y_1, \ldots, Y_k)) - \sum_{i=1}^k T(Y_1, \ldots, \nabla_X Y_i, \ldots, Y_k). \]

If $\nabla_X T = 0$ for each $X \in \mathfrak{X}(M)$, then we say that $T$ is **parallel**.

**Exercise 2.45.** Show that for any vector field $X$, $\nabla_X g = 0$. That is, the Riemannian metric is parallel.

By the torsion-free property for the Levi-Civita, namely $\nabla_X Y_i - \nabla_Y X_i = [X,Y_i]$, we may combine the above two propositions to obtain:

\[ (\mathcal{L}_X T)(Y_1, \ldots, Y_k) = (\nabla_X T)(Y_1, \ldots, Y_k) + \sum_{i=1}^k T(Y_1, \ldots, \nabla_Y X_i, \ldots, Y_k). \]

**Exercise 2.46.** Show that

\[(\mathcal{L}_X g)(Y, Z) = g(\nabla_Y X, Z) + g(Y, \nabla_Z X).\]

Given a covariant $k$-tensor

\[ T : \times^k T_M \to \mathbb{R}, \]

we define the covariant $(k + 1)$-tensor $\nabla T : \times^{k+1}T_M \to \mathbb{R}$ by

\[ (\nabla T)(X_0, X_1, \ldots, X_k) = (\nabla_{X_0} T)(X_1, \ldots, X_k) \]

for $p \in M$ and $X_i \in T_p M$ for $i = 0, 1, \ldots, k$.

**Exercise 2.47.** Show that if $T$ is a symmetric 2-tensor, i.e., $T(Y, Z) = T(Z, Y)$, then

\[ \nabla T)(X, Y, Z) = \nabla T)(X, Z, Y). \]

On the other hand, given a vector field $Y$ on $\mathcal{M}$, we may consider its covariant derivative as

\[ \nabla Y : T\mathcal{M} \to T\mathcal{M} \]

defined by

\[ (\nabla Y)(X) = \nabla_X Y. \]

Note that $\nabla Y|_{T_p\mathcal{M}} : T_p\mathcal{M} \to T_p\mathcal{M}$ is linear for each $p \in \mathcal{M}$. Because of this, we say that $\nabla Y$ is a $(1,1)$-tensor (or tensor of type $(1,1)$). Now given
a \((1,1)\)-tensor \(S : TM \to TM\), we may define its covariant derivative as the bilinear map
\[
(\nabla S)_p : T_p M \times T_p M \to T_p M
\]
given by
\[
(\nabla S)_p (X_p, Y_p) = (\nabla_{X_p} S)_p (Y_p)
\]
for \(X_p, Y_p \in T_p M\).

**Exercise 2.48.** Show that if \(Z\) is a \(C^\infty\) vector field on \(M\), then for any \(p \in M\) and \(X_p, Y_p \in T_p M\),
\[
(\nabla \nabla Z)_p (X_p, Y_p) \equiv (\nabla (\nabla Z))_p (X_p, Y_p) = (\nabla^2_{X,Y} Z)_p,
\]
where the right-hand side is defined by (2.46). Hence
\[
\text{Rm}(X,Y) Z = (\nabla \nabla Z)(X,Y) - (\nabla \nabla Z)(Y,X).
\]

Let \(\alpha\) be a 1-form. Then
\[
(\nabla \nabla \alpha)(X,Y,Z) - (\nabla \nabla \alpha)(Y,X,Z) =
\]
\[
= \nabla_X (\nabla \alpha)(Y,Z) - \nabla_Y (\nabla \alpha)(X,Z)
\]
\[
= X ((\nabla \alpha)(Y,Z)) - \nabla \alpha (\nabla_Y X, Z) - \nabla \alpha (Y, \nabla_X Z)
\]
\[
= X ((\nabla \alpha)(X,Z)) + \nabla \alpha (\nabla_Y X, Z) + \nabla \alpha (X, \nabla_Y Z)
\]
\[
= X ((\nabla Y \alpha)(Z)) - (\nabla |X,Y| \alpha)(Z) - \nabla Z \alpha (\nabla_X Z)
\]
\[
= Y ((\nabla X \alpha)(Z)) + \nabla X \alpha (\nabla_Y Z)
\]
\[
= [X,Y] (\alpha (Z)) - (\nabla |X,Y| \alpha)(Z) + \alpha (\nabla_Y (\nabla_X Z)) - \alpha (\nabla_X (\nabla_Y Z))
\]
\[
= -\alpha (\text{Rm}(X,Y) Z).
\]

12. **Connection 1-forms, curvature 2-forms and the Cartan structure equations**

Besides using a parametrization, one may use a local orthonormal basis of vector fields to compute the Levi-Civita connection and the Riemann curvature tensor.

12.1. **Exterior derivative of 1-forms.**

The **exterior derivative** of a 1-form \(\alpha\) is defined by
\[
(d\alpha)(X,Y) = \frac{1}{2} (X (\alpha (Y)) - Y (\alpha (X)) - \alpha ([X,Y]))
\]
for \(X,Y \in \Xi(M)\). Clearly, \((d\alpha)(Y,X) = - (d\alpha)(X,Y)\).

**Exercise 2.49.** Check that for any \(f \in C^\infty(M)\) and \(X,Y \in \Xi(M)\),
\[
(d\alpha)(fX,Y) = f (d\alpha)(X,Y) = (d\alpha)(X,fY).
\]
Hence, $d\alpha$ is an alternating covariant 2-tensor. We call such an object a 2-form.

**Exercise 2.50.** Show that for any $f \in C^\infty(M)$,

$$d^2 f \doteq d(df) = 0.$$  

**Lemma 2.51.** For any 1-form $\alpha$,

$$\begin{aligned}
(d\alpha)(X,Y) &= \frac{1}{2} ((\nabla_X \alpha)(Y) - (\nabla_Y \alpha)(X)).
\end{aligned}$$

**Proof.** We compute, using (2.26), that

$$\begin{aligned}
(\nabla_X \alpha)(Y) - (\nabla_Y \alpha)(X) &= X(\alpha(Y)) - \alpha(\nabla_X Y) - Y(\alpha(X)) + \alpha(\nabla_Y X) \\
&= X(\alpha(Y)) - Y(\alpha(X)) - \alpha([X,Y]),
\end{aligned}$$

where we used the torsion-free property of the Levi-Civita connection. $\square$  

The **wedge product** of two 1-forms $\beta$ and $\gamma$ is

$$\begin{aligned}
(\beta \wedge \gamma)(X,Y) &= \frac{1}{2} (\beta (X) \gamma (Y) - \beta (Y) \gamma (X)).
\end{aligned}$$

It is easy to see that $\beta \wedge \gamma$ is a 2-form and that $\beta \wedge \gamma = -\gamma \wedge \beta$.

**12.2. Orthonormal frames and their dual coframes.**

Now let $g$ be a Riemannian metric on $M$. Given $p \in M$ and $X \in T_pM$, define $g : T_pM \to T^*_pM$ by

$$g(X)(Y) \doteq g(X,Y)$$

for $Y \in T_pM$. ($g(X) \in T^*_pM$ since $Y \mapsto g(X,Y)$ is linear).

**Exercise 2.52.** Show that $g : T_pM \to T^*_pM$ is a bijective linear map. Show that $g^{-1} : T^*_pM \to T_pM$ is given by

$$g^{-1}(\alpha)(Y) = \alpha(Y).$$

Define the inner product $g^*$ on $T^*_pM$ by

$$g^*(\alpha, \beta) = g(g^{-1}(\alpha), g^{-1}(\beta))$$

for $\alpha, \beta \in T^*_pM$. We often write $\langle \alpha, \beta \rangle = g^*(\alpha, \beta)$.

Given $\omega, \eta \in T^*_pM$, define the bilinear map

$$\omega \otimes \eta : T_pM \times T_pM \to \mathbb{R}$$

by

$$\langle \omega \otimes \eta \rangle(X,Y) = \omega(X) \eta(Y).$$

$\omega \otimes \eta$ is called the **tensor product** of $\omega$ and $\eta$.

Let $\{e_i\}_{i=1}^n$ be an orthonormal basis for $T_pM$. Let $\{\omega^i\}_{j=1}^n$ be the dual basis of 1-forms, i.e., $\omega^i(e_i) = \delta^i_j$. 
Exercise 2.53.  

(1) Show that \( g^* (\omega^j, \omega^k) = \delta^{jk} \).

(2) Show that \( g = \sum_{j=1}^n \omega^j \otimes \omega^j \).

12.3. Connection 1-forms.

Let \( O \) be an open set in \( M \) and let \( \{ e_i \}_{i=1}^n \) be \( C^\infty \) vector fields that form an orthonormal basis at each point in \( O \). Define the connection 1-forms \( \omega^i_j \) on \( O \) by

\[
\nabla_X e_i = \sum_{j=1}^n \omega^i_j (X) e_j.
\]

Exercise 2.54. Prove that \( \omega^i_j = -\omega^j_i \) for \( i, j = 1, \ldots, n \).

Solution. For any \( X \in T O \),

\[
0 = X \langle e_i, e_j \rangle = \langle \nabla_X e_i, e_j \rangle + \langle e_i, \nabla_X e_j \rangle = \omega^i_j (X) + \omega^j_i (X).
\]

Let \( \{ \omega^j \}_{j=1}^n \) denote the dual basis of 1-forms on \( O \), i.e., \( \omega^j (e_i) = \delta^j_i \).

Note that

\[
g = \sum_{j=1}^n \omega^j \otimes \omega^j.
\]

From (2.71) and (2.26), we obtain

\[
(\nabla_X \omega^j) (e_j) = X (\omega^j (e_j)) - \omega^j (\nabla_X e_j)
\]

\[
= -\sum_{j=1}^n \omega^j_j (X) \omega^j.
\]

12.4. Curvature 2-forms.

The curvature 2-forms \( \Omega^j_i \) on \( O \) are defined by

\[
\Omega^j_i (X, Y) e_j = \frac{1}{2} \mathrm{Rm} (X, Y) e_i,
\]

so that \( \Omega^j_i (X, Y) = \frac{1}{2} \langle \mathrm{Rm} (X, Y) e_i, e_j \rangle \).
Theorem 2.55 (Cartan structure equations). The first and second Cartan structure equations are

\[ d\omega^i = \sum_{j=1}^{n} \omega^j \wedge \omega^i_j, \]

(2.74)

\[ \Omega^i_j = d\omega^i_j - \sum_{k=1}^{n} \omega^k_i \wedge \omega^j_k, \]

(2.75)

respectively.

Proof. (1) We compute

\[ d\omega^i (X,Y) = \frac{1}{2} (\nabla_X \omega^i)(Y) - \frac{1}{2} (\nabla_Y \omega^i)(X) \]

\[ = \sum_{j=1}^{n} \left( -\frac{1}{2} \omega^i_j (X) \omega^j (Y) + \frac{1}{2} \omega^i_j (Y) \omega^j (X) \right) \]

and (2.74) follows from the definition of wedge product.

(2) (2.46)

\[ \nabla_2 X,Y e_i = \nabla_X (\nabla_Y e_i) - \nabla_2 Y,X e_i \]

\[ = \sum_{k=1}^{n} \left( \nabla_X \left( \omega^k_i (Y) e_k \right) - \omega^k_i (\nabla_X Y) e_k \right) \]

\[ = \sum_{k=1}^{n} \left( \left( \nabla_X \omega^k_i \right) (Y) e_k + \omega^k_i (\nabla_X Y) e_k + \omega^k_i (Y) \nabla_X e_k - \omega^k_i (\nabla_X Y) e_k \right) \]

\[ = \sum_{k=1}^{n} \left( \left( \nabla_X \omega^k_i \right) (Y) e_k + \omega^k_i (Y) \nabla_X e_k \right). \]

From (2.47) and this, we have

\[ \Omega^i_j (X,Y) \]

\[ = \frac{1}{2} \langle \nabla_2 X,Y e_i - \nabla_2 Y,X e_i, e_j \rangle \]

\[ = \frac{1}{2} \sum_{k=1}^{n} \left( \left( \nabla_X \omega^k_i \right) (Y) e_k + \omega^k_i (Y) \nabla_X e_k - \left( \nabla_Y \omega^k_i \right)(X) e_k - \omega^k_i (X) \nabla_Y e_k, e_j \right) \]

\[ = \sum_{k=1}^{n} d\omega^k_i (X,Y) \langle e_k, e_j \rangle \]

\[ + \frac{1}{2} \sum_{k,l=1}^{n} \left( \omega^k_i (Y) \omega^l_k (X) - \omega^k_l (X) \omega^l_k (Y) \right) \langle e_k, e_l \rangle \]

and (2.75) follows. \qed
Exercise 2.56 (Formula for connection 1-forms). Using the first structure equation (2.74), show that
\[
d\omega^k (e_i, e_j) = \frac{1}{2} \left( \omega^k_i (e_j) - \omega^k_j (e_i) \right).
\]

Using this, derive the formula for the connection 1-forms:
\[
(2.76) \quad \omega^k_i (e_j) = d\omega^j (e_i, e_k) + d\omega^j (e_i, e_k) - d\omega^k (e_j, e_i).
\]

Note the similarity between this and the formula for the Christoffel symbols (2.24).

Solution. We compute that
\[
d\omega^j (e_j, e_k) + d\omega^j (e_i, e_k) - d\omega^k (e_j, e_i)
= \sum_{\ell=1}^n \left( (\omega^\ell \land \omega^j_k) (e_j, e_k) + (\omega^\ell \land \omega^j_i) (e_i, e_k) - (\omega^\ell \land \omega^k_i) (e_j, e_i) \right)
= \frac{1}{2} \left( \omega^j_k (e_k) - \omega^i_k (e_j) + \omega^j_i (e_k) - \omega^j_k (e_i) + \omega^k_i (e_j) \right)
= \omega^k_i (e_j).
\]

An equivalent formulation of the second Bianchi identity is the following. Let
\[
(d_\nabla \Omega)^j_i \overset{\cdot}{=} d\Omega^j_i + \sum_{k=1}^n \left( -\omega^k_i \land \Omega^j_k + \omega^j_k \land \Omega^k_i \right)
\]
denote the exterior covariant derivative of \( \Omega \) considered as a 2-form with values in \( T^*M \otimes TM \). Here \( T^*M \otimes TM \) denotes the vector bundle whose fiber at \( p \in M \) is the vector space of linear maps from \( T_pM \) to \( T_pM \).

Lemma 2.57 (Second Bianchi identity). We have
\[
(d_\nabla \Omega)^j_i = 0.
\]
Proof. Using the second structure equations (2.75), we compute

\[ d\Omega^j_i + \sum_{k=1}^n \left( -\omega^k_i \wedge \Omega^j_k + \omega^j_k \wedge \Omega^k_i \right) \]

\[ = \sum_{k=1}^n \left( -d\omega^k_i \wedge \omega^j_k + \omega^j_k \wedge d\omega^k_i \right) \]

\[ + \sum_{k=1}^n \left( -\omega^k_i \wedge (d\omega^j_k - \sum_{\ell=1}^n \omega^\ell_k \wedge \omega^j_\ell) + \omega^j_k \wedge (d\omega^k_i - \sum_{\ell=1}^n \omega^\ell_i \wedge \omega^k_\ell) \right) \]

\[ = \sum_{k,\ell=1}^n \left( \omega^k_i \wedge \omega^\ell_k \wedge \omega^j_\ell - \omega^j_k \wedge \omega^\ell_i \wedge \omega^k_\ell \right) \]

\[ = 0 \]

after switching \( k \) and \( \ell \) in one of the terms in the last line. \( \Box \)

13. The Gauss-Bonnet formula

For a surface \( M^2 \), since \( \omega^1_1 = \omega^2_2 = 0 \), we have that \( \omega^1_2 = -\omega^2_1 \) is the only nonzero connection 1-form. The Cartan structure equations simplify to

\[ d\omega^1 = \omega^2 \wedge \omega^1, \quad d\omega^2 = \omega^1 \wedge \omega^2, \]

and

\[ \Omega^1_2 = d\omega^1. \]

Hence, the Gauss curvature, defined as the only sectional curvature (2.55), is given by

\[ K \equiv \langle \text{Rm} (e_1, e_2) e_2, e_1 \rangle = 2\Omega^1_2 (e_1, e_2) = 2d\omega^1 (e_1, e_2) \]

using (2.73). The scalar curvature is

\[ R = \sum_{i,j=1}^2 \langle \text{Rm} (e_i, e_j) e_j, e_i \rangle = \langle \text{Rm} (e_1, e_2) e_2, e_1 \rangle + \langle \text{Rm} (e_2, e_1) e_1, e_2 \rangle = 2K. \]

Exercise 2.58. Let \( U \subset \mathbb{R}^2 \) be an open set with coordinates \((r, \theta)\). Suppose that \( f : U \to \mathbb{R}^+ \) depends only on \( r \) and write it as \( f(r) \). Show that the Gauss curvature of a (rotationally symmetric) metric on \( U \) of the form

\[ g = dr \otimes dr + f(r)^2 d\theta \otimes d\theta \]

is given by

\[ K = -\frac{f''(r)}{f(r)}. \]
SOLUTION. A moving coframe is given by $\omega^1 = dr$ and $\omega^2 = f(r) d\theta$, which is dual to the moving frame $e_1 = \frac{\partial}{\partial r}$ and $e_2 = \frac{1}{f(r)} \frac{\partial}{\partial \theta}$. From (2.76) we calculate that

$$\omega^1(e_j) = d\omega^2(e_j, e_1) + d\omega^2(e_2, e_1) - d\omega^1(e_j, e_2),$$

so that

$$\omega^1(e_1) = d\omega^1(e_2, e_1) = 0,$$

$$\omega^1(e_2) = d\omega^2(e_2, e_1) = -\frac{f''(r)}{f(r)}.$$

So

$$\omega^2 = -f'(r) d\theta,$$

$$d\omega^2 = -f''(r) dr \wedge d\theta.$$

We conclude that

$$K = 2 d\omega^1(e_1, e_2) = -\frac{f''(r)}{f(r)}.$$

**Exercise 2.59.** Show that if $(\mathcal{M}^2, g)$ is a Riemannian surface and $u : \mathcal{M} \to \mathbb{R}$ is a function, then

$$R(e^u g) = e^{-u} (R(g) - \Delta_g u),$$

where $\Delta_g$ is the Laplacian.

We use the method of moving frames to prove one of the most fundamental results in Riemannian geometry, the **Gauss–Bonnet formula**, which says that the integral of the Gauss curvature on a closed Riemannian surface $(\mathcal{M}^2, g)$ is equal to $2\pi$ times the Euler characteristic of $\mathcal{M}^2$.

**Theorem 2.60** (Gauss–Bonnet). If $(\mathcal{M}^2, g)$ is a closed oriented Riemannian surface, then

$$\int_{\mathcal{M}} K dA = 2\pi \cdot \chi(\mathcal{M}).$$

The proof of this formula will occupy the rest of the subsection. Let $\{e_1, e_2\}$ be a local positively-oriented orthonormal basis for $T\mathcal{M}$ in an open set $U \subset \mathcal{M}$ so that the **area form** is

$$dA = \omega^1 \wedge \omega^2.$$

Note that for any 2-form on a Riemannian surface,

$$\beta = 2\beta(e_1, e_2) \omega^1 \wedge \omega^2.$$

The Gauss-Bonnet integrand is locally the exterior derivative of the connection 1-form $-\omega^2$:

$$KdA = 2d\omega^1(e_1, e_2) \omega^1 \wedge \omega^2 = -d\omega^2.$$

(2.77)
We would like to apply Stokes’s theorem: for a compact surface $N$ with (possibly nonempty) boundary $\partial N$, for any 1-form $\alpha$,
\[
\int_N d\alpha = \int_{\partial N} \alpha.
\]
Since $\omega_1^2$ is only defined locally, we cannot apply Stokes’s theorem globally on $M$. However, we can choose an orthonormal frame field $\{e_1, e_2\}$ on $M$ minus a finite number of points. Indeed, let $V \in C^\infty(M)$ be a smooth vector field with isolated zeros at a finite number of points $p_1, \ldots, p_k$ (for example we may take $V$ to be the gradient of a Morse function on $M$).

**Remark 2.61.** A Morse function on $M$ is a $C^\infty$ function $f : M \to \mathbb{R}$ such that each critical point $p$ are nondegenerate, i.e., the matrix $\left( \frac{\partial^2 f}{\partial x^i \partial x^j}(p) \right)$ is nonsingular. In this case, the critical points are isolated (see Corollary 2.3 in [10]), and assuming $M$ is closed, there only finite many critical points. On any manifold there exists a Morse function (see Corollaries 6.7 and 6.8 in [10]).

On $M^2 - \{p_1, \ldots, p_k\}$ let
\[
e_1 = \frac{V}{|V|},
\]
and then $e_2$ is determined uniquely by the condition that $\{e_1, e_2\}$ be a positively-oriented orthonormal frame.

**Lemma 2.62.** For $\epsilon < \text{inj}(M^2, g)$, let $B(p_i, \epsilon) \doteq \{x \in M : d(x, p_i) < \epsilon\}$ for $i = 1, \ldots, k$. The Gauss–Bonnet integrand may be rewritten as
\[
\int_M KdA = \sum_{i=1}^k \lim_{\epsilon \to 0} \int_{\partial B(p_i, \epsilon)} \omega_1^2.
\]

**Proof.** By (2.77), Stokes’s theorem, and taking into account the orientation on $\bigcup_{i=1}^k \partial B(p_i, \epsilon)$ induced as the boundary of $M - \bigcup_{i=1}^k B(p_i, \epsilon)$, we have
\[
\int_M KdA = - \lim_{\epsilon \to 0} \int_M \bigcup_{i=1}^k B(p_i, \epsilon) d\omega_1^2 = \lim_{\epsilon \to 0} \int_{\bigcup_{i=1}^k \partial B(p_i, \epsilon)} \omega_1^2.
\]

We shall show that the boundary integral on the RHS of (2.78) tends to $2\pi$ times the index of the vector field $V$ (see the next section for the definition and properties of index).

---

3\text{inj}(M^2, g)$ is the injectivity radius of $g$. 
Lemma 2.63.

\[\lim_{\varepsilon \to 0} \int_{\partial B(p_i, \varepsilon)} \omega_1^2 = 2\pi \cdot \text{index}_{p_i}(V).\]

Assuming this lemma, we can complete the proof of the Gauss-Bonnet formula. Using (2.78), (2.79) and the Poincaré–Hopf theorem, which says that the index of a vector field is equal to the Euler characteristic of the underlying manifold, we have

\[\int_{\mathcal{M}} KdA = 2\pi \sum_{i=1}^{k} \text{index}_{p_i}(V) = 2\pi \cdot \chi_{\mathcal{M}} = 2\pi \cdot \chi(\mathcal{M}).\]

We conclude this section with the proof of Lemma 2.63. Given \(1 \leq i \leq k\), let \(\nu\) denote the unit outward normal vector field to \(\partial B(p_i, \varepsilon)\) and let \(T\) denote the unit tangent vector field to \(\partial B(p_i, \varepsilon)\) such that the frame \(\{\nu, T\}\) is positively-oriented. Parametrize \(\partial B(p_i, \varepsilon)\) by a unit speed path

\[\gamma_i : [0, L_i] \to \partial B(p_i, \varepsilon) \subset \mathcal{M}\]

such that \(d\gamma_i/ds = T\). We define the angle function

\[\theta : [0, L_i] \to \mathbb{R}\]

by

\[T(\gamma_i(s)) \parallel \cos \theta(s) \cdot e_1(\gamma_i(s)) + \sin \theta(s) \cdot e_2(\gamma_i(s)).\]

We require that \(\theta\) be smooth so that it is well-defined up to an integer multiple of \(2\pi\). The unit outward normal vector field \(\nu\) is given by

\[\nu(\gamma_i(s)) = \sin \theta(s) \cdot e_1(\gamma_i(s)) - \cos \theta(s) \cdot e_2(\gamma_i(s)).\]

The change in the angle \(\theta\) as you go around \(\gamma_i\) is

\[\theta(L_i) - \theta(0) = -2\pi (\text{index}_{p_i}(V) - 1).\]

The above discussion concerning the angle \(\theta\) is related to the boundary integral in (2.79) by the following.

Sublemma.

\[\omega_1^2(T) = \kappa - T\theta,\]

where \(\kappa \parallel - \langle \frac{\nabla}{\pi} T, \nu \rangle\) is the geodesic curvature of \(\partial B(p_i, \varepsilon)\).
14. Index of a vector field

In this section we discuss the index of a vector field on a 2-dimensional manifold. See [9] and the references therein for details.

Let \( ds \) denote the arc length element of \( \partial B(p_i, \varepsilon) \). By (2.80) and the fundamental theorem of calculus, we have

\[
\int_{\partial B(p_i, \varepsilon)} \omega_1^2 = \int_{\partial B(p_i, \varepsilon)} \omega_1^2(T) \, ds = \int_{\partial B(p_i, \varepsilon)} (\kappa - T\theta) \, ds = \int_{\partial B(p_i, \varepsilon)} \kappa \, ds - \theta(L_i) + \theta(0) = \int_{\partial B(p_i, \varepsilon)} \kappa \, ds + 2\pi (\text{index}_{p_i}(V) - 1).
\]

Since

\[
\lim_{\varepsilon \to 0} \int_{\partial B(p_i, \varepsilon)} \kappa \, ds = 2\pi,
\]

we conclude that

\[
\sum_{i=1}^{k} \lim_{\varepsilon \to 0} \int_{\partial B(p_i, \varepsilon)} \omega_1^2 = 2\pi \cdot \sum_{i=1}^{k} \text{index}_{p_i}(V),
\]

and both the lemma and the Gauss–Bonnet formula are proved.

14. Index of a vector field

In this section we discuss the index of a vector field on a 2-dimensional manifold. See [9] and the references therein for details.

We first consider a vector field \( V \) on an open subset \( U \) of \( \mathbb{R}^2 \) with an isolated 0 at a point \( p \in U \). Let

\[
S^1(p, \varepsilon) = \{ q \in \mathbb{R}^2 : |q - p| = \varepsilon \} \cong S^1
\]
(the isomorphism is given by \( q \mapsto q - p \) and assume that \( \epsilon > 0 \) is small enough so that \( V(q) \neq 0 \) for all \( q \in S^1(p, \epsilon) \). Consider the map 
\[ \bar{V} : S^1(p, \epsilon) \rightarrow S^1 \]
defined by
\[ \bar{V}(q) = \frac{V(q)}{|V(q)|}. \]
Then the \textbf{degree} (i.e., winding number) of \( \bar{V} \) is independent of \( \epsilon \) for \( \epsilon \) sufficiently small. The \textbf{index} of \( V \) at \( p \) is defined to be this degree and is denoted by \( \text{index}_p(V) \).

Here are some gradient vector fields with zeros at the origin:

\[
\nabla (x^2 + y^2) = (2x, 2y) \quad \text{has index } +1 \text{ at the origin } (0, 0)
\]

\[
\nabla (-x^2 - y^2) = (-2x, -2y) \quad \text{has index } +1 \text{ at } (0, 0)
\]

\[
\nabla (y^2 - x^2) = (-2x, 2y) \quad \text{has index } -1 \text{ at } (0, 0)
\]

Let \( V \) be a vector field on a 2-dimensional differentiable manifold with isolated zeros and let \( p \) be a zero of \( V \). Let \( x : U \rightarrow M \) be a parametrization with \( p \in x(U) \). The \textbf{index} of \( V \) at \( p \) is defined to be the index of the push forward \( d(x^{-1})(V|_{x(U)}) \) at \( x^{-1}(p) \). This definition is independent of the choice of parametrization.

Some nongradient vector fields with zeros at the origin:

\[
(-y, x) \quad \text{has index } +1 \text{ at the origin } (0, 0)
\]

\[
(x^2 - y^2, 2xy) \quad \text{has index } +2 \text{ at } (0, 0)
\]

Let \( f : \mathcal{M}^n \rightarrow \mathbb{R} \) be a Morse function, i.e., for each critical point (i.e., zero of the gradient \( \nabla f \) with respect to a Riemannian metric \( g \) on \( \mathcal{M} \)) \( p \) of \( f \), the symmetric matrix \( (\frac{\partial^2 f}{\partial x_i \partial x_j}(p)) \) is nonsingular. Then there are \( n \) real
eigenvalues, all which are nonzero. The **index of $f$ at $p$** is defined to be the number of negative eigenvalues of $(\frac{\partial f}{\partial x^i \partial x^j}(p))$, denoted by $\text{index}_p(f)$. Then

$$\text{index}_p(\nabla f) = (-1)^{\text{index}_p(f)}.$$  

Note that if $\det(\frac{\partial f}{\partial x^i \partial x^j}(p)) > 0$, then $(-1)^{\text{index}_p(f)} = 1$, whereas if $\det(\frac{\partial f}{\partial x^i \partial x^j}(p)) < 0$, then $(-1)^{\text{index}_p(f)} = -1$.

A main special case of the Poincaré–Hopf theorem says that for a Morse function $f$ on a closed manifold $M$ with critical points $p_1, p_2, \ldots, p_k$, the Euler characteristic of $M$ may be expressed as

$$\chi(M) = \sum_{i=1}^{k} \text{index}_{p_i}(\nabla f) = \sum_{i=1}^{k} (-1)^{\text{index}_{p_i}(f)}.$$

### 15. Stokes’s Theorem

Let $D \subset \mathbb{R}^2$ be a connected compact smooth domain, so that $\partial D$ is a smooth embedded circle. Given a smooth function $f : D \to \mathbb{R}$, define

$$\int_D f dx \wedge dy \div \int_D f dx dy,$$

where the left-hand side is the integral of a 2-form and the right-hand side is the Riemann integral of $f$. Let $\alpha = P dx + Q dy$ be a 1-form on $D$ (where $P$ and $Q$ are smooth functions on $D$). Then

$$d\alpha = dP \wedge dx + dQ \wedge dy = \frac{\partial P}{\partial y} dy \wedge dx + \frac{\partial Q}{\partial x} dx \wedge dy = \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx \wedge dy.$$  

Parametrize $\partial D$ counterclockwise by a map $c : S^1 \to \partial D$. The integral of $\alpha$ on $\partial D$ is given by

$$\int_{\partial D} \alpha = \int_{S^1} \alpha(c'(u)) du.$$  

Then Stokes’s Theorem, which says that

$$\int_D d\alpha = \int_{\partial D} \alpha,$$

is equivalent to

$$\int_D \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \int_{\partial D} P dx + Q dy.$$
16. Exercises on generalized geometry

Let $\mathcal{M}^n$ be a $C^\infty$ manifold. For each $p \in \mathcal{M}$, define an (indefinite) norm on $T_p\mathcal{M} \oplus T^*_p\mathcal{M}$ by

\[(2.81) \quad |X + \xi|^2 = \xi(X)\]

for $X \in T_p\mathcal{M}$ and $\xi \in T^*_p\mathcal{M}$.

**Exercise 2.64.** Show that the **parallelogram law** holds for this norm, i.e.,

\[(2.82) \quad |(X_1 + \xi_1) + (X_2 + \xi_2)|^2 + |(X_1 + \xi_1) - (X_2 + \xi_2)|^2 = 2|X_1 + \xi_1|^2 - 2|X_2 + \xi_2|^2\]

for $X_1, X_2 \in T_p\mathcal{M}$ and $\xi_1, \xi_2 \in T^*_p\mathcal{M}$.

By (2.82), **polarization** yields an (indefinite) inner product:

\[(2.83) \quad \langle X_1 + \xi_1, X_2 + \xi_2 \rangle \equiv \frac{1}{2}(\xi_1(X_2) + \xi_2(X_1)).\]

Let $\{e_i\}_{i=1}^n$ be a basis for $T_p\mathcal{M}$ and let $\{e^*_j\}_{j=1}^n$ be the dual basis for $T^*_p\mathcal{M}$, i.e.,

\[e^*_i(e_j) = \delta_{ij}.\]

**Exercise 2.65.** Define $f^+_i \equiv e_i + e^*_i$ and $f^-_i \equiv e_i - e^*_i$. Show that

\[(2.84) \quad \langle f^+_i, f^+_j \rangle = \pm \delta_{ij}, \quad \langle f^+_i, f^-_j \rangle = 0.\]

Given vector fields $X, Y$ and 1-forms $\xi, \eta$, the **Courant bracket** is defined by

\[(2.85) \quad [X + \xi, Y + \eta] \equiv [X, Y] + \mathcal{L}_X\eta - \mathcal{L}_Y\xi - \frac{1}{2}d(\langle \xi, Y \rangle - \langle \eta, X \rangle).\]

**Exercise 2.66.** Prove the following properties of the Courant bracket:

1. If $u = X + \xi$, $v = Y + \eta$ and $Z + \zeta$ are $C^\infty$ sections of $T\mathcal{M} \oplus T^*\mathcal{M}$, then

\[[u, v] + [(v, w), u] + [(w, u), v] = \frac{1}{3}d(\langle u, v \rangle, w) + \langle v, [u, w] \rangle + \langle w, [u, v] \rangle.\]

2. \[\begin{align*}
\{ f \} &\equiv f(u, v) - \langle u, v \rangle df, \\
X\langle v, w \rangle &\equiv \langle [u, v] + d\langle u, v \rangle, w \rangle + \langle v, [u, w] + d\langle u, w \rangle \rangle \\
&\text{for } f \in C^\infty(\mathcal{M}).
\end{align*}\]
Now let $g$ be a Riemannian metric on $M$. Given $p \in M$, define
\begin{equation}
V_p \doteq \{X + g(X) : X \in T_pM\} \subset T_pM \oplus T^*_pM, \\
V_p^\perp \doteq \{X - g(X) : X \in T_pM\} \subset T_pM \oplus T^*_pM,
\end{equation}
where $g(X) \in T^*_pM$ is given by (2.68). Let $X^+ \doteq X + g(X)$ and $X^- \doteq X - g(X)$. Then
\begin{equation}
\langle X^+, Y^+ \rangle = g(X, Y), \\
\langle X^-, Y^- \rangle = -g(X, Y).
\end{equation}

The projection $\pi_V : T_pM \oplus T^*_pM \to V_p$ is defined by
\begin{equation}
\pi_V (X + \xi) = \frac{X + g(\xi)}{2} + \frac{g(X) + \xi}{2}
\end{equation}
and the projection $\pi_{V^\perp} : T_pM \oplus T^*_pM \to V_p^\perp$ is defined by
\begin{equation}
\pi_{V^\perp} (X + \xi) = \frac{X - g(\xi)}{2} - \frac{g(X) - \xi}{2}.
\end{equation}
Then $\pi_V (X) = X^+$ and $\pi_{V^\perp} (X) = X^-.$

Let $V = \bigsqcup_{p \in M} V_p$, which is a (real) vector bundle over $M$. Given $X \in \Xi (M)$ and $v \in C^\infty (V)$, define
\begin{equation}
\nabla_X v \doteq \pi_V ([X^-, v]).
\end{equation}

**Exercise 2.67.** Show that $\nabla : \Xi (M) \times C^\infty (V) \to C^\infty (V)$ is an affine connection.

In a parametrization, each $V_p$ for $p \in x (U)$ is spanned by
\begin{align*}
\frac{\partial^+}{\partial x^i} &= \frac{\partial}{\partial x^i} + \sum_{j=1}^n g_{ij} dx^j = \frac{\partial}{\partial x^i} + g \left( \frac{\partial}{\partial x^i} \right) = \pi_V \left( \frac{\partial}{\partial x^i} \right), \\
\frac{\partial^-}{\partial x^i} &= \frac{\partial}{\partial x^i} - \sum_{j=1}^n g_{ij} dx^j = \frac{\partial}{\partial x^i} - g \left( \frac{\partial}{\partial x^i} \right) = \pi_{V^\perp} \left( \frac{\partial}{\partial x^i} \right).
\end{align*}

**Exercise 2.68.** Show that
\begin{equation}
\nabla_{\frac{\partial^+}{\partial x^j}} \frac{\partial^+}{\partial x^j} = 2 \sum_{k=1}^n \Gamma^k_{ij} \frac{\partial^+}{\partial x^k},
\end{equation}
where $\Gamma^k_{ij}$ are the Christoffel symbols (of the Levi-Civita connection).
Geodesics

The study of geodesics is one of the pillars of Riemannian geometry.

1. The first variation of arc length

Let $\gamma : [0, L] \to \mathcal{M}$ be a curve. We study ‘variations’ of this curve by considering a 1-parameter family of curves $\gamma_s : [0, L] \to \mathcal{M}$, where $s \in (-\varepsilon, \varepsilon)$, with $\gamma_0 = \gamma$. We say that the family $\{\gamma_s\}_{s \in (-\varepsilon, \varepsilon)}$ is $C^\infty$ if the map

$$\Gamma : [0, L] \times (-\varepsilon, \varepsilon) \to \mathcal{M}$$

defined by

$$\Gamma (s, t) = \gamma_s (t)$$

is $C^\infty$. We shall assume that both $\gamma$ and $\{\gamma_s\}_{s \in (-\varepsilon, \varepsilon)}$ are $C^\infty$. Note that the 1-parameter family of curves $\beta_t : s \mapsto \Gamma (s, t) = \gamma_s (t)$ is also $C^\infty$ (i.e., $\beta_t (s) \doteqdot \gamma_s (t)$).

Let $\frac{\partial}{\partial s} = e_1$ and $\frac{\partial}{\partial t} = e_2$ denote the coordinate vector fields on $[0, L] \times (-\varepsilon, \varepsilon) \subset \mathbb{R}^2$. Define the ‘$V$-ariation’ vector fields

$$d\Gamma_{(s, t)} \left( \frac{\partial}{\partial s} \right) \doteqdot \frac{\partial \Gamma}{\partial s} (s, t) \doteqdot V_s (t) \doteqdot V_t (s) = \frac{\partial \gamma_s}{\partial s} (t) = \beta_t' (s)$$

and the ‘$T$-angent’ vector fields

$$d\Gamma_{(s, t)} \left( \frac{\partial}{\partial t} \right) \doteqdot \frac{\partial \Gamma}{\partial t} (s, t) \doteqdot T_t (s) \doteqdot T_s (t) = \gamma_s' (t).$$

We may call $\Gamma$ a (parametrized) ‘surface’ (although $\Gamma$ may not be an immersion nor a regular surface) and we may call $\frac{\partial \Gamma}{\partial s}$ and $\frac{\partial \Gamma}{\partial t}$ vector fields along the surface $\Gamma$ (by analogy with vector fields along a curve).
Given \( s \in (-\varepsilon, \varepsilon) \), the vector fields \( T_s(t) = \gamma'_s(t) \) and \( V_s(t) = \beta'_s(0) \) along \( \gamma_s \) are called the tangent and variation vector fields, respectively. Also denote \( T(t) \parallel T_0(t) = \gamma_0'(t) \) and \( V(t) \parallel V_0(t) = \beta'_0(0) \), which are vector fields along \( \gamma \).

Now assume that \( \gamma = \gamma_0 \) has constant speed, that is, \( |\gamma'(t)| = |T_0(t)| = C \), where \( C \) is a constant. Consider the lengths of \( \gamma_s \)

\[
L(\gamma_s) = \int_0^L |\gamma'_s(t)| \, dt = \int_0^L |T_s(t)| \, dt = \int_0^L |T_t(s)| \, dt.
\]

We have

\[
\frac{d}{ds} L(\gamma_s) = \int_0^L \frac{\partial}{\partial s} |T_t(s)| \, dt.
\]

Now, by (2.36),

\[
\frac{\partial}{\partial s} |T_t(s)| = \frac{1}{2 |T_t(s)|} \frac{\partial}{\partial s} \left( |T_t(s)|^2 \right)
= \frac{1}{|T_t(s)|} \left( \left( \frac{D}{ds} T_t \right)(s), T_t(s) \right),
\]

where \( \frac{D}{ds} \) is covariant differentiation along \( \beta_t \).

**Claim:**

\[
(3.1) \quad \left( \frac{D}{ds} T_t \right)(s) = \left( \frac{D}{dt} V_s \right)(t),
\]

where \( \frac{D}{dt} \) is covariant differentiation along \( \gamma_s \).

Heuristically, we may believe this by writing

\[
\left( \frac{D}{ds} T_t \right)(s) - \left( \frac{D}{dt} V_s \right)(t) = \nabla_{d\Gamma(s)} \left( \frac{\partial}{\partial t} \right) - \nabla_{d\Gamma(s)} \frac{\partial}{\partial s} \frac{\partial}{\partial t}
= [d\Gamma(s), d\Gamma(t)]
= d\Gamma \left( \frac{\partial}{\partial s}, \frac{\partial}{\partial t} \right)
= 0.
\]

**Exercise 3.1.** Prove the claim.

**HINT:** Write \( T_t(s) \parallel T(t) = \sum_{i=1}^n T_i(s, t) \frac{\partial}{\partial x^i}, \quad V_s(t) \parallel V(t) = \sum_{j=1}^n V_j(s, t) \frac{\partial}{\partial x^j} \), and use the method of proof of Lemma 2.25.

Now, by the claim,

\[
\frac{\partial}{\partial s} |T_t(s)| = \frac{1}{|T_t(s)|} \left( \left( \frac{D}{dt} V_s \right)(t), T_s(t) \right)
= \frac{1}{|T_s(t)|} \left( \frac{d}{dt} \left( V_s(t), T_s(t) \right) - \left( V_s(t), \left( \frac{D}{dt} T_s \right)(t) \right) \right).
\]
Evaluating this at \( s = 0 \) while using \(|T_s(t)| \equiv 1\), we have
\[
\frac{\partial}{\partial s} \bigg|_{s=0} |T_t(s)| = \frac{1}{|T(t)|} \left( \frac{d}{dt} \langle V(t), T(t) \rangle - \langle V(t), \left( \frac{D}{dt} T \right)(t) \rangle \right).
\]

Hence, by the Fundamental Theorem of Calculus:

**Proposition 3.2** (First variation of arc length formula). If \( \gamma = \gamma_0 \) has constant speed, then
\[
\frac{d}{ds} \bigg|_{s=0} L(\gamma_s) = \langle V(L), T(L) \rangle - \langle V(0), T(0) \rangle - \int_0^L \frac{1}{|T(t)|} \langle V(t), \left( \frac{D}{dt} T \right)(t) \rangle dt.
\]

As a special case we may assume that the \( \gamma_s \) satisfy \( \gamma_s(0) = \gamma(0) \) and \( \gamma_s(L) = \gamma(L) \) for all \( s \in (-\varepsilon, \varepsilon) \). Then \( V(0) = 0 \) and \( V(L) = 0 \). Hence, if the two endpoints of each \( \{\gamma_s\} \) are independent of \( s \), then
\[
\frac{d}{ds} \bigg|_{s=0} L(\gamma_s) = -\int_0^L \frac{1}{|T(t)|} \langle V(t), \left( \frac{D}{dt} T \right)(t) \rangle dt.
\]

Given a curve \( \gamma \), if we have that \( \frac{d}{ds} \bigg|_{s=0} L(\gamma_s) = 0 \) for any family of curves \( \gamma_s \) with \( \gamma_0 = \gamma \), then
\[
-\int_0^L \frac{1}{|T(t)|} \langle V(t), \left( \frac{D}{dt} T \right)(t) \rangle dt = 0
\]
for any \( C^\infty \) vector field \( V \) along \( \gamma \). Hence \( \gamma \) satisfies
\[
\left( \frac{D}{dt} T \right)(t) = 0
\]
for all \( t \in [0, L] \). This equation for \( \gamma \) is called the geodesic equation. One often writes this as \( \nabla_T T = 0 \).

As another special case of (3.2), we have:

**Corollary 3.3.** If \( \gamma = \gamma_0 \) is a geodesic, then
\[
\frac{d}{ds} \bigg|_{s=0} L(\gamma_s) = \langle V(L), T(L) \rangle - \langle V(0), T(0) \rangle,
\]
where \( V(0) = \frac{\partial}{\partial s} \gamma(0) \) and \( V(L) = \frac{\partial}{\partial s} \gamma(L) \).

### 1. Vector fields along a map.

Let \( \varphi : \mathcal{N} \to \mathcal{M} \) be a \( C^\infty \) map. A vector field along \( \varphi \) is a map
\[
V : \mathcal{N} \to \mathcal{T}\mathcal{M}
\]
such that \( \pi \circ V = \varphi \). That is, \( V(x) \in T_{\varphi(x)}\mathcal{M} \) for \( x \in \mathcal{N} \). Let \( \Xi(\varphi) \) denote the set of \( C^\infty \) vector fields along \( \varphi \).
3. Geodesics

Generalizing Lemma 2.25, we have:

**Claim.** There exists a unique map $\tilde{\nabla} : \mathfrak{X}(\mathcal{N}) \times \mathfrak{X}(\varphi) \to \mathfrak{X}(\varphi)$, where we denote $\tilde{\nabla}(U,V) = \tilde{\nabla}_{U}V$, with the following properties: for any $U \in \mathfrak{X}(\mathcal{N})$ and $V,W \in \mathfrak{X}(\varphi)$ and $C^\infty$ function $f : \mathcal{N} \to \mathbb{R}$,

1. For each $y \in \mathcal{N}$, the tangent vector $\tilde{\nabla}_{U}yV \in \tilde{\nabla}_{U}V(y) \in T_{\varphi(y)}\mathcal{M}$ depends only on $U,y$, and $\varphi$.
2. $\tilde{\nabla}_{U}(V + W) = \tilde{\nabla}_{U}V + \tilde{\nabla}_{U}W$,
3. $\tilde{\nabla}_{U}(fV) = f\tilde{\nabla}_{U}V + U(f)V$,
4. If $\tilde{V} \in \mathfrak{X}(\mathcal{M})$ is such that $\tilde{V} \circ \varphi = V$, then $\tilde{\nabla}_{U}V = \nabla dtV$.

**Sketch of proof of claim.** Given $y \in \mathcal{N}$, let $V$ be a neighborhood of $\varphi(y)$ in $\mathcal{M}$ and let $\{E_{i}\}_{i=1}^{n}$ be a set of $C^\infty$ vector fields on $V$ such that for each point $p \in V$, $\{(E_{i})_{p}\}_{i=1}^{n}$ is a basis of $T_{p}\mathcal{M}$. For example, let $(U,x)$ be a parametrization of $\mathcal{M}$ such that $\varphi(y) \in x(U)$ and take $V = x(U)$ and $E_{i} = \frac{\partial}{\partial x_{i}}$.

For $V \in \mathfrak{X}(\varphi)$ we may write $V$ in $V$ as

$$V(y) = \sum_{i=1}^{n} V^{i}(y)(E_{i})_{\varphi(y)}.$$  

For $U,y \in T_{y}\mathcal{N}$, we define

$$\tilde{\nabla}_{U}yV = \sum_{i=1}^{n} (U_{y}(V^{i})(E_{i})_{\varphi(y)}) + V^{i}(y)\nabla_{d\varphi(U)}(E_{i})_{\varphi(y)}.$$  

**Exercise 3.4.** Show that definition (3.5) is independent of the choice of $\{E_{i}\}_{i=1}^{n}$ as above.

**Exercise 3.5.**

1. Show that $\tilde{\nabla}$ satisfies properties (1)–(4).
2. Show that if $\mathcal{N} = \mathcal{I}$ is an interval and $e_{1} = \frac{d}{dt} \in T\mathcal{I}$, then

$$\tilde{\nabla}_{e_{1}}V = \frac{\nabla dt}{dt}V,$$

where $\frac{\nabla dt}{dt}$ is as Lemma 2.25, corresponding to the Levi-Civita connection.
3. Show that the definition of $\tilde{\nabla}$ generalizes the definition of $\frac{d}{dt}$ in (2.31).
4. Using $\mathcal{N} = [0,L] \times (-\varepsilon,\varepsilon)$ and $\tilde{\nabla}$, interpret $\frac{d}{dt}T_{t}$ and $\frac{d}{dt}V_{s}$ in (3.1).

Let $\varphi : \mathcal{N} \to \mathcal{M}$ be a $C^\infty$ map. If $U,V$ are vector fields on $\mathcal{N}$, then $d\varphi(U)$ and $d\varphi(V)$ are vector fields along $\varphi$. 


1. The first variation of arc length

Claim.

(3.6) \( \tilde{d}_\varphi ([U,V]) = \tilde{\nabla}_{d\varphi(U)}d\varphi (V) - \tilde{\nabla}_{d\varphi(V)}d\varphi (U). \)

In particular, if \([U,V] = 0\), then \(\tilde{\nabla}_{d\varphi(U)}d\varphi (V) = \tilde{\nabla}_{d\varphi(V)}d\varphi (U).\)

Proof. This follows from

\[
\tilde{\nabla}_{d\varphi(U)}d\varphi (V) = d\varphi (\tilde{\nabla}_V^N V)
\]

and \(\tilde{\nabla}_U^N V - \tilde{\nabla}_U^N U = [U,V].\)

Claim. If \(U \in \Xi (\mathcal{N})\) and \(V, W \in \Xi (\varphi),\) then

(3.7) \( U \langle V, W \rangle = \langle \tilde{\nabla}_U V, W \rangle + \langle V, \tilde{\nabla}_U W \rangle. \)

Note that \(\langle V, W \rangle : \mathcal{N} \to \mathbb{R}.\)

Proof. For convenience, we use the Einstein summation convention.

First we compute that

(3.8)

\[
U_y \langle V, W \rangle = U_y \left( V^i (y) W^j (y) \langle E_i, E_j \rangle_{\varphi(y)} \right)
\]

\[
= U_y (V^i) W^j (y) \langle E_i, E_j \rangle_{\varphi(y)} + U_y (W^j) V^i (y) \langle E_i, E_j \rangle_{\varphi(y)}
\]

\[
+ V^i (y) W^j (y) U_y \langle E_i, E_j \rangle_{\varphi(y)}
\]

\[
= U_y (V^i) \langle E_i, W \rangle_{\varphi(y)} + U_y (W^j) \langle V, E_j \rangle_{\varphi(y)}
\]

\[
+ V^i (y) W^j (y) U_y \langle E_i, E_j \rangle_{\varphi(y)}. \]

Using (3.5), we compute that

(3.9)

\[
\left\langle \tilde{\nabla}_{U_y} V, W \right\rangle + \left\langle V, \tilde{\nabla}_{U_y} W \right\rangle
\]

\[
= \left\langle U_y (V^i) (E_i)_{\varphi(y)} + V^i (y) \nabla_{d\varphi(U_y)} (E_i)_{\varphi(y)}, W_{\varphi(y)} \right\rangle
\]

\[
+ \left\langle V_{\varphi(y)}, U_y (W^j) (E_i)_{\varphi(y)} + W^j (y) \nabla_{d\varphi(U_y)} (E_i)_{\varphi(y)} \right\rangle
\]

\[
= U_y (V^i) \langle E_i, W \rangle_{\varphi(y)} + U_y (W^j) \langle V, E_j \rangle_{\varphi(y)}
\]

\[
+ V^i (y) W^j (y) \left\langle \nabla_{d\varphi(U_y)} (E_i)_{\varphi(y)}, (E_j)_{\varphi(y)} \right\rangle
\]

\[
+ V^i (y) W^j (y) \left\langle (E_i)_{\varphi(y)}, \nabla_{d\varphi(U_y)} (E_i)_{\varphi(y)} \right\rangle. \]

That (3.8) and (3.9) are equal now follows from

\[
U_y \langle E_i, E_j \rangle_{\varphi(y)} = \left\langle \nabla_{d\varphi(U_y)} (E_i)_{\varphi(y)}, (E_j)_{\varphi(y)} \right\rangle + \left\langle (E_i)_{\varphi(y)}, \nabla_{d\varphi(U_y)} (E_i)_{\varphi(y)} \right\rangle,
\]

which in turn follows from the compatibility of \(\nabla\) with the metric \(g\) on \(\mathcal{M}.

Remark 3.6. A vector field along \(\varphi\) is equivalent to a section of the pulled back tangent bundle \(\varphi^*T\mathcal{M}\) in (2.59), i.e., a map \(V : \mathcal{N} \to \varphi^*T\mathcal{M}\) such that \(\pi \circ V = \text{id}_\mathcal{N},\) where \(\pi : \varphi^*T\mathcal{M} \to \mathcal{N}\) is the projection map. Let \(C^\infty (\varphi^*T\mathcal{M})\) denote the set of \(C^\infty\) sections of \(\varphi^*T\mathcal{M}.\) Then \(\tilde{\nabla}\) in (3.5) is equivalent to a
map \( \tilde{\nabla} : \Xi(M) \times C^\infty(\varphi^*TM) \to C^\infty(\varphi^*TM) \), which is called the induced (or pulled back) connection.

2. Geodesics and the exponential map

Geodesics provide a way to define a map, called the exponential map, from each tangent space to the manifold. The behavior of geodesics is related to the behavior of this map.

2.1. Exponential map, the Gauss lemma and the Hopf–Rinow theorem.

Let \((M^n, g)\) be a Riemannian manifold and let \(p \in M\). We now define the exponential map at \(p\). Given \(V \in T_pM\), let \(\varepsilon_V \in (0, \infty]\) denote the largest positive extended number such that the constant speed geodesic \(\gamma_V(s)\) emanating from \(p\) with \(\gamma_V'(0) = V\) is defined for \(s \in [0, \varepsilon_V]\). Note that \(|\gamma_V'(s)| \equiv |V|\).

Define
\[
\text{domain}(\exp_p) \doteq \{V \in T_pM : \varepsilon_V > 1\}.
\]

The exponential map at \(p\)
\[
\exp_p : \text{domain}(\exp_p) \subset T_pM \to M
\]
is defined by
\[
\exp_p(V) \doteq \gamma_V(1).
\]

Note that \(\gamma_V(r) = \gamma_{rV}(1) = \exp_p(rV)\) for \(V \in T_pM\) and \(r < \varepsilon_V\) \((\varepsilon_{rV} = \frac{1}{r} \varepsilon_V > 1)\). In particular, \(\exp_p(V) = \gamma_{\frac{V}{|V|}}(|V|)\). That is, the exponential map takes a vector \(V\) to the image of \(|V|\) for the unit speed geodesic starting in the direction of \(V\).

Exercise 3.7. Show that if \(p \in M\) and \(V \in T_pM\) are such that \(\varepsilon_V < \infty\), then for any \(s_i \to \varepsilon_V\), \(\lim_{i \to \infty} \gamma_V(s_i)\) does not exist.

Hint. Suppose that there exists \(s_i \to \varepsilon_V\) such that \(\lim_{i \to \infty} \gamma_V(s_i) \doteq q\) exists. Then there exists a subsequence and \(W \in T_qM\) such that \(\gamma_{\frac{W}{|W|}}(s_i) \to W\) and \(|W| = |V|\). Hence there exist \(\delta \in (0, \varepsilon_V)\) and a constant speed geodesic \(\beta : (-\delta, \delta) \to M\) with \(\beta(0) = q\) and \(\beta'(0) = W\). We claim that \(\beta(s) = \gamma_V(s + \varepsilon_V)\) for \(s \in (-\varepsilon_V, 0)\), which means that we may extend \(\gamma_V(s)\) past \(s = \varepsilon_V\), yielding a contradiction.

Claim. \(\text{domain}(\exp_p)\) is an open subset of \(T_pM\).

Define \(\text{domain}(\exp) \doteq \bigsqcup_{p \in M} \text{domain}(\exp_p)\) and
\[
\exp : \text{domain}(\exp) \subset TM \to M
\]
by \(\exp(p, V) \doteq \gamma_V(1)\).
2. Geodesics and the exponential map

Claim. The exponential map is $C^\infty$.

Proof. This follows from the $C^\infty$ dependence on their initial data of solutions to systems of first order ODE with $C^\infty$ coefficients.

Given $V \in \text{domain}(\exp_p)$, let

$$(d \exp_p)_V : T_V (T_p \mathcal{M}) \cong T_p \mathcal{M} \to T_{\exp_p(V)} \mathcal{M}$$

denote the differential of $\exp_p$ at $V$.

Claim. The map $(d \exp_p)_0 : T_0 (T_p \mathcal{M}) \cong T_p \mathcal{M} \to T_p \mathcal{M}$ is the identity map.

Proof. We compute that

$$(d \exp_p)_0 (V) = \left. \frac{d}{dt} \right|_{t=0} \exp_p (tV) = \left. \frac{d}{dt} \right|_{t=0} \gamma_V (t) = V.$$

By the inverse function theorem (Proposition 1.18), there exists $\varepsilon > 0$ such that $\exp_p|_{B(0,\varepsilon)} : B(0,\varepsilon) \to \exp_p (B(0,\varepsilon))$ is a diffeomorphism onto an open subset of $\mathcal{M}$.

Lemma 3.8 (Gauss). If $V \in \text{domain}(\exp_p)$ and $W \in T_V (T_p \mathcal{M}) \cong T_p \mathcal{M}$ is perpendicular to $V$, then $(d \exp_p)_V (W) \in T_{\exp_p(V)} \mathcal{M}$ is perpendicular to $(d \exp_p)_V (V) = \gamma'_V (1)$. That is, if $\langle W, V \rangle = 0$, then

$$\langle (d \exp_p)_V (W), (d \exp_p)_V (V) \rangle = 0.$$

Proof. Given $V \in \text{domain}(\exp_p)$ and $W \in T_p \mathcal{M}$, there exists $\delta > 0$ and a family of constant speed geodesics $\gamma_s : [0,1] \to \mathcal{M}, s \in (-\delta, \delta)$, defined by

$$\gamma_s(t) = \exp_p(t(V + sW)) \text{ for } 0 \leq t \leq 1.$$

Since each geodesic $\gamma_s$ is minimal, we have $L(\gamma_s) = |V + sW|_{g(p)}$. From this, we compute that

$$\left. \frac{d}{ds} \right|_{s=0} L(\gamma_s) = \left. \frac{1}{|V|} \frac{d}{ds} \right|_{s=0} |V + sW|_{g(p)}^2 = \frac{1}{|V|} \langle V, W \rangle_{g(p)}.$$

On the other hand, by the first variation of arc length formula (3.4), we have

$$\left. \frac{d}{ds} \right|_{s=0} L(\gamma_s) = \frac{1}{|V|} \left\langle \left. \frac{\partial \gamma_s}{\partial s} \right|_{s=0} (1) , \gamma'_0 (1) \right\rangle_{g(\exp_p(V))}$$

$$= \frac{1}{|V|} \langle (d \exp_p)_V (W), (d \exp_p)_V (V) \rangle$$
since \( V(0) = 0 \) and where the factor on the RHS of the first line is due to 
\(|\gamma'(1)| = |V|\). Hence
\[
\langle d\exp_p(W), d\exp_p(V) \rangle = \langle V, W \rangle_{g(p)}.
\]

\[\square\]

**Corollary 3.9.** If the distance function \( r(\cdot) \doteq d(\cdot, p) \) is \( C^\infty \) in a neighborhood of a point \( x \), then we have
\[
(3.10) \quad (\nabla r)(x) = \gamma_x'(r(x))
\]
where \( \gamma_x : [0, r(x)] \to \mathcal{M} \) is the unique minimal unit speed geodesic from \( p \) to \( x \). Thus, if \( \gamma_x = \gamma_{V_x} \) for some unit vector \( V_x \), then
\[
(\nabla r)(x) = (d\exp_p)_{r(x)}(V_x).
\]

**Proof.** Let \( \gamma_s : [0, r(x)] \to \mathcal{M}, s \in (-\delta, \delta) \), be any 1-parameter family of curves with \( \gamma_0 = \gamma_x \) and \( \gamma_s(0) = p \). Since \( \gamma_x \) is a minimal geodesic, we have
\[
L(\gamma_0) = d(p, \gamma_0(b)) \quad \text{while} \quad L(\gamma_s) \geq d(p, \gamma_s(r(x))).
\]
Hence, since \( \nabla r \) exists at \( x \),
\[
\left. \frac{d}{ds}\right|_{s=0} L(\gamma_s) = \langle \nabla r, X \rangle,
\]
where \( X \doteq \partial |_{s=0} \gamma_s(r(x)) \in T_x \mathcal{M} \). On the other hand, by the first variation formula (3.4),
\[
\left. \frac{d}{ds}\right|_{s=0} L(\gamma_s) = \left\langle \frac{\partial \gamma_x}{\partial t}(r(x)), X \right\rangle.
\]
Therefore
\[
\langle \nabla r, X \rangle = \left\langle \frac{\partial \gamma_x}{\partial t}(r(x)), X \right\rangle \quad \text{for all} \quad X \in T_x \mathcal{M}, \text{and} (3.10) \text{ follows.}\]
\[\square\]

Let \( \frac{\partial}{\partial r} \) denote both the radial unit outward pointing vector field on \( T_p \mathcal{M} - \{\vec{0}\} \) and its image under the derivative of the exponential map: \( \frac{\partial}{\partial r} \doteq d\exp_p(\frac{\partial}{\partial r}) \). For the latter to be well defined, we restrict \( \exp_p \) to a punctured ball \( B(\vec{0}, \varepsilon) - \{\vec{0}\} \subset T_p \mathcal{M}, \varepsilon > 0 \), where it is an embedding (since at \( \vec{0} \) the map \( d\exp_p = \text{id}_{T_p \mathcal{M}} \) is invertible, there exists such an \( \varepsilon \)). We also denote
\[
r(x) \doteq |\exp_p^{-1}(x)| \quad \text{for} \quad x \in \tilde{B}(p, \varepsilon) \doteq \exp_p(B(\vec{0}, \varepsilon)).
\]
Note that we have not yet shown that \( r(x) \) equals \( d(x, p) \) for \( x \in \tilde{B}(p, \varepsilon) \).

The Gauss lemma implies that at any point \( x \in \tilde{B}(p, \varepsilon) - \{p\} \), we have
\[
(3.11) \quad \frac{\partial}{\partial r} = \text{grad} \, r;
\]
that is, for every \( X \in T_x \mathcal{M}, \left\langle \partial/\partial r, X \right\rangle = X(r) \). Indeed, one sees this from writing \( X \) as the sum of its radial component and perpendicular vector and
applying the Gauss lemma: if
\[ X = a \frac{\partial}{\partial r} + \sum_{i=1}^{n-1} b^i \frac{\partial}{\partial \theta^i}, \]
where \( \{r, \theta^1, \ldots, \theta^{n-1}\} \) are spherical coordinates, then
\[ \left\langle \frac{\partial}{\partial r}, X \right\rangle = a = X(r), \]
where the first equality follows from the Gauss lemma: \( \left\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial \theta^i} \right\rangle = 0 \).

Lemma 3.10.

1. For every \( V \in B(\vec{0}, \varepsilon) \), \( \gamma_V : [0, 1] \to \mathcal{M} \) is the unique path, up to reparametrization, joining \( p \) and \( \gamma_V(1) = \exp_p(V) \) whose length realizes the distance \( d(p, \exp_p(V)) = |V| \). In particular, short geodesics are minimal and
\[ r(x) = d(x, p) \quad \text{for} \quad x \in \tilde{B}(p, \varepsilon). \]

2. For every \( q \) outside of the ball \( \tilde{B}(p, \varepsilon) \), there exists a point \( q' \in \partial \tilde{B}(p, \varepsilon) \) such that
\[ d(p, q) = d(p, q') + d(q', q) = \varepsilon + d(q', q). \]

Proof. (1) This follows from the Gauss lemma, in particular on integrating the fact that for any path \( \beta \) we have
\[ \left| \dot{\beta}(u) \right| \geq \left\langle \dot{\beta}, \frac{\partial}{\partial r} \right\rangle = \frac{d}{d u} r(\beta(u)) \]
as long as \( \beta \) stays inside \( \tilde{B}(p, \varepsilon) \).

Exercise 3.11. Prove part (2) of Lemma 3.10.

Part (2) is useful in proving the Hopf–Rinow theorem below. It also implies that for all \( q \notin \tilde{B}(p, \varepsilon) \), we have \( d(q, p) \geq \varepsilon \). On the other hand, if \( q \in \tilde{B}(p, \varepsilon) \), then there exists \( V \in B(\vec{0}, \varepsilon) \) such that \( \exp_p(V) = q \) so that \( d(q, p) = |V| < \varepsilon \). Hence
\[ \tilde{B}(p, \varepsilon) = B(p, \varepsilon) \supseteq \{x \in \mathcal{M} : d(x, p) < \varepsilon \}. \]

Remark 3.12. If there exists a (unit speed) minimal geodesic \( \gamma \) from \( p \) to \( q \), then it is easy to show that there exists \( q' \in \partial \tilde{B}(p, \varepsilon) \) satisfying (3.12); namely we define \( q' \equiv \gamma(\varepsilon) \). Part of the Hopf-Rinow theorem is to prove the converse of this under the assumption \( (\mathcal{M}, d) \), where \( d \) is the Riemannian distance, is complete as a metric space.

Now we recall the following well-known result.
Theorem 3.13 (Hopf–Rinow). Let \((\mathcal{M}^n, g)\) be a Riemannian manifold. Then the following are equivalent:

1. \((\mathcal{M}, d)\) is a complete metric space.
2. There exists \(p \in \mathcal{M}\) such that \(\text{exp}_p\) is defined on all of \(T_p \mathcal{M}\).
3. For all \(p \in \mathcal{M}\), \(\text{exp}_p\) is defined on all of \(T_p \mathcal{M}\).

Any one of these conditions implies

4. For any \(p, q \in \mathcal{M}\) there exists a smooth minimal geodesic from \(p\) to \(q\).

**Proof.** Clearly (3) implies (2). One part of the statement of this theorem is that given \(p \in \mathcal{M}\), if \(\text{exp}_p\) is defined on all of \(T_p \mathcal{M}\), then for any \(q \in \mathcal{M}\) there exists a smooth minimal geodesic from \(p\) to \(q\). To see this, we may assume \(q \notin B(p, \varepsilon)\) so that there exists \(q' \in \partial B(p, \varepsilon)\) satisfying (3.12), and we consider the unit speed geodesic \(\gamma : [0, \infty) \to \mathcal{M}\) with \(\gamma(0) = p\) and \(\gamma(\varepsilon) = q'\). Take \(u_0 \in [0, d(p, q)]\) to be the largest value of \(u\) such that

\[
(3.13) \quad d(p, q) = d(p, \gamma(u)) + d(\gamma(u), q)
\]

for all \(v \in [0, u]\). Note that the set of all such \(u\) is closed. Since

\[
(3.14) \quad d(p, q) = d(p, \gamma(u)) + d(\gamma(u), q) \quad \text{for } u \in [0, \varepsilon],
\]

we have \(u_0 \geq \varepsilon\). If \(u_0 < d(p, q)\), then there exists \(r_0 \in (0, d(\gamma(u_0), q))\) such that \(\text{exp}_{\gamma(u_0)}\) is an embedding when restricted to \(B(\hat{0}, r_0) \subset T_{\gamma(u_0)} \mathcal{M}\). This gives us a point \(q'' \in \partial B(\gamma(u_0), r_0)\) such that

\[
(3.15) \quad d(\gamma(u_0), q) = d(\gamma(u_0), q'') + d(q'', q).
\]

Now (3.13), (3.15) and the triangle inequality imply

\[
d(p, q'') = d(p, \gamma(u_0)) + d(\gamma(u_0), q').
\]

Let \(\sigma : [0, d(\gamma(u_0), q'')] \to \mathcal{M}\) be the unique unit speed minimal geodesic from \(\gamma(u_0)\) to \(q''\). Therefore the concatenated path \(\gamma|_{[0, u_0]} \sim \sim \sigma\) from \(p\) to \(q''\) is minimal, and by the first variation of arc length formula, \(\gamma|_{[0, u_0]} \sim \sim \sigma = \gamma|_{[0, d(p, q'')]\)}\) is a smooth geodesic. Hence

\[
d(p, q) = d(p, \gamma(d(p, q''))) + d(\gamma(d(p, q'')), q),
\]

which is a contradiction. Hence \(u_0 = d(p, q)\) so that \(\gamma|_{[0, d(p, q)]}\) is a minimal geodesic from \(p\) to \(q\). Note that allowing \(p \in \mathcal{M}\) to be arbitrary, the above argument tells us (3) implies (4).

To show that completeness of the metric space \((\mathcal{M}, d)\) follows from \(\text{exp}_p\) being defined on all of \(T_p \mathcal{M}\) for some \(p\), we note that the result in the above paragraph implies that corresponding to a sequence of points \(\{q_i\}\) in \(\mathcal{M}\) is a sequence of unit speed minimal geodesics \(\gamma_i\) from \(p\) to \(q_i\). If
\{q_i\} is Cauchy, then there exists a subsequence such that $\gamma_i$ converges to a geodesic $\gamma_\infty$ (since $\{q_i\}$ is Cauchy, the lengths $L_i$ of the intervals where $\gamma_i$ are defined converge; and since the unit $(n-1)$-sphere is compact, there exists a subsequence such that $\{\dot{\gamma}_i(0)\}$ converges). The continuous dependence of solutions on their initial data then implies $q_i$ limits to the endpoint of $\gamma_\infty$. This completes the proof that (2) implies (1).

Finally we show that completeness of the metric space $(\mathcal{M},d)$ implies $\exp_p$ is defined on all of $T_p\mathcal{M}$ for all $p$. Given $V \in T_p\mathcal{M}$, let $t_0 \in (0,\infty]$ be the largest value of $t$ such that $\gamma$ is defined on $[0,t)$. If $t_0 < \infty$, then we may define $\gamma(t_0) = \lim_{t \to t_0^-} \gamma(t)$ for any sequence $t_i \to t_0$. Since $(\mathcal{M},d)$ is complete, and the sequence $\{\gamma(t_i)\}$ is Cauchy, the limit exists. It is then easy to show that $\gamma$ may be smoothly extended past $t_0$ as a geodesic, contradicting $t_0 < \infty$.

Thus (1) implies (3). □


Definition 3.15. A point $x \in \mathcal{M}$ is a conjugate point of $p \in \mathcal{M}$ if $x$ is a singular value of $\exp_p : T_p\mathcal{M} \to \mathcal{M}$. That is, there exists $V \in T_p\mathcal{M}$ such that $x = \exp_p(V)$ and $d(\exp_p)_V : TV(T_p\mathcal{M}) \to T_x\mathcal{M}$ is not injective.

A fact which we will prove later is the following. If $p \in \mathcal{M}$ and $x = \exp_p(V)$, where $d(\exp_p)_V$ is not injective, then
$$d(p,\gamma_V(s)) < s |V|$$
for each $s > 1$ in the domain of $\gamma_V$. In particular, a constant speed geodesic emanating from $p$ is not minimizing after its first conjugate point (it may stop minimizing earlier; see the definition of cut point below).

Exercise 3.16. Describe the exponential map on the unit $n$-sphere.

Solution. Let $p \in S^n \equiv S^n(1)$ and let $V \in T_pS^n - \{0\}$. Then
$$\gamma_V(s) = \cos(|V|s)p + \sin(|V|s)\frac{V}{|V|}.$$
So
$$\exp_p(V) = \cos(|V|)p + \sin(|V|)\frac{V}{|V|}.$$

Exercise 3.17. Compute $d(\exp_p)_V : TV(T_pS^n) \to TV(\exp_p(V)S^n)$. When is $d(\exp_p)_V$ not injective?

Solution. Since $d(\exp_p)_0 = id_{T_pS^n}$, we assume $V \neq 0$. We compute
$$d(\exp_p)_V(W) = -\sin(|V|)\frac{\langle W, V \rangle}{|V|}p + \cos(|V|)\frac{\langle W, V \rangle}{|V|}\frac{V}{|V|}$$
$$+ \sin(|V|)\frac{W}{|V|} - \sin(|V|)\frac{\langle W, V \rangle}{|V|^3}V.$$
using $W(|V|) = \frac{\langle W, V \rangle}{|V|}$. Let $V^\perp = \{ U \in T_p S^n : \langle U, V \rangle = 0 \}$ and define the projection

$$\pi : T_p S^n \to V^\perp \subset T_p S^n$$

by

$$\pi(W) = W - \left\langle W, \frac{V}{|V|} \right\rangle \frac{V}{|V|}.$$

Then

$$d(\exp_p)_V(W) = \left\langle W, \frac{V}{|V|} \right\rangle \left( \cos (|V|) \frac{V}{|V|} - \sin (|V|) p \right) + \sin (|V|) \pi(W).$$

Note that

$$\left\langle \cos (|V|) \frac{V}{|V|} - \sin (|V|) p, \pi(W) \right\rangle = 0.$$

Hence

$$|d(\exp_p)_V(W)|^2 = \left\langle W, \frac{V}{|V|} \right\rangle^2 \left| \cos (|V|) \frac{V}{|V|} - \sin (|V|) p \right|^2 + \sin^2 (|V|)|\pi(W)|^2.$$

Hence $d(\exp_p)_V$ is not injective if and only if $V \neq 0$ and $\sin (|V|) = 0$, i.e., $|V| \in \pi \mathbb{Z}$. In this case,

$$\ker (d(\exp_p)_V) = V^\perp \subset T_p S^n.$$

2.2. Cut locus and injectivity radius.

Let $(M^n, g)$ be a complete Riemannian manifold.

**Definition 3.18.** A function $f : M \to \mathbb{R}$ is a (globally) **Lipschitz function** with Lipschitz constant equal to $L$ if for all $x, y \in M$ we have

$$|f(x) - f(y)| \leq L d(x, y).$$

Given $p \in M$, from the triangle inequality it is easy to see that distance function $r(\cdot) \equiv d(\cdot, p)$ is a Lipschitz function with Lipschitz constant equal to 1, i.e.,

$$|r(x) - r(y)| \leq d(x, y).$$

Given a unit speed geodesic $\gamma : [0, \infty) \to M$ with $\gamma(0) = p$, either

1. $\gamma$ is a geodesic ray, i.e., minimal on each finite subinterval, or
2. there exists a unique $r_\gamma \in (0, \infty)$ such that $d(\gamma(s), p) = s$ for $s \leq r_\gamma$ and $d(\gamma(s), p) < s$ for $s > r_\gamma$. 

In other words, $\gamma(s)$ stops minimizing as $s$ surpasses $r_\gamma$. We say that $\gamma(r_\gamma)$ is a **cut point** to $p$ along $\gamma$. Note that if $\gamma(s)$ is a conjugate point to $p$ along $\gamma$, then $s \geq r_\gamma$. The **cut locus** $\text{Cut}(p)$ of $p$ in $\mathcal{M}$ is the set of all cut points of $p$.

Now let

$$D_p \triangleq \{ V \in T_p \mathcal{M} : d(\exp_p(V), p) = |V| \},$$

which is a closed subset of $T_p \mathcal{M}$. We define $C_p \triangleq \partial D_p$ to be the **cut locus of $p$ in the tangent space**. We have $\text{Cut}(p) = \exp_p(C_p)$. We have

$$\exp_p|_{\text{int}(D_p)} : \text{int}(D_p) \to \mathcal{M} - \text{Cut}(p)$$

is a diffeomorphism. We call $\text{int}(D_p)$ the **interior to the cut locus in the tangent space** $T_p \mathcal{M}$.

**Lemma 3.19.** A point $\gamma(s)$ is a cut point to $p$ along $\gamma$ if and only if $s$ is the smallest positive number such that either $\gamma(s)$ is a conjugate point to $p$ along $\gamma$ or there exist two distinct minimal geodesics joining $p$ and $\gamma(s)$.

**Proof.** See Lemma 5.2 in Cheeger and Ebin [3].

Given $V \in T_p \mathcal{M}$ and $s > 0$, we have $\gamma_V(s) = \exp_p(sV)$. For each unit vector $V \in T_p \mathcal{M}$ there exists at most a unique $r_V \in (0, \infty)$ such that $\gamma_V(r_V)$ is a cut point of $p$ along $\gamma_V$. Furthermore, if we set $r_V = \infty$ when $\gamma_V$ is a ray, then the map from the unit tangent space at $p$ to $(0, \infty]$ given by $V \mapsto r_V$ is a continuous function (see §11.6 of Bishop and Crittenden’s book [1] for example). Hence we have

$$C_p = \partial D_p = \{ r_V V : V \in T_p \mathcal{M}, \ |V| = 1, \ \gamma_V \text{ is not a ray} \}$$

has measure zero with respect to the Euclidean measure on $(T_p \mathcal{M}, g(p))$.\footnote{For $C_p$ is the radial graph of a continuous function.}

Since $\exp_p$ is a smooth function, we conclude that

**Lemma 3.20.** $\text{Cut}(p) = \exp_p(C_p)$ has measure zero with respect to the Riemannian measure on $(\mathcal{M}, g)$.

**Proposition 3.21.** We have that $r$ is $C^\infty$ in a neighborhood of a point $x \in \mathcal{M}$ if and only if $x \in \mathcal{M} - (\text{Cut}(p) \cup \{p\})$.

**Proof.** Now if $x \notin \text{Cut}(p)$ and $x \neq p$, then $r$ is $C^\infty$ at $x$ and $|\nabla r(x)| = 1$ by (3.10).

Since $\text{Cut}(p)$ has measure zero, we have

$$|\nabla r| = 1 \text{ a.e. on } \mathcal{M}.$$

\footnote{See Section 2.5 of [11] for a discussion of Riemannian measure.}
An alternate proof of Lemma 3.20 can be given as follows. Since \( r \) is locally Lipschitz, we have that \( r \) is differentiable a.e. On the other hand it is easy to see that \( r \) is not \( C^1 \) at those points \( x \) in \( \text{Cut} (p) \) for which there are two distinct minimal geodesics joining \( p \) to \( x \). Thus, this set of points in \( \text{Cut} (p) \) has measure zero. We also know that by Sard's theorem, the points in \( \text{Cut} (p) \) which are conjugate points form a measure zero set (since these points are singular values of \( \exp_p \)). We conclude that \( \text{Cut} (p) \) has measure zero.

**Definition 3.22.** The *injectivity radius* \( \text{inj} (p) \) of a point \( p \in M \) is defined to be the supremum of all \( r > 0 \) such that \( \exp_p \) is an embedding when restricted to \( B(\vec{0}, r) \). Equivalently,

1. \( \text{inj} (p) \) is the distance from \( \vec{0} \) to \( C_p \) with respect to \( g (p) \),
2. \( \text{inj} (p) \) is the Riemannian distance from \( p \) to \( \text{Cut} (p) \).

The injectivity radius of a Riemannian manifold \((M^n, g)\) is defined to be

\[
\text{inj} (M, g) \doteq \inf \{ \text{inj} (p) : p \in M \}.
\]

When \( M \) is closed, the injectivity radius is always positive.
Introduction to
Geometric Analysis

In geometric analysis the use of the covariant derivative, Hessian and Laplacian is ubiquitous.

1. The Hessian and Laplacian of a function

Let $(\mathcal{M}^n, g)$ be a Riemannian manifold. Given a $C^\infty$ function $f : \mathcal{M} \rightarrow \mathbb{R}$, its differential $df$ is the $C^\infty$ 1-form defined by

$$df(X) = Xf$$

for each $X \in T_p\mathcal{M}$, $p \in \mathcal{M}$. The covariant derivative of $df$ is a covariant 2-tensor:

$$(\nabla df)(X,Y) = X(df(Y)) - df(\nabla_XY) + \nabla_Y(Xf)$$

for each $X,Y \in T_p\mathcal{M}$, $p \in \mathcal{M}$. We call $\nabla df$ the Hessian of $f$.

From (4.1) we see that

$$(\nabla df)(X,Y) - (\nabla df)(Y,X) = X(Yf) - (\nabla_XY)(f) - Y(Xf) + (\nabla_YX)(f) = 0$$

since $XY - YX = [X,Y] = \nabla_XY - \nabla_YX$. Hence the Hessian of $f$ is a symmetric (covariant) 2-tensor:

$$(\nabla df)(X_p,Y_p) = (\nabla df)(Y_p,X_p)$$
for any $p \in \mathcal{M}$ and $X_p, Y_p \in T_p \mathcal{M}$.

Since the covariant derivative of $f$ is the same as the differential of $f$, one often uses the notation $\nabla f = df$. This notation is redundant with the gradient of $f$ defined to be the unique vector field $\nabla f = \text{grad } f$ such that

$$\left< (\text{grad } f)_p, X_p \right> = X_p f$$

for any $X \in T_p \mathcal{M}, p \in \mathcal{M}$. In any case, one often writes $\nabla df = \nabla \nabla f$. It is also customary to write the components of $\nabla \nabla f$ as

$$\nabla_i \alpha_j = (\nabla \alpha) \left( \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) = \frac{\partial^2 f}{\partial x^i \partial x^j} - \sum_{k=1}^n \Gamma^k_{ij} \alpha_k.$$

**Exercise 4.1.** Show that for a 1-form $\alpha$,

$$\nabla_i \alpha_j = (\nabla \alpha) \left( \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) = \frac{\partial^2 f}{\partial x^i \partial x^j} - \sum_{k=1}^n \Gamma^k_{ij} \alpha_k.$$

In particular, taking $\alpha = df$ yields

$$\nabla_i \nabla_j f = (\nabla \nabla f) \left( \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) = \frac{\partial^2 f}{\partial x^i \partial x^j} - \sum_{k=1}^n \Gamma^k_{ij} \frac{\partial f}{\partial x^k}.$$

**Exercise 4.2.** Show that for any vector field $X$ we have

$$(\mathcal{L}_X g)_{ij} = (\mathcal{L}_X g) \left( \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) = \nabla_i X^*_j + \nabla_j X^*_i,$$

where $X^*$ is the 1-form defined by $X_p^*(Y_p) = g(X_p, Y_p)$ for each $Y_p \in T_p \mathcal{M}$. In particular,

$$(\mathcal{L}_{\text{grad } f} g)_{ij} = (\mathcal{L}_{\text{grad } f} g) \left( \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) = 2\nabla_i \nabla_j f.$$

Note that $(\text{grad } f)^* = df$.

For any symmetric covariant 2-tensor $\beta$, we may define its trace by

$$(\text{trace}_g \beta)(p) = \sum_{i=1}^n \beta(e_i, e_i),$$

where $\{e_i\}_{i=1}^n$ is an orthonormal basis of tangent vectors at $p$. This definition is independent of the choice of $\{e_i\}_{i=1}^n$.

**Exercise 4.3.** Define $\beta_{ij} = \beta(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j})$. Show that in $x(U)$,

$$\text{trace}_g \beta = \sum_{i,j=1}^n g^{ij} \beta_{ij}.$$
The **Laplacian** of a function \( f : \mathcal{M} \to \mathbb{R} \) is the trace of its Hessian:

\[
\Delta f \doteq \sum_{i=1}^{n} (\nabla df) (e_i, e_i) = \sum_{i,j=1}^{n} g^{ij} \nabla_i \nabla_j f = \sum_{i,j=1}^{n} g^{ij} \left( \frac{\partial^2 f}{\partial x^i \partial x^j} - \Gamma^k_{ij} \frac{\partial f}{\partial x^k} \right).
\]

**Exercise 4.4.** Show that in \( x(\mathcal{U}) \),

\[
\Delta f = \frac{1}{\sqrt{|g|}} \sum_{i,j=1}^{n} \frac{\partial}{\partial x^i} \left( \sqrt{|g|} g^{ij} \frac{\partial f}{\partial x^j} \right),
\]

where \( |g| \doteq \det (g_{ij}) \).

Let \( \alpha \) be a 1-form on \( \mathcal{M} \). Its **divergence** is defined by

\[
\text{div} \ \alpha \doteq \sum_{i=1}^{n} (\nabla \alpha) (e_i, e_i) = \sum_{i,j=1}^{n} g^{ij} \nabla_i \alpha_j.
\]

We have \( \Delta f = \text{div} \ (df) \).

The definition of Laplacian extends to tensors. In particular, if \( T \) is a covariant \( k \)-tensor on \( \mathcal{M} \), then \( \nabla T \) is a covariant \((k+1)\)-tensor and \( \nabla (\nabla T) \) is a covariant \((k+2)\)-tensor. The **rough Laplacian** of \( T \) is defined as the trace of the second covariant derivative by

\[
\Delta T (X_1, \ldots, X_k) \doteq \sum_{a=1}^{n} \nabla \nabla T (e_a, e_a, X_1, \ldots, X_k)
\]

for all tangent vectors \( X_1, \ldots, X_k \). In particular, \( \Delta T \) is also a covariant \( k \)-tensor, defined by

\[
\Delta T = \sum_{a=1}^{n} \nabla \nabla T (e_a, e_a, \cdots, \cdots).
\]

**Exercise 4.5.**

1. **Show that if** \( \alpha \) **is a 1-form, then**

\[
(\nabla \nabla \alpha) (X, Y, Z) = (\nabla_X (\nabla_Y \alpha) - \nabla_{X Y} \alpha) (Z).
\]

2. **Show that**

\[
|\nabla \alpha|^2 = \sum_{a=1}^{n} |\nabla_{e_a} \alpha|^2 = \sum_{a,b=1}^{n} (\nabla_{e_a} \alpha (e_b))^2.
\]

**Exercise 4.6.** Show that if \( \beta \) **is a symmetric covariant 2-tensor**, then

\[
\sum_{i=1}^{n} (\nabla_X \beta) (e_i, e_i) = X \left( \text{trace}_{g} \beta \right).
\]

In particular, taking \( \beta = \nabla df \), we obtain

\[
\sum_{i=1}^{n} (\nabla_X (\nabla df)) (e_i, e_i) = X(\Delta f).
\]
Recall that by (2.64) that if \( \alpha \) is a 1-form, then
\[
\nabla_X (\nabla \alpha) (Y, Z) - \nabla_Y (\nabla \alpha) (X, Z) = -\alpha (Rm (X, Y) Z).
\]
Thus
\[
\Delta (d f) (X) = \sum_{i=1}^{n} (\nabla_{e_i} (\nabla d f)) (e_i, X)
\]
\[
= \sum_{i=1}^{n} (\nabla_{e_i} (\nabla d f)) (X, e_i)
\]
\[
= \sum_{i=1}^{n} ((\nabla_X (\nabla d f)) (e_i, e_i) - d f (Rm (e_i, X) e_i))
\]
\[
= X (\Delta f) + d f (Rc (X)),
\]
where \( Rc (X) \) is defined by (2.57) and where we used (2.63) for the second equality. Hence

**Lemma 4.7** (Commutator of \( \Delta \) and \( f \) acting on functions). For any \( C^\infty \) function \( f \)
\[
(4.7) \quad \Delta (d f) = d (\Delta f) + Rc (\text{grad} f).
\]

**Exercise 4.8.** Show that
\[
d f (Rc (X)) = Rc (\text{grad} f).
\]

**Lemma 4.9.** If \( \alpha \) is a 1-form, then
\[
(4.8) \quad \Delta |\alpha|^2 = 2 \langle \Delta \alpha, \alpha \rangle + 2 |\nabla \alpha|^2,
\]
where \( \Delta \alpha \) is the rough Laplacian of \( \alpha \).

**Proof.** We have by (2.27),
\[
\left( d |\alpha|^2 \right) (X) = X \langle \alpha, \alpha \rangle
\]
\[
= 2 \langle \nabla_X \alpha, \alpha \rangle.
\]
By (4.1), (2.27), and (4.5), we have
\[
\left( \nabla d |\alpha|^2 \right) (X, Y) = X \left( Y \left( |\alpha|^2 \right) \right) - \left( \nabla_X Y \right) \left( |\alpha|^2 \right)
\]
\[
= 2 X \langle \nabla_Y \alpha, \alpha \rangle - 2 \langle \nabla_{\nabla_X Y} \alpha, \alpha \rangle
\]
\[
= 2 \langle \nabla_X (\nabla_Y \alpha), \alpha \rangle + 2 \langle \nabla_Y \alpha, \nabla_X \alpha \rangle - 2 \langle \nabla_{\nabla_X Y} \alpha, \alpha \rangle
\]
\[
= 2 \langle \langle \nabla \nabla \alpha \rangle (X, Y), \cdot \rangle, \alpha \rangle + 2 \langle \nabla_Y \alpha, \nabla_X \alpha \rangle.
\]
Now taking \( X = Y = e_a \), summing \( a \) from 1 to \( n \), and using (4.6), we obtain (4.8). \( \square \)
By (4.7) and (4.8), we have
\[ \Delta |df|^2 = 2 \langle d(\Delta f) + \text{Rc} (\nabla \! f), df \rangle + 2 |\nabla df|^2 \]
\[ = 2 \langle d(\Delta f), df \rangle + 2 \text{Rc} (\nabla \! f, \nabla \! f) + 2 |\nabla df|^2. \]

We often write this formula as
\[ \Delta |\nabla \! f|^2 = 2 \langle \nabla (\Delta f), \nabla \! f \rangle + 2 \text{Rc} (\nabla \! f, \nabla \! f) + 2 |\nabla \nabla \! f|^2. \]

Notice that
\[ g^* (df, dh) = g (\nabla \! f, \nabla \! h) \]
for any \( h \in C^\infty (\mathcal{M}) \).

Let \( \gamma \) be a covariant \( k \)-tensor. Then its trace, on the first two components, is the covariant \((k - 2)\)-tensor defined by
\[ (\text{trace}_{1,2} \gamma)_p (X_1, \ldots, X_{k-2}) = \sum_{a=1}^{n} \gamma_p (e_a, e_a, X_1, \ldots, X_{k-2}), \]
where \( \{e_a\}_{a=1}^{n} \) is an orthonormal basis of \( T_p \mathcal{M} \). Note that the rough Laplacian of \( \gamma \) is given by
\[ \Delta \gamma = \text{trace}_{1,2} (\nabla \nabla \gamma). \]

**Exercise 4.10.** Show that the definition of \( \text{trace}_{1,2} \gamma \) is independent of the choice of orthonormal basis.

Let \( \beta \) be a covariant \( k \)-tensor. Then its divergence is the covariant \((k - 1)\)-tensor defined by
\[ \text{div} (\beta) = \text{trace}_{1,2} (\nabla \beta). \]

If \( f \in C^\infty (\mathcal{M}) \), then we have
\[ \text{div} (\nabla df) = \text{trace}_{1,2} (\nabla \nabla df) = \Delta (df). \]

The contracted second Bianchi identity says that
\[ \text{div} (\text{Rc}) = \frac{1}{2} dR. \]

Suppose that \( (\mathcal{M}^n, g) \) is a \( C^\infty \) Riemannian manifold and \( f \in C^\infty (\mathcal{M}) \) satisfy the equation
\[ \text{Rc} + \nabla df = 0. \]
This equation is known as the quasi-Einstein metric equation or the Ricci soliton equation. Taking the trace of this equation, we have
\[ R + \Delta f = 0. \]

On the other hand, taking the divergence of this equation yields
\[ \frac{1}{2} dR + \Delta (df) = 0. \]
4. Introduction to Geometric Analysis

By (4.7), we then have
\[ 0 = \frac{1}{2} dR + d(\Delta f) + \text{Rc} (\text{grad } f) \]
\[ = -\frac{1}{2} dR + \text{Rc} (\text{grad } f) \]
since \( d(\Delta f) = -dR \) by (4.11). We also write this as

\[ (4.12) \quad 2 \text{Rc} (\nabla f) = \nabla R. \]

Now by (4.10),
\[ \left\langle \nabla |df|^2, X \right\rangle = \left( d|df|^2 \right)(X) \]
\[ = 2 (\nabla df)(X, \nabla f) \]
\[ = -2 \text{Rc} (X, \nabla f) \]
\[ = -2 \left\langle \text{Rc} (\nabla f), X \right\rangle. \]

Hence
\[ \nabla |df|^2 = -2 \text{Rc} (\nabla f). \]

Combining this with (4.12) yields
\[ \nabla \left( R + |df|^2 \right) = 0. \]

Since we always assume \( \mathcal{M} \) is connected, we conclude that
\[ R + |df|^2 = C \]
on \( \mathcal{M} \), where \( C \) is a constant.

From (4.9) while using (4.10), we obtain
\[ -\Delta R = -\langle \nabla R, \nabla f \rangle + 2 |\text{Rc}|^2 \]
since \( \Delta f = -R = |df|^2 - C \) and since \( 2 \text{Rc} (\nabla f, \nabla f) = \langle \nabla R, \nabla f \rangle \) by (4.12).

That is,
\[ \Delta R + 2 |\text{Rc}|^2 = \langle \nabla R, \nabla f \rangle. \]

2. Hypersurfaces in Riemannian manifolds

Throughout this section \( (\mathcal{M}^n, g) \) shall denote a Riemannian manifold.
2. Hypersurfaces in Riemannian manifolds

2.1. Level sets.

Let \( f : \mathcal{M} \rightarrow \mathbb{R} \) be a \( C^\infty \) function and let \( c \) be a regular value, i.e., \( df(x) \neq 0 \) for each \( x \in f^{-1}(c) \). Then the level set \( f^{-1}(c) \) is a \( C^\infty \) hypersurface. A choice of unit normal vector field \( \nu \) to \( f^{-1}(c) \) is

\[
\nu = \frac{\nabla f}{|\nabla f|},
\]

where \( \nabla f \) denotes the gradient of \( f \) and where \( \nu \) is well defined since \( |\nabla f(x)| = |df(x)| > 0 \) for each \( x \in f^{-1}(c) \). Using the definition of the second fundamental form, we compute that

\[
\mathrm{II} (X,Y) = \langle \nabla_X \nu, Y \rangle = \langle \nabla_X \nu^* (Y) \rangle
\]

\[
= \langle \nabla_X \frac{df}{|\nabla f|} (Y) \rangle
\]

\[
= \frac{\nabla_X (df)}{|\nabla f|} (Y) + X \left( \frac{1}{|\nabla f|} \right) (df) (Y)
\]

\[
= \frac{(df)(X,Y)}{|\nabla f|},
\]

where \( \nu^* = g(\nu) \) as in (2.68) and since \( (df)(Y) = 0 \).

2.2. Parametrized hypersurfaces.

Let

\[
F : \mathcal{N}^{n-1} \rightarrow \mathcal{M}^n
\]

be an immersion, so that \( dF_y : T_y\mathcal{N} \rightarrow T_{F(y)}\mathcal{M} \) is an injection for each \( y \in \mathcal{N} \). We call \( F \) a parametrized hypersurface. Then \( dF_y(T_y\mathcal{N}) \subset T_{F(y)}\mathcal{M} \) is a hyperplane. Assume that there exists a continuous (and hence \( C^\infty \)) choice of unit normal vector field \( \nu(y) \in T_{F(y)}\mathcal{M} \). That is,

\[
\nu : \mathcal{N} \rightarrow T\mathcal{M}
\]

is a \( C^\infty \) map such that \( \pi \circ \nu = F \), where \( \pi : T\mathcal{M} \rightarrow \mathcal{M} \) is the projection map, \( \langle \nu (y), X \rangle = 0 \) for each \( X \in dF_y(T_y\mathcal{N}) \), and \( |\nu| = 1 \).

Let \( T_y \triangleq dF_y(T_y\mathcal{N}) \) for \( y \in \mathcal{N} \). The first fundamental form \( I_y : T_y \times T_y \rightarrow \mathbb{R} \) is defined by

\[
I_y (X,Y) \triangleq g_{F(y)} (X,Y)
\]

for \( X,Y \in T_y \). We may define the pulled back metric \( \gamma \) on \( \mathcal{M} \) by \( \gamma \triangleq F^* (g) \), i.e.,

\[
\gamma (U,V) \triangleq g (dF (U), dF (V)) = I_y (dF (U), dF (V))
\]
for each \( U, V \in T_yN, y \in N \), where the second equality holds since \( dF(U), dF(V) \in T_y \).

We now define the second fundamental form. Given \( y \in N \) and \( U \in T_yN \), let \( c: (-\varepsilon, \varepsilon) \rightarrow N \) be a \( C^\infty \) curve with \( c'(0) = U \). Consider \( \nu \) as a vector field along the curve \( F \circ c: (-\varepsilon, \varepsilon) \rightarrow M \), i.e.,

\[
\nu(t) \equiv \nu(c(t)) \in T_{(F \circ c)(t)}M.
\]

The \textbf{(pulled back) second fundamental form} \( (F^* \Pi)_y: T_yN \times T_yN \rightarrow \mathbb{R} \) is defined by

\[
(F^* \Pi)_y(U, V) \equiv g_{F(y)} \left( \frac{D}{dt} \nu, dF \right)
\]

for \( U, V \in T_yN \), where \( \frac{D}{dt} \nu \) is the covariant derivative of \( \nu \) at \( t = 0 \) along the curve \( F \circ c \) with \( c'(0) = U \).

The \textbf{second fundamental form} \( \Pi_y : T_y \times T_y \rightarrow \mathbb{R} \) is then defined by

\[
\Pi_y(X, Y) \equiv (F^* \Pi)_y(dF^{-1}(X), dF^{-1}(Y))
\]

\[
= g_{F(y)} \left( \frac{D}{dt} \nu, Y \right)
\]

for \( X, Y \in T_y \), \( y \in N \), and where \( \frac{D}{dt} \nu \) is the covariant derivative of \( \nu \) at \( t = 0 \) along the curve \( F \circ c \) with \( c'(0) = dF^{-1}(X) \). This is often written using simpler notation as

\[
\Pi_y(X, Y) = g_{F(y)}(\nabla_X \nu, Y).
\]

Now let \( (U, x) \) be a parametrization of \( N \). Then we have the coordinate vector fields \( \{ \frac{\partial}{\partial x^i} \}_{i=1}^{n-1} \) on \( x(U) \subset N \). Define

\[
\frac{\partial F}{\partial x^i}(y) \equiv dF_y(\frac{\partial}{\partial x^i}) \in T_y \subset T_{f(y)}M
\]

for \( y \in x(U) \). Note that \( \{ \frac{\partial F}{\partial x^i}(y) \}_{i=1}^{n-1} \) is a basis of \( T_y \) for each \( y \in x(U) \).

The components of the first fundamental form are

\[
I_{ij}(y) \equiv \gamma_{ij}(y)
\]

\[
= \gamma \left( \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right)_y = I_y \left( \frac{\partial F}{\partial x^i}(y), \frac{\partial F}{\partial x^j}(y) \right)
\]

for \( i, j = 1, \ldots, n - 1 \).
The components of the second fundamental form are
\[ \Pi_{ij} (y) \doteq (F^* \Pi)_{ij} \left( \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) \]
\[ = \Pi_y \left( \frac{\partial F}{\partial x^i} (y), \frac{\partial F}{\partial x^j} (y) \right) \]
\[ = g_{F(y)} \left( \frac{D}{dt} \nu, \frac{\partial F}{\partial x^j} \circ c_i \right), \]
where \( \frac{D}{dt} \nu \) is the covariant derivative of \( \nu \) at \( t = 0 \) along the curve \( F \circ c_i \) with \( c'_i (0) = \frac{\partial}{\partial x^i} y \). Consider \( \left( \frac{\partial F}{\partial x^i} \circ c_i \right) (t) \) as a vector field along \( F \circ c_i \). Then
\[ \Pi_{ij} (y) = g_{F(y)} \left( \frac{D}{dt} \nu, \frac{\partial F}{\partial x^j} \circ c_i \right). \]

3. Distance functions

Suppose \( r : \mathcal{M} \to \mathbb{R} \) is \( C^\infty \) on an open subset \( \mathcal{V} \) and \( |\nabla r| \equiv 1 \) on \( \mathcal{V} \). Given \( c \in \mathbb{R} \), let \( \mathcal{S}_c \doteq r^{-1} (c) \cap \mathcal{V} \), which is a \( C^\infty \) hypersurface. Then a choice of unit normal vector field \( \nu \) to \( \mathcal{S}_c \) is
\[ \nu \doteq \nabla r. \]

The corresponding second fundamental form is
\[ \Pi (X, Y) = \langle \nabla_X \nu, Y \rangle = (\nabla dr) (X, Y) \]
for \( X, Y \in T_p \mathcal{S}_c, \ p \in \mathcal{V} \). Let \( \{e_i\}_{i=1}^n \) be an orthonormal basis for \( T_p \mathcal{M} \) with \( e_n = \nu \). Then \( \{e_i\}_{i=1}^{n-1} \) is an orthonormal basis for \( T_p \mathcal{S}_c \). The mean curvature of \( \mathcal{S}_c \) is
\[ H = \sum_{i=1}^{n-1} \Pi (e_i, e_i) = \sum_{i=1}^{n-1} (\nabla dr) (e_i, e_i). \]

Now \( 0 = \nu |\nabla r|^2 = 2 \langle \nabla \nu, \nabla r \rangle = 2(\nabla dr) (\nu, \nabla r) = 2(\nabla dr) (e_n, e_n). \) Hence
\[ \Delta r = \sum_{i=1}^n (\nabla dr) (e_i, e_i) = H. \]

Note that \( H : \mathcal{V} \to \mathbb{R} \) is a \( C^\infty \) function and \( H (p) \) is equal to the mean curvature at \( p \) of the hypersurface \( \mathcal{S}_c \), where \( c = r (p) \).
We compute using (4.9) that
\[ \nu(H) = \langle \nabla r, \nabla \Delta r \rangle \]
\[ = \Delta |\nabla r|^2 - Rc \langle \nabla r, \nabla r \rangle - |\nabla dr|^2 \]
\[ = -Rc \langle \nabla r, \nabla r \rangle - |\nabla II|^2 \]
since \( \Delta |\nabla r|^2 = 0 \). This implies
\[ \nu(H) \leq -Rc \langle \nabla r, \nabla r \rangle - H^2 \]
\[ - \frac{n-1}{n} \frac{H^2}{n-1}. \]

Let \((\mathcal{M}^n, g)\) be a complete Riemannian manifold with \(Rc \geq 0\). Then
\[ \nu(H) \leq -\frac{H^2}{n-1}. \] (4.13)

Now fix a point \(O \in \mathcal{M}\) and let \(r(p) \equiv d(p, O)\). There exists \(\varepsilon > 0\) such that \(r\) is \(C^\infty\) in \(B(O, \varepsilon) - \{O\}\). Moreover,
\[ \lim_{p \to O} r(p) H(p) = n - 1. \] (4.14)

By (4.13), we have
\[ \nu(r^2H) \leq -\frac{r^2H^2}{n-1} + 2rH \leq n - 1. \]

Let \(\gamma : [0, \varepsilon) \to \mathcal{M}\) be any unit speed geodesic with \(\gamma(0) = O\). By the Gauss lemma, \(\gamma'(t) = \nu(\gamma(t))\) and hence
\[ \frac{d}{dt} (r^2H)(\gamma(t)) = \nu(r^2H)(\gamma(t)) \leq n - 1 \]
for each \(t \in (0, \varepsilon)\). Since \(\lim_{t \to 0} (r^2H)(\gamma(t)) = 0\), we obtain
\[ (r^2H)(\gamma(t)) \leq (n - 1) t = (n - 1) r(\gamma(t)) \]
for each \(t \in (0, \varepsilon)\). Therefore
\[ H(p) \leq \frac{n-1}{r(p)} \]
for each \(p \in B(O, \varepsilon) - \{O\}\).

4. Volume

Sakai [11]

Let \(\mathcal{M}^n\) be a \(C^\infty\) differentiable manifold with a countable atlas \(\{(U_\alpha, x_\alpha)\}_{\alpha \in A}\) such that the closure \(x_\alpha(U_\alpha) \subset \mathcal{M}\) is compact. Then there exist open subsets \(W_\alpha \subset U_\alpha\), for \(\alpha \in A\), such that each \(W_\alpha\) is compact and \(\bigcup_{\alpha \in A} x_\alpha(W_\alpha) = \mathcal{M}\). We define a subset \(S \subset \mathcal{M}\) to be \textbf{measurable} if for each \(\alpha \in A\) we have that \(x_\alpha^{-1}(S \cap W_\alpha) \subset \mathbb{R}^n\) is Lebesgue measurable.
5. Covariant tensors and differentiable forms

Now let \( g \) be a Riemannian metric on \( M \). Let \( p \in M \) and let \( (U, x) \) be a parametrization with \( p \in x(U) \). Then \( \left\{ \frac{\partial}{\partial x^i} \right\}_{i=1}^n \) is a basis of \( T_pM \), where \( \frac{\partial}{\partial x^i} = \frac{\partial}{\partial x^i} p \). The volume, with respect to \( g_p \), of the parallelepiped spanned by \( \left\{ \frac{\partial}{\partial x^i} \right\}_{i=1}^n \) is equal to

\[
\sqrt{\det g_{ij}(p)}.
\]

We call this the \textit{infinitesimal volume} or \textit{volume element} since it is at \( p \), i.e., only at a point.

Let \( d\mu_{\mathbb{R}^n} \) denote the Lebesgue measure on \( \mathbb{R}^n \). Let \( R \subset x(U) \) be a subset such that \( x^{-1}(R) \subset \mathbb{R}^n \) is a measurable set. Given \( f : R \to \mathbb{R} \) such that \( f \circ x|_{x^{-1}(R)} : x^{-1}(R) \to \mathbb{R} \) is a measurable function, we define

\[
\int_R f \, d\mu = \int_{x^{-1}(R)} f(x(q)) \sqrt{\det g \left( \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right)(x(q))} \, d\mu_{\mathbb{R}^n}(q)
\]

provided the function \( q \mapsto f(x(q)) \sqrt{\det g_{ij}(x(q))} \), for \( q \in x^{-1}(R) \), is integrable.

For simplicity, assume that \( S \subset M \) is compact and that \( f : S \to \mathbb{R} \) is bounded and measurable. Let \( \{\psi_\alpha\}_{\alpha \in A} \) be a partition of unity subordinate to the covering \( \{x_\alpha(W_\alpha)\}_{\alpha \in A} \). We define

\[
\int_S f \, d\mu \doteq \sum_{\alpha \in A} \int_{x_\alpha^{-1}(S) \cap W_\alpha} (\psi_\alpha f)(x_\alpha(q)) \sqrt{\det g \left( \frac{\partial}{\partial x^i_\alpha}, \frac{\partial}{\partial x^j_\alpha} \right)(x_\alpha(q))} \, d\mu_{\mathbb{R}^n}(q).
\]

Note that \( \text{supp}(\psi_\alpha f) \subset x_\alpha(W_\alpha) \). The \textit{volume} of \( S \) is defined by

\[
\text{Vol}(S) \doteq \int_S d\mu.
\]

In particular, if \( M \) is compact, then the volume of \( M \) is

\[
\text{Vol}(M) \doteq \int_M d\mu.
\]

5. Covariant tensors and differentiable forms

Let \( V \) be an \( n \)-dimensional real vector space and let \( V^* \) denote its dual vector space. Given \( k \in \mathbb{Z}^+ \), define \( \bigotimes^k V^* \) to be the vector space of multilinear functions on \( \times^k V \doteq V \times \cdots \times V \) (\( k \)-fold product). That is, \( T \in \bigotimes^k V^* \) iff

\[
T : \times^k V \to \mathbb{R}
\]

satisfies that for each \( i = 1, 2, \ldots, k \),

\[
T(V_1, \ldots, V_{i-1}, cV + W, V_{i+1}, \ldots, V_k) = cT(V_1, \ldots, V_{i-1}, V, V_{i+1}, \ldots, V_k) + T(V_1, \ldots, V_{i-1}, W, V_{i+1}, \ldots, V_k)
\]
for \( c \in \mathbb{R} \) and \( V, W \in \mathbb{V} \). We call \( T \) a \textbf{covariant} \( k \)-\textbf{tensor} (for \( \mathbb{V} \)). Of course, \( \otimes^1 \mathbb{V}^* = \mathbb{V}^* \).

Given \( \alpha_1, \ldots, \alpha_k \in \mathbb{V}^* \), define their \textbf{tensor product}

\[
\alpha_1 \otimes \cdots \otimes \alpha_k \in \bigotimes^k \mathbb{V}^*
\]

by

\[
(\alpha_1 \otimes \cdots \otimes \alpha_k)(V_1, \ldots, V_k) = \alpha_1(V_1) \cdots \alpha_k(V_k).
\]

It is clear that this map is multilinear.

**Exercise 4.11.** Suppose that \( \{e_j\}_{j=1}^n \) is a basis for \( \mathbb{V} \) and let \( \{e^*_i\}_{i=1}^n \) be its dual basis of \( \mathbb{V}^* \), i.e., \( e_i^*(e_j) = \delta_{ij} \). Show that

\[
\{e^*_i \otimes \cdots \otimes e^*_k\}_{i_1 \ldots i_k = 1}^n
\]

is a basis of \( \bigotimes^k \mathbb{V}^* \). In particular, \( \dim(\bigotimes^k \mathbb{V}^*) = n^k \). Show for any \( T \in \bigotimes^k \mathbb{V}^* \) that

\[
(4.15) \quad T = \sum_{i_1, \ldots, i_k = 1}^n T(e_{i_1}, \ldots, e_{i_k}) e^*_{i_1} \otimes \cdots \otimes e^*_{i_k}.
\]

More generally, a multilinear function \( S : (\times^k \mathbb{V}) \times (\times^\ell \mathbb{V}^*) \to \mathbb{R} \) is called a \((k, \ell)\)-\textbf{tensor} (for \( \mathbb{V} \)). Let \( \bigotimes^k \mathbb{V}^* \otimes \bigotimes^\ell \mathbb{V} \) denote the space of all such functions. Given \( \alpha_1, \ldots, \alpha_k \in \mathbb{V}^* \) and \( W_1, \ldots, W_\ell \in \mathbb{V} \), define

\[
\alpha_1 \otimes \cdots \otimes \alpha_k \otimes W_1 \otimes \cdots \otimes W_\ell \in \bigotimes^k \mathbb{V}^* \otimes \bigotimes^\ell \mathbb{V}
\]

by

\[
(\alpha_1 \otimes \cdots \otimes \alpha_k \otimes W_1 \otimes \cdots \otimes W_\ell)(V_1, \ldots, V_k, \beta_1, \ldots, \beta_\ell) = \alpha_1(V_1) \cdots \alpha_k(V_k) \beta_1(W_1) \cdots \beta_\ell(W_\ell)
\]

for \( V_1, \ldots, V_k \in \mathbb{V} \) and \( \beta_1, \ldots, \beta_\ell \in \mathbb{V}^* \).

**Exercise 4.12.** Show that

\[
\{e^*_{i_1} \otimes \cdots \otimes e^*_{i_k} \otimes e_{j_1} \otimes \cdots \otimes e_{j_\ell}\}_{i_1, \ldots, i_k, j_1, \ldots, j_\ell = 1}^n
\]

is a basis of \( \bigotimes^k \mathbb{V}^* \otimes \bigotimes^\ell \mathbb{V} \), which thus has dimension \( n^{k+\ell} \).

A multilinear function \( A : \times^k \mathbb{V} \to \mathbb{R} \) is said to be \textbf{alternating} (or \textbf{fully antisymmetric}) if

\[
A(V_1, \ldots, V_k) = \text{sign}(I) A(V_{i_1}, \ldots, V_{i_k})
\]

for any \( V_1, \ldots, V_k \in \mathbb{V} \) and permutation \( I : (1, \ldots, k) \to (i_1, \ldots, i_k) \), where \( \text{sign}(I) \) is the sign of \( I \). We call \( A \) a \textbf{\( k \)-form}. We denote the set of \( k \)-forms by \( \Lambda^k \mathbb{V}^* \). Clearly \( \Lambda^k \mathbb{V}^* \subset \bigotimes^k \mathbb{V}^* \).
The **wedge product** of a \( k \)-form \( A \) and a \( \ell \)-form \( B \) is the \((k + \ell)\)-form defined by

\[
(A \wedge B) (V_1, \ldots, V_{k+\ell}) = \frac{1}{(k+\ell)!} \sum_J \text{sign}(J) A(V_{j_1}, \ldots, V_{j_k}) B(V_{j_{k+1}}, \ldots, V_{j_{k+\ell}})
\]

for \( V_1, \ldots, V_{k+\ell} \in \mathcal{V} \), where the sum on the RHS is over all permutations \( J : (1, \ldots, k + \ell) \mapsto (j_1, \ldots, j_{k+\ell}) \). The wedge product is associative:

\[
(A \wedge B) \wedge C = A \wedge (B \wedge C)
\]

for any \( k \)-form \( A \), \( \ell \)-form \( B \) and \( m \)-form \( C \). We have that

\[
\{ e^*_i \wedge \cdots \wedge e^*_i \}_{1 \leq i_1 \prec \cdots \prec i_k \leq n}
\]

is a basis of \( \Lambda^k \mathcal{V}^* \). In particular, \( \dim(\Lambda^k \mathcal{V}^*) = \binom{n}{k} \).

Let \( \mathcal{M} \) be a \( C^\infty \) differentiable manifold. The **covariant \( k \)-tensor bundle** of \( \mathcal{M} \) is

\[
\bigotimes^k T^* \mathcal{M} = \bigsqcup_{p \in \mathcal{M}} \bigotimes^k T^*_p \mathcal{M}.
\]

Equivalently, we may define it as the point-set

\[
\bigotimes^k T^* \mathcal{M} = \left\{ (p, T) : p \in \mathcal{M}, T \in \bigotimes^k T^*_p \mathcal{M} \right\}.
\]

We have the projection map \( \pi : \bigotimes^k T^* \mathcal{M} \to \mathcal{M} \) defined by \( \pi (p, T) \equiv p \).

A \( C^\infty \) **covariant \( k \)-tensor** on \( \mathcal{M} \) is a \( C^\infty \) section of the vector bundle \( \bigotimes^k T^* \mathcal{M} \), i.e., it is a \( C^\infty \) map \( T : \mathcal{M} \to \bigotimes^k T^* \mathcal{M} \) such that \( \pi \circ T = \text{id}_\mathcal{M} \). We denote by \( C^\infty (\bigotimes^k T^* \mathcal{M}) \) the set of \( C^\infty \) covariant \( k \)-tensors on \( \mathcal{M} \). Note that a covariant 1-tensor on \( \mathcal{M} \) is the same as a 1-form on \( \mathcal{M} \).

Let \( (U, x) \) be a parametrization of \( \mathcal{M} \). Then \( \{dx^i\}_{i=1}^n \) are \( C^\infty \) 1-forms on \( x(U) \) such that \( \{dx^i_p\}_{i=1}^n \) is a basis of \( T^*_p \mathcal{M} \) for each \( p \in x(U) \). Hence

\[
\{dx^i_p \otimes \cdots \otimes dx^i_p\}_{i_1, \ldots, i_k=1}^n
\]

is a basis of \( \bigotimes^k T^*_p \mathcal{M} \). If \( T \in C^\infty (\bigotimes^k T^* \mathcal{M}) \), then by (4.15) we have on \( x(U) \),

\[
T = \sum_{i_1, \ldots, i_k=1}^n T \left( \frac{\partial}{\partial x^{i_1}}, \ldots, \frac{\partial}{\partial x^{i_k}} \right) dx^{i_1}_p \otimes \cdots \otimes dx^{i_k}_p.
\]

It is customary to write

\[
T_{i_1 \cdots i_k} \equiv T \left( \frac{\partial}{\partial x^{i_1}}, \ldots, \frac{\partial}{\partial x^{i_k}} \right),
\]
4. Introduction to Geometric Analysis

so that

\[ T = \sum_{i_1, \ldots, i_k=1}^{n} T_{i_1 \ldots i_k} dx^{i_1}_p \otimes \cdots \otimes dx^{i_k}_p. \]

For example, a Riemannian metric \( g \) may be written on \( x(U) \) as

\[ g = \sum_{i,j=1}^{n} g_{ij} dx^i \otimes dx^j. \]

Given a covariant \( k \)-tensor \( T \) on \( M^n \) and a smooth map \( \varphi : N^m \to M \), we define the pull back of \( T \) to \( N \) by

\[ (\varphi^* T)(V_1, \ldots, V_k) \mapsto T(d\varphi(V_1), \ldots, d\varphi(V_k)) \]

for all \( V_1, \ldots, V_k \in T_xN \), \( x \in N \). Then \( \varphi^* T \) is a covariant \( k \)-tensor on \( N \).

**Exercise 4.13.** Corresponding to the definition of a \((k,\ell)\)-tensor for \( \nabla \), define a \((k,\ell)\)-tensor of \( M \).

A (differential) \( k \)-form on \( M \) is a covariant \( k \)-tensor on \( M \) which is alternating (or fully antisymmetric), i.e.,

\[ T(V_1, \ldots, V_k) = \text{sign}(I) T(V_{i_1}, \ldots, V_{i_k}) \]

for any permutation \( I : (1, \ldots, k) \mapsto (i_1, \ldots, i_k) \). As in (4.16), the wedge product of a \( k \)-form \( T \) and a \( \ell \)-form \( U \) on \( M \) is given by

\[ (T \wedge U)(V_1, \ldots, V_{k+\ell}) \]

\[ \mapsto \frac{1}{(k+\ell)!} \sum_J \text{sign}(J) T(V_{j_1}, \ldots, V_{j_k}) U(V_{j_{k+1}}, \ldots, V_{j_{k+\ell}}). \]

For each \( p \in x(U) \), let \( \Lambda^k T^*_p M \) denote the space of alternating multilinear functions on \( x^k T_p M \). From (4.17), we have

\[ \{dx^{i_1}_p \wedge \cdots \wedge dx^{i_k}_p \}_{1 \leq i_1 < \cdots < i_k \leq n} \]

is a basis of \( \Lambda^k T^*_p M \). Define

\[ \Lambda^k T^* M \doteq \left\{ (p, T) : p \in M, \ T \in \Lambda^k T^*_p M \right\}. \]

A section of \( \Lambda^k T^* M \) is the same as a differential \( k \)-form.
The exterior derivative of a $k$-form $\alpha$ satisfies

\begin{equation}
(d\alpha)(V_0, \ldots, V_k) = \frac{1}{k+1} \sum_{j=0}^{k} (-1)^j V_j \left( \alpha \left( V_0, \ldots, \widehat{V_j}, \ldots, V_k \right) \right) + \frac{1}{k+1} \sum_{0 \leq i < j \leq k} (-1)^{i+j} \alpha \left( [V_i, V_j], V_0, \ldots, \widehat{V_i}, \ldots, \widehat{V_j}, \ldots, V_k \right),
\end{equation}

where $\widehat{V_j}$ means to omit $V_j$. This generalizes definition (2.65) for the exterior derivative of a 1-form.

Using the product rule and the fact that $\nabla$ is torsion-free, we may express $d\alpha$ in terms of covariant derivatives as

\begin{equation}
(d\alpha)(V_0, \ldots, V_k) = \frac{1}{k+1} \sum_{j=0}^{k} (-1)^j \left( \nabla V_j \alpha \right) \left( V_0, \ldots, \widehat{V_j}, \ldots, V_k \right) \equiv (d\nabla \alpha)(V_0, \ldots, V_k).
\end{equation}

The RHS is called the **exterior covariant derivative** of $\alpha$. Formula (4.20) says that, acting on differential forms, the exterior derivative $d$ is equal to the exterior covariant derivative $d\nabla$.

For example, if $\alpha$ is a 1-form, then

\[ (d\alpha)(X, Y) = \frac{1}{2} \left\{ (\nabla_X \alpha)(Y) - (\nabla_Y \alpha)(X) \right\}. \]

If $\beta$ is a 2-form, then

\[ (d\beta)(X, Y, Z) = \frac{1}{3} \left\{ (\nabla_X \beta)(Y, Z) + (\nabla_Y \beta)(Z, X) + (\nabla_Z \beta)(X, Y) \right\}. \]
Bibliography


