An Introduction to Riemannian Geometry

We include in these notes a presentation of the basics of differential geometry with a view to Riemannian geometry. We refer the reader to the classics on the subject for a more comprehensive and careful treatment [**2, 3, 4, 5, 7**]. In addition to these texts, our exposition has benefited from the book [**1**] and lecture notes by Ben Andrews.

Mat Langford

Knoxville, December 2019

Last updated September, 2022

Contents

1. Paracompactness, partitions of unity, and manifolds

Recall that a family $\mathcal{A} \doteqdot \{U_\alpha\}_{\alpha \in \mathcal{J}}$ of subsets $U_\alpha \subset X$ of a topological space X is a **cover** of (or **covers**) X if $\cup_{\alpha \in \mathcal{J}} U_\alpha = X$.

Definition 1.1. *A cover $\mathcal{A}$ of a topological space X is*

- ***open*** *if each $A \in \mathcal{A}$ is open.*
- *a **refinement** of a cover $\mathcal{B}$ of X if every element of $\mathcal{A}$ lies in some element of $\mathcal{B}$.*
- ***locally finite*** *if every point of X has a neighborhood which intersects only finitely many elements of $\mathcal{A}$.*

Example 1.2.

- *The cover of a metric space by balls of radius $1/2$ is a refinement of the cover by balls of radius 1. (Neither is locally finite in general.)*
- *The cover of $\mathbb{R}$ by the sets $(n-1, n+1)$ for $n \in \mathbb{N}$ is locally finite.*
- *The cover of the interval $(-1, 1)$ by the sets $(-1/n, 1/n)$ for $n \in \mathbb{N}$ is not locally finite.*

A subcover is trivially a refinement. The converse is not true, of course; however, any finite refinement corresponds to a finite subcover by selecting the sets of the original cover containing those of the refinement. This proves the following reformulation of compactness.

Proposition 1.3. *A topological space is compact if and only if every open cover admits a finite open refinement.*

This reformulation motivates a natural generalization of compactness, whereby the finiteness condition in the preceding proposition is replaced by local finiteness.

Definition 1.4. *A topological space is **paracompact** if every open cover admits a locally finite open refinement.*

Our task now is to convince the reader that this property is both strong enough to efficiently prove useful/intuitive theorems while at the same time weak enough to be satisfied by a large class of interesting spaces.

Example 1.5. *The following topological spaces are paracompact:*

- *Compact spaces.*
- *Discrete spaces: The open cover consisting of all singleton sets is a locally finite open refinement of any open cover. Of course, non-finite discrete spaces are not compact as the cover by singleton sets has no non-trivial subcover.*

- *Euclidean spaces, $\mathbb{R}^n$: Let $\mathcal{A}$ be any open cover of $\mathbb{R}^n$ and consider, for each $i \in \mathbb{N}$, the open annulus $B_{i+1} \setminus \overline{B}_{i-1}$, where $B_i \doteqdot B_i(0)$ is the ball of radius i about the origin. Since the closed annulus $\overline{B}_{i+1} \setminus B_{i-1}$ is compact and contains $B_{i+1} \setminus \overline{B}_{i-1}$, the open cover $\mathcal{A}_i \doteqdot \{U \cap B_{i+1} \setminus \overline{B}_{i-1} : U \in \mathcal{A}\}$ of $B_{i+1} \setminus \overline{B}_{i-1}$ admits a finite subcover $\mathcal{B}_i \doteqdot \{U_1 \cap B_{i+1} \setminus \overline{B}_{i-1}, \dots, U_{N_i} \cap B_{i+1} \setminus \overline{B}_{i-1}\}$, say. We claim that the collection of open sets $\mathcal{B} \doteqdot \cup_{i\in\mathbb{N}}\mathcal{B}_i$ is a locally finite refinement of $\mathcal{A}$: Since the annuli $B_{i+1}\setminus\overline{B}_{i-1}$ cover $\mathbb{R}^n$, $\mathcal{B}$ must also since it is a union of covers of each of these annuli. Furthermore, each element of $\mathcal{B}$ is of the form $U \cap B_{i+1} \setminus \overline{B}_{i-1} \subset U$ for some $U \in \mathcal{A}$; this proves that $\mathcal{B}$ is a refinement of $\mathcal{A}$. To see that it is locally finite, choose for each $x \in \mathbb{R}^n$ an annular neighborhood $B_{i+1} \setminus \overline{B}_{i-1}$; by construction, this annulus intersects only finitely many elements of $\mathcal{B}$.*
- *Every second countable locally compact Hausdorff space is paracompact (cf. parts (3) $\Longrightarrow$ (4) and (4) $\Longrightarrow$ (1) of the proof of Theorem 1.15 below).*
- *The Sorgenfrey line $\mathbb{R}_\ell$ is paracompact and Hausdorff but neither compact, locally compact, nor second countable.*

The proof that $\mathbb{R}^n$ is paracompact has a useful generalization.

Definition 1.6. *An* ***exhaustion*** *of a topological space X is a countable family $\{A_i\}_{i\in\mathbb{N}}$ of subsets $A_i \subset X$ which covers X and satisfies $A_i \subset A_{i+1}$.*

Note that the inclusion condition is the opposite of that of a *nested* sequence of subsets.

Proposition 1.7. *Let X be a Hausdorff space. If X admits an exhaustion $\{K_i\}_{i\in\mathbb{N}}$ by compact sets K_i satisfying*[1] *$K_i \subset \operatorname{int} K_{i+1}$ then it is paracompact.*

Proof. Let X be a Hausdorff space which admits such an exhaustion $\{K_i\}_{i\in\mathbb{N}}$ and let $\mathcal{A}$ be an open cover of X. Set $K_0 \doteqdot \emptyset$ and $B_i \doteqdot \operatorname{int} K_i$. Since X is Hausdorff, compact sets are closed and hence the sets $B_{i+1} \setminus K_{i-1}$ are open. Note that the collection of these 'annuli' covers X. Indeed, for each $x \in X$ there exists some $i \in \mathbb{N}$, and hence some *smallest* $i \in \mathbb{N}$, such that $x \in B_{i+1}$. Then $x \in B_{i+1} \setminus B_i \subset B_{i+1} \setminus K_{i-1}$. Moreover, since each annulus $B_{i+1} \setminus K_{i-1}$ lies in the compact set K_{i+1}, the cover $\mathcal{A}_i \doteqdot \{U \cap B_{i+1} \setminus K_i : U \in \mathcal{A}\}$ of $B_{i+1} \setminus K_i$ admits a finite subcover, $\mathcal{B}_i \doteqdot \{U_1 \cap B_{i+1} \setminus K_i, \dots, U_{N_i} \cap B_{i+1} \setminus K_i\}$, say. The open cover $\mathcal{B} \doteqdot \cup_{i\in\mathbb{N}}\mathcal{B}_i$ of X is a locally finite refinement of $\mathcal{A}$ (cf. the proof that $\mathbb{R}^n$ is paracompact in Example 1.5). $\square$

Example 1.8.

[1] Such an exhaustion is sometimes called a **compact exhaustion**.

- *Paracompact subspaces of paracompact spaces need not be closed: The open interval* $(0,1) \subset \mathbb{R}$ *is paracompact (it is homeomorphic to* $\mathbb{R}$*) but not closed in* $\mathbb{R}$.
- *Subspaces of paracompact spaces need not be paracompact: Consider the space given by equipping the minimal uncountable well-ordered set* S_Ω *with the order topology and let* $\overline{S}_\Omega = S_\Omega \cup \{\Omega\}$ *be its one-point compactification. The product* $\overline{S}_\Omega \times \overline{S}_\Omega$ *is compact and hence paracopompact but the subspace* $S_\Omega \times \overline{S}_\Omega$ *is not paracompact since it is Hausdorff but not normal (see Exercise 1.2).*
- *The product of two paracompact spaces need not be paracompact: The Sorgenfrey plane* $\mathbb{R}_\ell \times \mathbb{R}_\ell$ *Hausdorff but not normal, and hence not paracompact.*

One of the most useful aspects of paracompactness is its relation to partitions of unity.

Definition 1.9. *Let* X *be a topological space. A* ***partition of unity*** *for* X *is a collection* $\{\rho_\alpha\}_{\alpha\in J}$ *of continuous functions* $\rho_\alpha : X \to [0,1]$ *which is*

- ***locally finite****: each point* $x \in X$ *has a neighborhood* U *on which only finitely many* ρ_α *are not identically zero; and*
- ***unitary****:* $\sum_{\alpha\in J} \rho_\alpha(x) = 1$ *for all* $x \in X$.

A partition of unity $\{\rho_\alpha\}_{\alpha\in J}$ *is* ***subordinate to*** *(or* ***dominated by****) an open cover* $\mathcal{A}$ *of* X *if, for each* $\alpha \in J$*, there is some* $U_\alpha \in \mathcal{A}$ *such that* $\operatorname{spt}\rho_\alpha \subset U_\alpha$*, where* $\operatorname{spt}\rho_\alpha \doteqdot \overline{\{x \in X : \rho_\alpha(x) \neq 0\}}$ *is the* ***support*** *of* ρ_α.

Partitions of unity are typically used to 'glue together' locally defined objects to obtain global ones (important examples being the construction of metrics and connections on differentiable manifolds, and the introduction of a well-defined notion of integration and the proof of Stokes' Theorem on Riemannian manifolds). Another important and elegant application concerns the embedding of manifolds in Euclidean spaces (see Theorem 1.17 below).

Proposition 1.10. *Let* X *be a topological space. Suppose that every open cover of* X *admits a subordinate partition of unity. Show that* X *is paracompact.*

Proof. See Exercise 1.3. □

Definition 1.11. *A topological space* X *is* ***locally Euclidean*** *if there is some* $n \in \mathbb{N}$ *such that every point of* X *has a neighborhood which is homeomorphic to* $\mathbb{R}^n$.

Definition 1.12. *A* ***manifold*** *is a locally Euclidean, paracompact Hausdorff space.*

Example 1.13. *The line with two origins is locally Euclidean and paracompact. It is not a manifold, however, since it is not Hausdorff.*

The **dimension** of a manifold is defined to be the dimension n of its Euclidean model space $\mathbb{R}^n$ (it is well-defined by invariance of domain). We will often describe a manifold M of dimension $n \in \mathbb{N}$ as an n-**manifold** and use the notation 'M^n' to indicate that a given manifold M has dimension n.

Proposition 1.14. *Every locally Euclidean Hausdorff space is locally compact and normal.*

Proof. See Exercise 1.2. □

Often, manifolds are defined to be locally Euclidean, **second countable** Hausdorff spaces. The following theorem implies, in particular, that the two definitions are equivalent when the underlying space has countably many connected components.

Theorem 1.15. *Let X be a connected locally Euclidean Hausdorff space. The following statements are equivalent:*

(1) *X is paracompact.*
(2) *X is a union of countably many compact sets.*
(3) *X is second countable.*
(4) *X admits an exhaustion $\{K_i\}_{i\in\mathbb{N}}$ by compact sets K_i satisfying $K_i \subset \operatorname{int} K_{i+1}$.*
(5) *Every open cover of X admits a subordinate partition of unity.*

Proof. (1) $\Longrightarrow$ (2): Suppose that X is paracompact. Let $\{U_\alpha\}_{\alpha\in J}$ be a cover of X by open sets U_α each of which is homeomorphic to $\mathbb{R}^n$ and let $\varphi_\alpha : U_\alpha \to \mathbb{R}^n$ be corresponding homeomorphisms. Consider the refinement $\mathcal{A} \doteqdot \{\varphi_\alpha^{-1}(B_r(x)) : \alpha \in J, r > 0, x \in \mathbb{R}^n\}$ of $\{U_\alpha\}_{\alpha\in J}$ consisting of the preimages $\varphi_\alpha^{-1}(B_r(x))$ of all balls $B_r(x) \subset \mathbb{R}^n$. Observe that each $B \in \mathcal{A}$ has compact closure. Since X is paracompact, we can find a locally finite open refinement $\mathcal{B}$ of $\mathcal{A}$. Note that each $B \in \mathcal{B}$ has compact closure. We will show that $\mathcal{B}$ is countable.

Let us refer to a finite ordered collection $\{B_1, \ldots, B_N\} \subset \mathcal{B}$ as a **string** (joining B_1 and B_N) if $N = 1$ or if $B_{i-1} \cap B_i \neq \emptyset$ for each $i = 2, \ldots, N$. Fix some $B_0 \in \mathcal{B}$.

Claim 1.16. *For every $B \in \mathcal{B}$ there exists a string joining B_0 and B.*

Proof of Claim 1.16. Let $\gamma : [0,1] \to X$ be a path satisfying $\gamma(0) \in B_0$ and $\gamma(1) \in B$. Since $\gamma([0,1])$ is compact, it can be covered by a finite subset $\mathcal{C}$ of $\mathcal{B}$ which includes the elements B_0 and B. We shall call a string $\mathcal{C}' = \{C_0, \dots, C_k\} \subset \mathcal{C}$ a **left** (respectively, **right**) **string** if $C_0 = B_0$ (respectively, $C_0 = B$). Since $\mathcal{C}$ is finite, there are a finite number of left strings and a finite number of right strings. Suppose that no left string contains B. Then no right string contains B_0 as any such string would be a left string containing B. It follows that the union of all left strings is disjoint from the union of all right strings, whence we conclude that $\gamma([0,1])$ is disconnected, a contradiction. □

Consider the function $N : \mathcal{B} \to \mathbb{N}$ which assigns to each $B \in \mathcal{B}$ the length of the shortest string joining B_0 and B. To prove the desired implication, it suffices to show that the set $\mathcal{B}_k \doteqdot N^{-1}(k)$ is finite for each $k \in \mathbb{N}$.

Certainly, $\mathcal{B}_1 = \{B_0\}$ is finite. So suppose, for some $k \in \mathbb{N}$, that $\mathcal{B}_i$ is finite for each $i \leq k$. Then the union

$$K_k \doteqdot \overline{\cup_{i=1}^{k} \cup \mathcal{B}_i} \subset \cup_{i=1}^{k} \cup_{B \in \mathcal{B}_i} \overline{B}$$

is compact. Since $\mathcal{B}$ is locally finite, each point $x \in K_k$ admits a neighborhood U_x which only finitely many $B \in \mathcal{B}$ intersect. The corresponding cover of K_k admits a finite subcover, so K_k is intersected by only finitely many $B \in \mathcal{B}$. But $\mathcal{B}_{k+1} \doteqdot N^{-1}(k+1)$ is a subset of the set of all B which intersect K_k. (Removing B from a minimal $(k+1)$-string connecting B to B_0 leaves a minimal k-string).

(2) $\Longrightarrow$ (3): Suppose that X is covered by a countable collection of compact subsets, $\{K_i\}_{i \in \mathbb{N}}$. Then, as each K_i is covered by finitely many open sets each of which is homeomorphic to $\mathbb{R}^n$, X is covered by countably many open sets each of which is homeomorphic to $\mathbb{R}^n$. Since $\mathbb{R}^n$ is second countable, each of these open sets admits a countable basis. Combining these bases yields a countable basis for X.

(3) $\Longrightarrow$ (4): Suppose that X admits a countable basis, $\mathcal{B} = \{B_i\}_{i=1}^{n}$. Since X is locally compact, we can arrange that each $B_i \in \mathcal{B}$ has compact closure since discarding those that don't still leaves a basis. The desired exhaustion $\{K_i\}_{i \in \mathbb{N}}$ is now constructed recursively via

$$K_1 \doteqdot \overline{B}_1 \quad \text{and}$$
$$K_{k+1} \doteqdot \overline{B_1 \cup \cdots \cup B_{i_k}}\,,$$

where i_k is the smallest integer such that $K_k \subset B_1 \cup \cdots \cup B_{i_k}$.

(4) $\Longrightarrow$ (1): This implication follows from Proposition 1.7.

(1) $\implies$ (5): Suppose that X is paracompact. Let $\{U_\alpha\}_{\alpha\in A}$ be a cover of X by open sets each of which is homeomorphic to $\mathbb{R}^n$ and choose homeomorphisms $\varphi_\alpha : U_\alpha \to \mathbb{R}^n$. Now, since X is normal, it is regular. Recall that this implies that every neighbourhood of every point contains a closed neighbourhood of the point. It follows that any other open cover $\{W_\gamma\}_{\gamma\in C}$ admits a locally finite refinement $\{V_\beta\}_{\beta\in B}$ such that each $\overline{V}_\beta$ lies in some $U_\alpha \cap W_\gamma$. We will use this fact to construct a partition of unity subordinate to $\{W_\gamma\}_{\gamma\in C}$.

Indeed, choose for each $\beta \in B$ some $\alpha \in A$ such that $V_\beta \subset U_\alpha$ and set

$$\hat{\rho}_\beta(x) \doteqdot \begin{cases} d(\varphi_\alpha(x), \mathbb{R}^n \setminus \varphi_\alpha(V_\beta)) & \text{if} \quad x \in \overline{V}_\beta \\ 0 & \text{if} \quad x \in X \setminus V_\beta\,. \end{cases}$$

Then $\operatorname{spt} \hat{\rho}_\beta = \overline{V}_\beta \subset W_\gamma$ for some $\gamma \in C$, the sum $\sum_{\beta\in B} \hat{\rho}_\beta(x)$ is well defined for each $x \in X$ and, by the pasting lemma, $\hat{\rho}_\beta$ is continuous. The desired partition of unity is then defined by setting

$$\rho_\beta(x) \doteqdot \frac{\hat{\rho}_\beta(x)}{\sum_{\beta\in B} \hat{\rho}_\beta(x)}\,.$$

(5) $\implies$ (1): This is the content of Exercise 1.3. □

Note that the equivalence (1) $\iff$ (5) survives the transition to an arbitrary number of connected components, whereas the equivalences (1) $\iff$ (i), $i = 2, 3$, or 4, only remain true when X has a countable number of connected components.

We finish by proving Whitney's "easy" embedding theorem.

Theorem 1.17 (Whitney's "easy" embedding theorem)**.** *Every compact manifold embeds in some Euclidean space.*

Proof. Let X be a compact n-manifold. Then X admits a cover by finitely many open sets $U_1, \dots U_N$, each of which is homeomorphic to $\mathbb{R}^n$. By Theorem 1.15, X admits a partition of unity $\rho_1, \dots, \rho_N$ such that $\operatorname{spt} \rho_i \subset U_i$ for each $i = 1, \dots, N$. Choose homeomorphisms $\psi_i : U_i \to B_1(0)$ for each $i = 1, \dots, N$ and define a map $\Psi : X \to \mathbb{R}^N \times \mathbb{R}^{nN}$ by

$$F(x) \doteqdot (\rho_1(x), \dots, \rho_N(x), \rho_1(x)\psi_1(x), \dots, \rho_N(x)\psi_N(x))\,,$$

where we define $\rho_i(x)\psi_i(x) \doteqdot 0$ if x is not in the domain U_i of ψ_i. By the pasting lemma, F is continuous. We claim that F is an embedding. Since F is a continuous map from a compact space into a Hausdorff space, it suffices to show that it is injective. To this end, suppose that $F(x) = F(y)$ for some pair of points $x, y \in M$. Equating the first N components of $F(x)$ and $F(y)$ yields

$$\rho_i(x) = \rho_i(y) \quad \text{for each} \quad i = 1 \dots, N\,.$$

Since ρ_i is non-negative and sums pointwise to unity, at least one of these numbers, $\rho_k(x) = \rho_k(y)$ say, is non-zero. Equating the final nN components of $F(x)$ and $F(y)$ then implies that $\psi_k(x) = \psi_k(y)$. Since ψ_k is a homeomorphism, we conclude that $x = y$. □

Exercises.

Exercise 1.1. *The Sorgenfrey line $\mathbb{R}_\ell$ is the real line equipped with the topology generated by the right-half open, left-half closed intervals, $[a, b)$. Show that $\mathbb{R}_\ell$ is paracompact.*

Exercise 1.2. *Prove the following propositions:*

(a) *Every paracompact Hausdorff space is normal.*
(b) *Every closed subspace of a paracompact space is paracompact.*
(c) *A locally connected topological space is paracompact if and only if all of its connected components are paracompact.*

Exercise 1.3. *Let X be a topological space. Suppose that every open cover of X admits a subordinate partition of unity. Show that X is paracompact.*

Exercise 1.4. *Show that a topological space is locally Euclidean if and only if each point admits a neighbourhood which is homeomorphic to an open subset of Euclidean space.*

Exercise 1.5. *Let X be a locally Euclidean Hausdorff space. Prove the following propositions:*

(a) *X is locally compact.*
(b) *If X is connected then X is path connected.*
(c) *Given $x \in X$ choose a neighborhood U_x and a homeomorphism $\varphi_x : U_x \to \mathbb{R}^n$. The collection*
$$\mathcal{B} \doteqdot \{\varphi_x^{-1}(B_r(z)) : x \in X, r > 0, z \in \mathbb{R}^n\}$$
is a basis for X.

Exercise 1.6. *Let X be a locally Euclidean Hausdorff space, let $\{U_\alpha\}_{\alpha\in J}$ be a cover of X by open sets U_α each of which is homeomorphic to $\mathbb{R}^n$ and let $\varphi_\alpha : U_\alpha \to \mathbb{R}^n$ be corresponding homeomorphisms. Consider the refinement $\mathcal{A} \doteqdot \{\varphi_\alpha^{-1}(B_r(x)) : \alpha \in J, r > 0, x \in \mathbb{R}^n\}$ of $\{U_\alpha\}_{\alpha\in J}$ consisting of the preimages $\varphi_\alpha^{-1}(B_r(x))$ of all open balls $B_r(x) \subset \mathbb{R}^n$.*

(a) *Show that each $B \in \mathcal{A}$ has compact closure contained in some U_α.*

Suppose further that X is paracompact and let $\mathcal{B}$ be a locally finite refinement of $\mathcal{A}$.

(b) *Show that each $B \in \mathcal{B}$ has compact closure contained in some U_α.*

Exercise 1.7. *Let X be a paracompact Hausdorff space.*

(a) *Show that every open cover $\mathcal{U}$ of X admits a locally finite open refinement $\mathcal{V}$ such that each $V \in \mathcal{V}$ satisfies $\overline{V} \subset U$ for some $U \in \mathcal{U}$* HINT*: Use the fact that X is regular.*

(b) *Deduce using Urysohn's Lemma that every open cover of X admits a subordinate partition of unity.*

2. Differentiable manifolds

Recall that a topological space is an m-**manifold** if it is paracompact, Hausdorff and locally homeomorphic to $\mathbb{R}^m$.

Definition 2.1 (Charts and atlases)**.** *Let M^m be a (topological) m-manifold. A **chart** for M^m is an embedding $\varphi : U \hookrightarrow \mathbb{R}^m$, where U is an open subset of M^m. An **atlas** for M^m is a collection $\{\varphi_\alpha : U_\alpha \to \mathbb{R}^m\}_{\alpha \in J}$ of charts $\varphi_\alpha : U_\alpha \to \mathbb{R}^m$ such that $M^m = \cup_{\alpha \in J} U_\alpha$. The maps $\varphi_\alpha \circ \varphi_\beta^{-1}|_{\varphi_\beta(U_\alpha \cap U_\beta)} : \varphi_\beta(U_\alpha \cap U_\beta) \to \varphi_\alpha(U_\alpha \cap U_\beta)$ are called **transition maps**.*

We will usually describe charts $\varphi : U \to \mathbb{R}^m$ using their component functions $\varphi = (x^1, \ldots, x^m) : U \to \mathbb{R}^m$.

Recall that a function from an open subset U of $\mathbb{R}^m$ to $\mathbb{R}$ is **smooth**[2] if it admits all partial derivatives to all orders at all points of U, while a map from an open subset of $\mathbb{R}^m$ to $\mathbb{R}^n$ is **smooth** if each of its component functions is smooth.

Definition 2.2 (Differentiable structures)**.** *Let M^m be an m-manifold. An atlas $\mathcal{A} = \{\varphi_\alpha : U_\alpha \to \mathbb{R}^m\}_{\alpha \in J}$ for M^m is **differentiable** if each of its transition maps is smooth.*

*Two differentiable atlases $\mathcal{A}$ and $\mathcal{B}$ are **compatible** if their union is also a differentiable atlas (equivalently, if for every chart $\varphi : U \to \mathbb{R}^m$ in $\mathcal{A}$ and every $\eta : V \to \mathbb{R}^m$ in $\mathcal{B}$ the maps $\varphi \circ \eta^{-1}|_{\eta(U \cap V)}$ and $\eta \circ \varphi^{-1}|_{\varphi(U \cap V)}$ are smooth). Given a differentiable atlas $\mathcal{A}$, the **maximal atlas** containing $\mathcal{A}$ is the union of all atlases compatible with $\mathcal{A}$.*

*A **differentiable structure** for M^m is given by equipping it with a maximal atlas, $\mathcal{A}$ say. The resulting pair $(M^m, \mathcal{A})$ is called a **differentiable manifold** (usually simply denoted by M^m).*

We will also refer to differentiable manifolds as **smooth manifolds**.

Example 2.3. *The following provide examples of differentiable manifolds.*

(1) *The m-dimensional Euclidean space $\mathbb{R}^m$ equipped with the differentiable structure induced by the atlas consisting of the identity map.*
(2) *The m-dimensional sphere $S^m \doteqdot \{x \in \mathbb{R}^{m+1} : |x| = 1\}$ equipped with the differentiable structure induced by the atlas consisting of stereographic projections about two antipodal points.*[3]

[2] It is possible to generalize most, if not all, of the notions in these notes regarding smooth structures to include structures carrying weaker differentiability properties, such as C^k or $C^{k,\alpha}$.

[3] Stereographic projection $\varphi : S^m \to \mathbb{R}^m$ about a point $o \in S^m$ maps a point $p \in S^m \setminus \{o\}$ to the point $\varphi(p) \doteqdot \mathrm{ray}(p) \cap (\mathbb{R}^m \times \{0\}) \subset \mathbb{R}^m \times \{0\} \cong \mathbb{R}^m$, where $\mathrm{ray}(p)$ is the ray through p emanating from o.

(3) *The m-dimensional sphere S^m equipped with the differentiable structure induced by the atlas consisting of the $2(m+1)$ projections*

$$\pi_i^{\pm} : \{(x_1, \dots, x_{m+1}) \in S^m : \pm x_i > 0\} \to \mathbb{R}^m$$

defined by

$$\pi_i^{\pm}(x_1, \dots, x_{m+1}) \doteqdot (x_1, \dots, x_{\hat{i}}, \dots, x_{m+1}) \,,$$

where the hat on the index indicates that this term is not present.

(4) *The m-dimensional real projective space (the space of lines through the origin)*

$$\mathbb{R}P^m \doteqdot (\mathbb{R}^{m+1} \setminus \{0\})/\sim \,, \quad \textit{where } p \sim q \iff p = \lambda q \textit{ for some } \lambda \neq 0 \,,$$

equipped with the differentiable structure induced by the atlas consisting of the $(m+1)$ charts

$$\varphi_i : \{[(x_1, \dots, x_{m+1})] : x_i \neq 0\} \to \mathbb{R}^m$$

defined by

$$\varphi_i([(x_1, \dots, x_{m+1})]) \doteqdot \left(\frac{x_1}{x_i}, \dots, \frac{x_{\hat{i}}}{x_i}, \dots, \frac{x_{m+1}}{x_i}\right) \,,$$

where the hat again indicates that this term is not present.

(5) *The m-dimensional torus $T^m = \underbrace{S^1 \times \cdots \times S^1}_{m\text{-times}}$ equipped with the differentiable structure induced by the product atlas.*

(6) *Open subsets of differentiable manifolds equipped with the restriction atlas.*

(7) *Cartesian products of differentiable manifolds equipped with the product atlas.*

(8) *The general linear group*

$$\mathrm{GL}(\mathbb{R}^n) \doteqdot \{A \in n \times n\textit{-matrices} : \det A \neq 0\}$$

is a smooth manifold (it is an open subset of $\mathbb{R}^{n \times n}$).

Definition 2.4 (Smooth maps, functions and curves)**.** *A function $f : M \to \mathbb{R}$ from a differentiable manifold M to $\mathbb{R}$ is* ***differentiable*** *if given some chart (U, φ) for M the composition $f \circ \varphi^{-1} : \varphi(U) \to \mathbb{R}$ is smooth.*

A map $f : M^m \to N^n$ from a differentiable manifold M^m to a differentiable manifold N^n is ***differentiable*** *if given any charts $\varphi : U \to \mathbb{R}^m$ for M^m and $\eta : V \to \mathbb{R}^n$ for N^n the composition $\eta \circ f \circ \varphi^{-1} : \varphi(U) \to \eta(V)$ is smooth.*

A differentiable map $f : M \to N$ is a ***diffeomorphism*** *if it admits a differentiable inverse $f^{-1} : N \to M$. If such a map exists, M and N are said to be* ***diffeomorphic****.*

We shall usually refer to differentiable functions and differentiable maps as **smooth functions** and **smooth maps**, respectively.

Note that, in order to check that a function $f : M^m \to \mathbb{R}$ is smooth, we need not check that $f \circ \varphi^{-1}$ is smooth for *every* chart $\varphi : U \to \mathbb{R}^m$ for M^m (which is clearly impractical) — we are free to choose any particular compatible atlas $\{\varphi_\alpha\}_{\alpha\in A}$ and check that $f\circ\varphi_\alpha$ is smooth for each α. Indeed, if $\varphi : U \to \mathbb{R}^m$ and $\eta : V \to \mathbb{R}^m$ are two charts for M^m with nontrivial overlap $U \cap V$ then

$$f \circ \eta^{-1}|_{\eta(U\cap V)} = (f \circ \varphi^{-1}) \circ (\varphi \circ \eta^{-1})|_{\eta(U\cap V)} .$$

Since $\varphi \circ \eta^{-1}|_{\eta(U\cap V)}$ is smooth, smoothness of $f \circ \eta^{-1}|_{\eta(U\cap V)}$ follows from smoothness of $f \circ \varphi^{-1}|_{\varphi(U\cap V)}$. Similarly, in order to check that a map $f : M^m \to N^n$ smooth, it suffices to check smoothness with respect to any compatible atlas $\{\varphi_\alpha\}_{\alpha\in A}$ for M^m and any family of compatible charts $\{\eta_\beta\}_{\beta\in B}$ for N^n covering $f(M^m)$.

2.1. Differentiable manifolds-with-boundary. An n-**manifold-with-boundary** is a paracompact Hausdorff space M^n each of whose points admits a neighborhood locally homeomorphic to either $\mathbb{R}^n$ or the closed half-space $\mathbb{R}^n_+ \doteqdot \mathbb{R}^n \cap \{p \in \mathbb{R}^n : p \cdot e_n \geq 0\}$. If a point p admits a neighborhood homeomorphic to $\mathbb{R}^n$ then p is an **interior point**. Else, p is a **boundary point**. The set of boundary points is called the **boundary** of M^n and denoted ∂M^n. A **chart** for M^n is a pair (U, φ), where U is an open subset of M^n and $\varphi : U \to \varphi(U) \subset \mathbb{R}^n_+$ is a homeomorphism. If $\varphi(U)$ does not lie in the interior of $\mathbb{R}^n_+$ then φ is called a **boundary chart**.

Recall that a map from an arbitrary subset $A \subset \mathbb{R}^m$ to $\mathbb{R}^n$ is deemed to be **smooth** if it extends to a smooth map on an open subset of $\mathbb{R}^n$ containing A.

(**Differentiable**) **Atlases**, **transition maps**, **differentiable structures**, **differentiable manifolds-with-boundary** and **differentiable maps** between them may now be defined in the obvious way.

Example 2.5. *The following provide examples of differentiable manifolds-with-boundary.*

(1) *All differentiable manifolds are differentiable manifolds-with-boundary (albeit with empty boundary).*
(2) *The halfspace $\mathbb{R}^n_+$ equipped with the atlas induced by the identity chart is a differentiable manifold-with-boundary. Its boundary is $\partial\mathbb{R}^n_+ = \mathbb{R}^{n-1} \times \{0\}$.*
(3) *The closed unit ball $\overline{B}^n$ and the compliment $\mathbb{R}^n \setminus B^n$ of the open unit ball B^n can be equipped with atlases making them smooth manifolds-with-boundary, with boundary $\partial B^n = S^{n-1}$ (see Exercise 2.9).*

2.2. Additional notes. When $n \leq 3$, every (topological) manifold admits a unique (up to diffeomorphism) smooth structure. This is an "easy" theorem when $n = 1$ but much harder when $n = 2$ (a theorem of Radó) or $n = 3$ (a theorem of Moise).

If $n \geq 5$, Euclidean space $\mathbb{R}^n$ admits only the standard smooth structure (a theorem of Stallings). In extreme contrast to these results, Clifford Taubes (building on work of Mike Freedman and Simon Donaldson) showed that there exists a *continuum* of smooth structures on $\mathbb{R}^4$!

The case of spheres is even more interesting. For example, it is known that there are in general many "exotic" smooth structures (i.e. structures not diffeomorphic to the standard one) on high dimensional spheres and exactly 28 non-diffeomorphic smooth structures on the 7-sphere (by work of Michel Kervaire and John Milnor). Moreover, it remains unknown (as of the time of writing) whether or not there exist exotic smooth structures on S^4.

By work of Mike Freedman, the diffeomorphism problem is closely related to the four dimensional "smooth Poincaré conjecture", which asserts that the sphere with its standard differentiable structure is the only smooth homotopy sphere. Indeed, Freedman proved that the conjecture is true, for the four sphere, in the topological category. The smooth version just stated remains open. The conjecture is settled in both categories in all other dimensions due work of Poincaré (dimension two), Stephen Smale (dimensions five and higher) and Grisha Perelman (dimension three).

Exercises.

Exercise 2.1. *Let $\mathcal{A}$ be a differentiable atlas for a manifold M. Show that the maximal atlas containing $\mathcal{A}$ exists and is unique.*

Exercise 2.2. *Prove that each of the atlases in Example 2.3 does indeed induce a differentiable structure.*

Exercise 2.3. *Show that the two atlases for S^m defined in Example 2.3 are compatible, and hence induce the same differentiable structure.*

Exercise 2.4. *Consider the atlas $\mathcal{B} \doteqdot \{x \mapsto x^{29}\}$ for $\mathbb{R}$.*

(a) *Show that $\mathcal{B}$ is a differentiable atlas.*
(b) *Show that $\mathcal{B}$ is* not *compatible with the standard atlas $\mathcal{A} \doteqdot \{x \mapsto x\}$.*
(c) *Find a diffeomorphism between $(\mathbb{R}, \mathcal{A})$ and $(\mathbb{R}, \mathcal{B})$.*

Exercise 2.5. *Show that a map $F : M \to N$ between smooth manifolds M and N is smooth if and only if the map $f \circ F : M \to \mathbb{R}$ is smooth for every smooth $f : N \to \mathbb{R}$.*

Exercise 2.6. *The complex projective space $\mathbb{C}P^n$ is the space of complex lines in $\mathbb{C}^{n+1}$; that is, $\mathbb{C}P^n \doteqdot (\mathbb{C}^{n+1} \setminus \{0\})/\sim$, where*

$$x \sim y \quad \Longleftrightarrow \quad y = \lambda x \ \text{ for some } \ \lambda \in \mathbb{C} \setminus \{0\}\,.$$

Find a differentiable atlas which makes $\mathbb{C}P^n$ a $2n$-dimensional differentiable manifold.

Exercise 2.7. *Show that the multiplication and inversion maps*

$$\begin{aligned} \mathrm{GL}(\mathbb{R}^n) \times \mathrm{GL}(\mathbb{R}^n) &\to \mathrm{GL}(\mathbb{R}^n) \\ (A, B) &\mapsto A \cdot B \end{aligned}$$

and

$$\begin{aligned} \mathrm{GL}(\mathbb{R}^n) &\to \mathrm{GL}(\mathbb{R}^n) \\ A &\mapsto A^{-1}, \end{aligned}$$

respectively, are smooth.

A differentiable manifold equipped with a group structure is called a ***Lie group*** *if its multiplication and inversion maps are smooth.*

Exercise 2.8. *Let M^n be a manifold-with-boundary. Show that ∂M^n is an $(n-1)$-manifold. If M^n is smooth, show that the boundary charts induce a smooth atlas for ∂M^n.*

Exercise 2.9. *Find atlases which make $\overline{B}^n$ and $\mathbb{R}^n \setminus B^n$, respectively, differentiable manifolds-with-boundary.*

3. The tangent space and tangent maps

There are three natural notions of tangency for a smooth manifold M^n. The first, and most abstract, definition views tangency through charts, using the intuitive identification of points and vectors in $\mathbb{R}^n$.

Definition 3.1. *Let M be a smooth manifold(-with-boundary) and p a point of M. The **tangent space** (chart version) to M at p is the vector space T_pM consisting of equivalence classes $[(\varphi, u)]$ of pairs (φ, u) of charts $\varphi : U \to \mathbb{R}^n$ for M with $p \in U$ and vectors $u \in \mathbb{R}^n$ under the equivalence relation*

$$(\varphi, u) \sim (\eta, v) \quad \Longleftrightarrow \quad d(\eta \circ \varphi^{-1})_{\varphi(p)} u = v \,,$$

equipped with the linear structure defined by

$$[(\varphi, u)] + \lambda[(\eta, v)] \doteqdot [(\eta, d(\eta \circ \varphi^{-1})_{\varphi(p)} u + \lambda v)], \quad \lambda \in \mathbb{R} \,.$$

The basic idea is this: we think of a vector as being an “arrow” telling us which direction we are “pointing” inside the manifold. This information is encoded by viewing the direction through a chart — i.e. seeing which way we point “downstairs” in the chart. The equivalence relation allows us to pull this information back “upstairs” to the manifold, by removing the ambiguity of the choice of chart through which to determine the direction.

Another way to think about it is the following: we have a local description for M using charts, and we know what a vector is “downstairs” in each chart. We want to define a space of vectors T_pM “upstairs” in such a way that the derivative map $d\varphi_p$ of the chart map φ makes sense as a linear operator between the vector spaces T_pM and $\mathbb{R}^n$, and so that the chain rule continues to hold. Now, given two distinct charts $\varphi : U \to \mathbb{R}^n$ and $\eta : V \to \mathbb{R}^n$ about p, we would have for any vector $u \in T_pM$ vectors $v = d\varphi_p u \in \mathbb{R}^n$ and $w = d\eta_p u \in \mathbb{R}^n$. Now, if the chain rule is to hold, then we would have

$$d\varphi^{-1}_{\varphi(p)} v = d\eta^{-1}_{\eta(p)} w$$

(since both are equal to u) and this would hold only if

$$d(\eta \circ \varphi^{-1})_{\varphi(p)} v = w \,.$$

Observe that for any chart $\varphi : U \to \mathbb{R}^n$ about p we may define a map $d\varphi_p : T_pM \to \mathbb{R}^n$ by

$$d\varphi_p[(\varphi, u)] \doteqdot u \,.$$

This map is a linear isomphism and satisfies (formally) the chain rule[4].

[4] This will agree with our more general definition of derivatives of maps between manifolds below.

Observe also that each chart $\varphi : U \to \mathbb{R}^n$ about a given point p gives rise to a natural basis $\{e_i|_p\}_{i=1}^n$ for T_pM defined by

$$e_i|_p \doteqdot [(\varphi, e_i)]\,,$$

where $\{e_i\}_{i=1}^n$ is the canonical basis for $\mathbb{R}^n$. Equivalently,

$$e_i|_p \doteqdot d\varphi^{-1}_{\varphi(p)} e_i\,,$$

where $d\varphi^{-1}_{\varphi(p)} \doteqdot (d\varphi_p)^{-1}$.

The second definition expresses tangent vectors from the point of view of "velocities" of motions of "point particles" through the manifold.

Definition 3.2. *Let M be a smooth manifold(-with-boundary) and p a point of M. The **tangent space** (path version) to M at p is the vector space P_pM consisting of equivalence classes $[\gamma]$ of smooth curves $\gamma : I \to M$ such that $0 \in I$ and $\gamma(0) = p$ under the equivalence relation*

$$\gamma \sim \sigma \quad \Longleftrightarrow \quad (\varphi\circ\gamma)'(0) = (\varphi\circ\sigma)'(0)$$

for some (and hence any) chart $\varphi : U \to \mathbb{R}^n$ with $p \in U$, equipped with the linear structure defined by

$$[\gamma] + \lambda[\sigma] \doteqdot \left[s \mapsto \varphi^{-1}\left(\varphi(p) + s\left[(\varphi\circ\gamma)'(0) + \lambda(\varphi\circ\sigma)'(0)\right]\right)\right],\quad \lambda \in \mathbb{R}$$

for some (and hence any) chart $\varphi : U \to \mathbb{R}^n$ with $p \in U$.

In this definition, we think of tangent vectors as corresponding to the instantaneous velocities of point particles travelling along curves in the manifold. This information is once again encoded by viewing the velocities downstairs through a chart, and then pulling back upstairs via the obvious equivalence relation.

Observe that each chart $\varphi : U \to \mathbb{R}^n$ about a given point p gives rise to a natural basis $\{\dot{x}_i|_p\}_{i=1}^n$ for P_pM defined by

$$\dot{x}_i|_p \doteqdot [s \mapsto \varphi^{-1}(\varphi(p) + se_i)]\,,$$

where $\{e_i\}_{i=1}^n$ is the canonical basis for $\mathbb{R}^n$.

The final definition of tangency expresses the notion of a tangent vector as a (directional) derivative operator on functions.

Definition 3.3. *Let M be a smooth manifold(-with-boundary) and p a point of M. The **tangent space** (derivation version) to M at p is the vector space D_pM of smooth derivations at p. That is, the space of $\mathbb{R}$-linear maps $u : C^\infty(M) \to \mathbb{R}$ satisfying the Leibniz rule*

$$u(fg) = (uf)g(p) + f(p)(ug)\,,\quad f, g \in C(M)\,,$$

equipped with the linear structure defined by

$$(u + \lambda v)f \doteqdot uf + \lambda(vf)\,.$$

Observe that each chart $\varphi : U \to \mathbb{R}^n$ about a given point p gives rise to a natural basis $\{\partial_i|_p\}_{i=1}^n$ for D_pM defined by

$$\partial_i|_p \doteqdot f \mapsto d(f \circ \varphi^{-1})_{\varphi(p)} e_i \,,$$

where $\{e_i\}_{i=1}^n$ is the canonical basis for $\mathbb{R}^n$.

Theorem 3.4. *The maps $\alpha : T_pM \to P_pM$, $\beta : P_pM \to D_pM$ and $\delta : D_pM \to T_pM$ defined by*

$$\alpha([(\varphi, u)]) \doteqdot [s \mapsto \varphi^{-1}(\varphi(p) + su)] \,,$$

$$\beta([\gamma])f \doteqdot \frac{d}{ds}\Big|_{s=0} (f \circ \gamma)$$

and[5]

$$\delta(u) \doteqdot [(\varphi, (u\varphi^1, \ldots, u\varphi^n))] \,,$$

for some (and hence any) chart $\varphi : U \to \mathbb{R}^n$ with $p \in U$, are isomorphisms.

Proof. It suffices to show that the triple composition

$$T_pM \underset{\alpha}{\to} P_pM \underset{\beta}{\to} D_pM \underset{\delta}{\to} T_pM$$

is the identity map. This is a straightforward consequence of the definitions:

$$\begin{aligned}
(\delta \circ \beta \circ \alpha)([(\varphi, u)]) &= (\delta \circ \beta)([s \mapsto \varphi^{-1}(\varphi(p) + su)]) \\
&= \delta\left(\left[f \mapsto \frac{d}{ds}\Big|_{s=0} f(\varphi^{-1}(\varphi(p) + su))\right]\right) \\
&= \delta\left(\left[f \mapsto d(f \circ \varphi^{-1})_{\varphi(p)} \cdot u\right]\right) \\
&= \left[\left(\varphi, \sum_{i=1}^n \left(d(\varphi^i \circ \varphi^{-1})_{\varphi(p)} \cdot u\right) e_i\right)\right] \\
&= [(\varphi, u)] \,.
\end{aligned}$$

This completes the proof; however, since it will prove illustrative, we shall also reach the claim through the remaining two possible triple compositions.

For the composition

$$P_pM \underset{\beta}{\to} D_pM \underset{\delta}{\to} T_pM \underset{\alpha}{\to} P_pM \,,$$

[5] In order to ensure that δ is well defined, we need to extend $\varphi : U \to \mathbb{R}^n$ smoothly to M (not necessarily as a chart) in such a way that δ is independent of the extension. This can be achieved by multiplying φ with a smooth cut-off function which vanishes outside of a neighborhood $W \Subset U$ of p and is 1 in a neighborhood $V \Subset W$ of p and then making φ zero outside of U. We leave it to the reader to check that such a cut-off function exists and that $u\varphi^i$ is independent of the resulting extension.

we have

$$\begin{aligned}(\alpha\circ\delta\circ\beta)([\gamma]) &= (\alpha\circ\delta)\left(\left[f\mapsto (f\circ\gamma)'(0)\right]\right)\\ &= \alpha\left(\left[\left(\varphi, \sum_{i=1}^{n}\left(\varphi^i\circ\gamma\right)'(0)e_i\right)\right]\right)\\ &= \left[s\mapsto \varphi^{-1}\left(\varphi(p)+s\left(\varphi\circ\gamma\right)'(0)\right)\right]\\ &= [\gamma]\ .\end{aligned}$$

For the composition

$$D_pM \underset{\delta}{\to} T_pM \underset{\alpha}{\to} P_pM \underset{\beta}{\to} D_pM\ ,$$

we proceed as follows:

$$\begin{aligned}(\beta\circ\alpha\circ\delta)(u) &= (\beta\circ\alpha)\left(\left[\left(\varphi, \sum_{i=1}^{n}(u\varphi^i)e_i\right)\right]\right)\\ &= \beta\left(\left[s\mapsto\varphi^{-1}\left(\varphi(p)+s\sum_{i=1}^{n}(u\varphi^i)e_i\right)\right]\right)\\ &= \left[f\mapsto \frac{d}{ds}\Big|_{s=0} f\circ\varphi^{-1}\left(\varphi(p)+s\sum_{i=1}^{n}(u\varphi^i)e_i\right)\right]\\ &= \left[f\mapsto\sum_{i=1}^{n}(u\varphi^i)d(f\circ\varphi^{-1})_{\varphi(p)}\cdot e_i\right]\ .\end{aligned}$$

We claim that this is the same as u. indeed, by Taylor's theorem (applied to $f\circ\varphi^{-1}$),

$$f(q) = f(p) + \sum_{i=1}^{n} G_i(q)(\varphi^i(q)-\varphi^i(p))$$

for some smooth functions G_i with $G_i(p) = d(f\circ\varphi^{-1})_{\varphi(p)}\cdot e_i$. Thus,

$$uf = \sum_{i=1}^{n} G_i(p)u\varphi^i = \sum_{i=1}^{n} u\varphi^i d(f\circ\varphi^{-1})_{\varphi(p)}\cdot e_i\ . \qquad \square$$

From now on, we will denote the **tangent space** to M at p by T_pM, implicitly identifying each of the three versions using the natural isomorphisms.

Modulo the natural isomorphisms, $\{\partial_i|_p\}_{i=1}^n$ is called the **coordinate basis** for T_pM corresponding to the chart $\varphi : U\to\mathbb{R}^n$.

Note that we may identify the tangent spaces $T_p\mathbb{R}^n$ of $\mathbb{R}^n$ with $\mathbb{R}^n$ using the identity chart by identifying a vector $v=\sum_{i=1}^n v^i\partial_i|_p\in T_p\mathbb{R}^n$ with the point $\sum_{i=1}^n v^ie_i\in\mathbb{R}^n$.

Recall that the classical differential df_p at $p \in \mathbb{R}^n$ of a smooth map $f : U \underset{\text{open}}{\subset} \mathbb{R}^n \to \mathbb{R}^m$ is the linear map from $\mathbb{R}^n \to \mathbb{R}^m$ given by

$$df_p(v) \doteqdot \sum_{i=1}^{n} \sum_{j=1}^{m} u^i \left.\frac{\partial f^j}{\partial x^i}\right|_p e_j \,.$$

Definition 3.5. *The* ***differential*** *(a.k.a., the* ***tangent map*** *or* ***push-forward****) at $p \in M^m$ of a smooth map $f : M^m \to N^n$ is the linear map $df_p : T_pM^m \to T_{f(p)}N^n$ defined, with respect to some (and hence any) pair of charts $\varphi : U \to \mathbb{R}^m$ about p and $\eta : V \to \mathbb{R}^n$ about $F(p)$, by*

$$df_p[(\varphi, u)] = [(\eta, d(\eta \circ f \circ \varphi^{-1})_{\varphi(p)} u)] \,.$$

The ***rank*** *of f at p is the rank of df_p, which is defined as*

$$\operatorname{rank}(df_p) \doteqdot \operatorname{rank}\left(d(\eta \circ f \circ \varphi^{-1})_{\varphi(p)}\right)$$

for some (and hence any) pair of charts $\varphi : U \to \mathbb{R}^m$ about p and $\eta : V \to \mathbb{R}^n$ about $f(p)$. The ***nullity*** *of f at p is defined as* $\operatorname{null}(df_p) \doteqdot m - \operatorname{rank}(df_p)$.

In terms of paths, df_p is given by

$$df_p[\gamma] = [f \circ \gamma] \,,$$

while, in terms of derivations, it is given by

$$(df_p \cdot u)g = u(g \circ f) \,.$$

Observe that (upon identifying $T_p\mathbb{R}^n$ with $\mathbb{R}^n$) the differential of a smooth map $f : \mathbb{R}^n \to \mathbb{R}^m$ is just the usual differential. We also observe that (identifying $T_r\mathbb{R}$ with $\mathbb{R}$) the differentials dx^j_p of the component functions of any chart $\varphi = (x^1, \ldots, x^n) : U \to \mathbb{R}^n$ for M satisfy

$$dx^j_p(\partial_i|_p) = \delta^j_i \doteqdot \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \,. \end{cases}$$

Definition 3.6. *A smooth map $f : M \to N$ between smooth manifolds(-with-boundary) M and N is called a* ***local diffeomorphism/immersion/submersion****, respectively, if its linearization $df_p : T_pM \to T_{f(p)}N$ is an isomorphism/monomorphism/epimorphism, respectively, for all $p \in M$. The map f is called an* ***embedding*** *if it is a diffeomorphism onto its image. An immersion $\gamma : I \to M$ from an interval $I \subset \mathbb{R}$ into a smooth manifold M is called a* ***regular curve*** *in M.*

Example 3.7. *Projection mappings $\pi_M : M \times N \to M$ are submersions. Inclusion mappings $\iota_M : M \to M \times N$ are embeddings.*

Definition 3.8. *A subset M of a smooth n-manifold-with-boundary N is a smooth m-**submanifold**(**-with-boundary**) if for every point p in M there exists a neighborhood V of p in N, an open set $U \subset \mathbb{R}^m$ ($\mathbb{R}^m_+$) and a smooth map $\xi : U \to N$ such that ξ is a homeomorphism onto $M \cap V$ and $d\xi|_q$ is injective for every $q \in U$.*

Definition 3.8 characterizes a submanifold as a subset which may be written locally as the image of an embedding. The following proposition provides useful equivalent definitions of a submanifold. Its proof is an application of the classical inverse function theorem, which we now recall.

Theorem 3.9 (Classical inverse function theorem)**.** *Let f be a smooth function from an open neighbourhood of $x \in \mathbb{R}^n$ to $\mathbb{R}^n$. If the derivative $df_x : \mathbb{R}^n \to \mathbb{R}^n$ at x is an isomorphism, then there exist neighbourhoods U of x and V of $f(x)$ such that $f|_U$ is a diffeomorphism from U to V.*

Proposition 3.10. *Let M be a subset of a smooth manifold N. The following are equivalent:*

(a) *M is an m-dimensional submanifold.*

(b) *M is an m-dimensional manifold and admits a differentiable structure with respect to which the inclusion $\iota : M \to N$ is an embedding.*

(c) *For every $p \in M$ there exists an open set $V \subset N$ containing p, an open set $W \subset \mathbb{R}^n$ and a diffeomorphism $F : V \to W$ such that $F(M \cap V) = W \cap (\mathbb{R}^m \times \{0\})$.*

(d) *For every point $p \in M$ and every chart $y : W \to \mathbb{R}^n$ for N about p, there exists an open set $V \subset W$ containing p, an open set $U \subset \mathbb{R}^m$, a smooth map $f : U \to \mathbb{R}^{n-m}$ and a permutation $\sigma \in S_n$ such that*

$$M \cap V = \left\{p \in V : \left(y^{\sigma(m+1)}(p), \dots, y^{\sigma(n)}(p)\right) = f\left(y^{\sigma(1)}(p), \dots, y^{\sigma(m)}(p)\right)\right\}.$$

(e) *For every $p \in M$ there exists an open subset V of N containing p and a submersion $\pi : V \to \mathbb{R}^{n-m}$ such that $M \cap V = \pi^{-1}(0)$.*

Proof. We will not prove the implications "(b) implies (a)", "(c) implies (a)", "(d) implies (c)" or "(c) implies (e)", as these are more or less immediate. It then suffices to establish that "(a) implies (b)", "(a) implies (d)" and "(e) implies (d)".

We first establish that *(a) implies (b)*. If M is an m-dimensional submanifold of N, then, given $p \in M$, we can find an open subset U of $\mathbb{R}^m$, a neighbourhood V of p in N, and a smooth map $\xi : U \to N$ which is a homeomorphism onto $M \cap V$ and has injective derivative. Without loss of generality, the closure of V is contained in the domain of a chart ψ for N (since otherwise we may simply intersect it with an open set whose closure lies in the domain of a chart ψ of our choice, and then replace U by the preimage of the intersection). If we denote by $\tilde{V}$ the image of V under ψ and by $\tilde{M}$ the image of M under ψ, then the composition $\tilde{\xi} \doteqdot \psi \circ \xi : U \to V$ is a homeomorphism of U onto $\tilde{M} \cap V$ with injective derivative.

Set $y \doteqdot \psi(p)$ and $x \doteqdot \xi^{-1}(p)$. Since $d\tilde{\xi}_x$ is injective, we can choose a bijection $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ such that rows $\sigma(1)$ to $\sigma(m)$ of $d\tilde{\xi}_p$ are linearly independent. Thus, if we define the projection $\pi : \mathbb{R}^n \to \mathbb{R}^m$ by $\pi(z_1, \ldots, z_n) \doteqdot (z_{\sigma(1)}, \ldots, z_{\sigma(m)})$, then $d(\pi \circ \tilde{\xi})_x$ is an isomorphism, and hence the inverse function theorem provides open sets $A \subset U$ about x and $B \subset \pi(\tilde{V}) \subset \mathbb{R}^m$ about $\pi(y)$ and a smooth inverse $\eta : B \to A$ of $\pi \circ \tilde{\xi}|_A$.

We claim that the set $\mathcal{A}$ consisting of the maps

$$\varphi \doteqdot \pi \circ \psi|_W : W \to B\,, \quad \text{where } \; W \doteqdot \xi(A)$$

thus defined for each $p \in M$ is a differentiable atlas for M. Indeed, since $p = \xi(x) \in \xi(A)$, the domains certainly cover M. Moreover, each map φ is the inverse of the homeomorphism $\xi \circ \eta$. Finally, for any pair of points p and q in M, the corresponding transition map

$$\begin{aligned} \varphi_p \circ \varphi_q^{-1}|_{W_p \cap W_q} &= \pi_p \circ \psi_p \circ \xi_q \circ \eta_q|_{W_p \cap W_q} \\ &= \pi_p \circ (\psi_p \circ \psi_q^{-1}) \circ (\psi_q \circ \xi_q) \circ \eta_q|_{W_p \cap W_q} \end{aligned}$$

is smooth whenever the overlap is nontrivial. The corresponding maximal atlas therefore provides a differentiable structure for M.

Finally, with respect to the induced topology, the inclusion $\iota : M \to N$ is a homeomorphism and, for any pair of charts φ for M and ψ for N as above, we have $\psi \circ \iota \circ \varphi^{-1} = (\psi \circ \xi) \circ \eta$, which has injective derivative. That is, ι is an embedding, which establishes (b).

To see that *(a) implies (d)*, we simply proceed as above, defining the components of f to be the components $\sigma(m+1), \ldots, \sigma(n)$ of the map $\tilde{\xi} \circ \eta|_B$.

It remains then to prove that *(e) implies (d)*. Given $p \in M$, let $\pi : V \to \mathbb{R}^{n-m}$ be a submersion of a neighbourhood V of p in N such that $\pi^{-1}(\{0\}) = M \cap V$. Without loss of generality, the closure of V lies in the domain of a chart ψ for N. Set $x \doteqdot \psi^{-1}(p)$ and $U \doteqdot \psi(V)$. Since $d(\pi \circ \psi^{-1})_x$ is a surjective, we may find a bijection $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ such that its columns $\sigma(m+1), \ldots, \sigma(n)$ are linearly independent. Define $P : U \to \mathbb{R}^n$ by $P(z) \doteqdot (z_{\sigma 1}, \ldots, z_{\sigma(m)}, \pi(z))$. Then dP_x is an isomorphism, and hence the inverse function theorem provides an inverse P^{-1} for P restricted to some neighbourhood of x. Observe that (d) holds with $f(z_1, \ldots, z_m)$ given by the components $\sigma(m+1), \ldots, \sigma(n)$ of $P^{-1}(z_1, \ldots, z_m, 0)$. $\square$

A version of the proposition also holds in the context of manifolds-with-boundary.

Proposition 3.10 yields a rich supply of examples of manifolds. Indeed, by parts (a) and (b), the image of a smooth map $\xi : U \underset{\text{open}}{\subset} \mathbb{R}^m \to \mathbb{R}^n$ which is a homeomorphism onto its image and has everywhere injective derivative is a smooth m-manifold. By parts (b) and (d), the graph

$$\operatorname{graph} u \doteqdot \{(x, u(x)) : x \in \Omega \underset{\text{open}}{\subset} \mathbb{R}^n\}$$

of any smooth function $u : \Omega \to \mathbb{R}$ is a smooth n-manifold. By parts (b) and (e), the zero set

$$\operatorname{zero} u \doteqdot \{x \in \Omega \underset{\text{open}}{\subset} \mathbb{R}^{n+1} : u(x) = 0\}$$

of any smooth function $u : \Omega \to \mathbb{R}$ with nowhere-vanishing gradient is a smooth n-manifold.

Example 3.11. *By Proposition 3.10 (e), the following level sets are submanifolds of Euclidean space.*

(1) *The sphere,* $S^n = \{x \in \mathbb{R}^{n+1} : |x| = 1\}$*;*
(2) *The cylinder* $\{(x, y) \in \mathbb{R}^m \times \mathbb{R}^n : |x| = 1\}$*;*
(3) *The torus* $T^2 = \{(x, y) \in \mathbb{R}^2 \times \mathbb{R}^2 : |x| = |y| = \frac{1}{2}\}$*;*
(4) *The special linear group* $\mathrm{SL}(\mathbb{R}^n) \doteqdot \{L \in \mathrm{GL}(\mathbb{R}^n) : \det L = 1\}$ *;*
(5) *The orthogonal group* $\mathrm{O}(\mathbb{R}^n) \doteqdot \{L \in \mathrm{GL}(n) : LL^T = I\}$.

The inverse function theorem can be generalized to smooth maps between smooth manifolds(-with-boundary) by applying the classical version in charts.

Theorem 3.12 (Inverse function theorem)**.** *Let $f : M \to N$ be a smooth map between smooth manifolds(-with-boundary) M and N. If $df_p : T_pM \to T_{f(p)}N$ is an isomorphism then there exist neighborhoods U of p and V of $f(p)$ such that $f|_U$ is a diffeomorphism from U to V.*

Proof. Let $\varphi : U \to \mathbb{R}^n$ and $\eta : V \to \mathbb{R}^n$ be charts for M and N about p and $f(p)$, respectively. If $df_p : T_pM \to T_{f(p)}N$ is an isomorphism, then so is $d(\eta \circ f \circ \varphi^{-1})_{\varphi(p)}$. So the inverse function theorem provides open subsets $A \subset U$ and $B \subset V$ such that $\eta \circ f \circ \varphi^{-1}|_A : A \to B$ is a diffeomorphism. That is, f is a diffeomorphism from $\varphi^{-1}(A)$ to $\eta^{-1}(B)$. $\square$

The implicit function theorems are generalized similarly.

Theorem 3.13 (Implicit function theorem (surjective version))**.** *Let $f : M \to N$ be a smooth map between smooth manifolds(-with-boundary) M and N. If $df(p) : T_pM \to T_{f(x)}M$ is an epimorphism then there exists a neighborhood U of p such that $f^{-1}(\{f(p)\}) \cap U$ is a smooth submanifold of M of dimension* $\operatorname{null}(df(p))$.

Theorem 3.14 (Implicit function theorem (injective version)). *Let $f : M \to N$ be a smooth map between smooth manifolds(-with-boundary) M and N. If $df(p) : T_pM \to T_{f(p)}M$ is a monomorphism then there exists a neighborhood U of p such that $f|_U$ is an embedding.*

Exercises.

Exercise 3.1. *Show that the equivalence relation in Definition 3.2 does not depend on the choice of chart.*

Exercise 3.2. *Consider the differentiable manifold defined by equipping $\mathbb{R}^n$ with its standard differentiable structure (induced by the identity chart $\mathrm{id} : \mathbb{R}^n \to \mathbb{R}^n$). Given $p \in \mathbb{R}^n$, $T_p\mathbb{R}^n$ is isomorphic to $\mathbb{R}^n$ via the identification*

$$[(\mathrm{id}, u)] \mapsto u\,.$$

Write down explicit isomorphisms $P_p\mathbb{R}^n \to \mathbb{R}^n$ and $D_p\mathbb{R}^n \to \mathbb{R}^n$.

Exercise 3.3. *Let M^m be a smooth manifold and p a point of M^m. Show that, given any chart $\varphi : U \to \mathbb{R}^n$ about a point p, the induced bases for T_pM, P_pM and D_pM coincide under the corresponding natural isomorphisms.*

Exercise 3.4. *Show that the rank of a smooth map does not depend on the choice of charts.*

Exercise 3.5. *Let φ be a chart for a differentiable manifold. Show that φ^{-1} is a local diffeomorphism.*

Exercise 3.6. *Show that the projection map $\pi : S^n \to \mathbb{R}P^n$ defined by $\pi(x) = [x]$ is a local diffeomorphism.*

Exercise 3.7. *Let M be a differentiable manifold.*

- *Show that, for any $y \in M$, the inclusion map $\iota : M \to M \times M$ defined by $\iota(x) = (x, y)$ is an embedding.*
- *Show that the projection map $\pi : M \times M \to M$ defined by $\iota(x, y) = x$ is a submersion.*

Exercise 3.8. *Let G be a Lie group. Show that, for any $x \in G$, left multiplication by x, $y \mapsto L_x(y) \doteqdot x \cdot y$, is a local diffeomorphism.*

Exercise 3.9. *Let G be a Lie group. Show that the multiplication map is a submersion from $G \times G$ to G.*

Exercise 3.10. *Show that the inclusion of S^n (equipped with the atlas induced by stereographic projection) into $\mathbb{R}^{n+1}$ is an embedding.*

Exercise 3.11. *Prove that $\mathbb{C}P^1$ is diffeomorphic to S^2.* Hint*:* Consider the atlas for S^2 given by the two stereographic projections, and the atlas for $\mathbb{C}P^1$ given by the two projections $[z_1, z_2] \mapsto z_2/z_1 \in \mathbb{C} \cong \mathbb{R}^2$ and $[z_1, z_2] \mapsto$

z_1/z_2. Now define a map between the two manifolds by defining it between corresponding charts in such a way that it agrees on overlaps.

Exercise 3.12. *Consider* $S^3 \subset \mathbb{R}^4 \cong \mathbb{C}^2$. *Define* $\pi : S^3 \to \mathbb{C}P^1 \cong S^2$ *to be the restriction of the canonical projection* $(z_1, z_2) \mapsto [z_1, z_2]$. *Show that* π *is a submersion.*

Exercise 3.13. *What are the dimensions of the submanifolds* $\mathrm{SL}(\mathbb{R}^n)$ *and* $\mathrm{O}(\mathbb{R}^n)$ *of* $\mathrm{GL}(\mathbb{R}^n)$*?*

4. Some differential topology

We will now briefly present some beautiful applications of differential geometry to topology, following Milnor [**6**].

Definition 4.1. *Let $f : M \to N$ be a smooth map between smooth manifolds-with-boundary M and N. A point $x \in M$ is **critical** (a.k.a. **singular**) if $df(x) : T_xM \to T_{f(x)}N$ is not surjective (that is, if $\operatorname{Rank}(df(x)) < \dim(N)$). Else, x is a **regular** point. A value $y \in N$ is **regular** if $f^{-1}(\{y\})$ contains no singular points. Else, y is a **critical** (a.k.a. **singular**) value.*

Note that every point $y \in N \setminus f(M)$ is counted as a regular point (since $f^{-1}(\{y\})$ is empty, it cannot contain any singular point). We will denote the set of regular values of a smooth map f by $\operatorname{reg} f$. By the implicit function theorem (surjective version), the level sets $f^{-1}(\{y\})$ corresponding to a regular value y are smooth submanifolds of M of dimension $\dim(M) - \dim(N)$.

Theorem 4.2 (Regular value theorem)**.** *Let $f : M \to N$ be a smooth map between smooth manifolds-with-boundary M and N. If $y \in N$ is a regular value for both f and $f|_{\partial M}$ then $f^{-1}(\{y\})$ is a smooth submanifold-with-boundary of M and $\partial f^{-1}(\{y\}) = f^{-1}(\{y\}) \cap \partial M$ of dimension $\dim(M) - \dim(N)$.*

Example 4.3. *Define $u : \mathbb{R}^{n+1} \to \mathbb{R}$ by $u(x) \doteqdot \frac{1}{2}|x|^2$. Then $Du(x) = x$. That is, given $v \in T_x\mathbb{R}^{n+1} \cong \mathbb{R}^{n+1}$, $D_vu(x) = \langle v, x\rangle$. It follows that all points $x \in \mathbb{R}^{n+1}$ except for the origin are regular and all values $t > 0$ are regular. So the level sets $u^{-1}(\{t\}) = \{|x|^2 = 2t\} = S^n_{\sqrt{2t}}$ are smooth submanifolds of $\mathbb{R}^{n+1}$.*

If M is compact and $\dim(M) = \dim(N)$ then $f^{-1}(\{y\})$ is a compact zero dimensional submanifold-with-boundary of M and hence a finite set of points. In this case, we can define the function $\#f^{-1} : \operatorname{reg} f \to \mathbb{N}$ by $\#f^{-1}(y) \doteqdot |f^{-1}(\{y\})|$.

Lemma 4.4. *Let $f : M \to N$ be a smooth map from a smooth, compact manifold M to a smooth manifold N of the same dimension. The function $\#f^{-1}$ is locally constant.*

Proof. Given $y \in \operatorname{reg} f$ there exist finitely many points $x_1, \dots, x_N \in M$ such that $f^{-1}(\{y\}) = \{x_1, \dots, x_N\}$. By the inverse function theorem, we can find pairwise disjoint neighborhoods $U_1, \dots, U_N$ of $x_1, \dots, x_N$, respectively, on which f is a diffeomorphism. Set $V_i = f(U_i)$ for each $i = 1, \dots, N$. We claim that $\#f^{-1}$ is constant on the neighborhood

$$V \doteqdot (V_1 \cap \cdots \cap V_N) \setminus f(M \setminus (U_1 \cup \cdots \cup U_N))$$

of y. Indeed, given $z \in V$ there exist, for each $i = 1, \ldots, N$, points $z_i \in U_i$ such that $f(z_i) = z$. So $\#f^{-1}(z) \geq \#f^{-1}(y)$. On the other hand, if $w \in f^{-1}(\{z\})$ then $w \in U_i$ for some i and hence $w = z_i$. $\square$

Theorem 4.5 (Fundamental theorem of algebra)**.** *Every non-constant polynomial over $\mathbb{C}$ has a zero in $\mathbb{C}$.*

Proof. Recall that the stereographic projection maps $\overline{\varphi} : S^2 \setminus \{(0,0,-1)\} \to \mathbb{C}$ and $\underline{\varphi} : S^2 \setminus \{(0,0,1)\} \to \mathbb{C}$ are given by

$$\overline{\varphi}(x,y,z) = \left(\frac{x}{1+z}, \frac{y}{1+z}\right) \quad \text{and} \quad \underline{\varphi}(x,y,z) = \left(\frac{x}{1-z}, \frac{y}{1-z}\right)$$

and their inverses by

$$\overline{\varphi}^{-1}(u+iv) = \frac{\left(2u, 2v, 1+u^2+v^2\right)}{1+u^2+v^2} \quad \text{and} \quad \underline{\varphi}^{-1}(u+iv) = \frac{\left(2u, 2v, -1+u^2+v^2\right)}{1+u^2+v^2}.$$

A polynomial $p : \mathbb{C} \to \mathbb{C}$, $z \mapsto p(z) = a_0 z^n + \cdots + a_n$, $a_0 \neq 0$, induces a map $f : S^2 \to S^2$ via

$$f(X) \doteqdot \begin{cases} (\underline{\varphi}^{-1} \circ p \circ \underline{\varphi})(X) & \text{if } X \neq (0,0,1) \\ (0,0,1) & \text{if } X = (0,0,1)\,. \end{cases}$$

We claim that f is smooth (with respect to the differentiable structure on S^2 induced by stereographic projection). It suffices to check that f is smooth at $(0,0,1)$, and for this it suffices to check that the map

$$\overline{\varphi} \circ f \circ \overline{\varphi}^{-1} = \overline{\varphi} \circ \underline{\varphi}^{-1} \circ p \circ \underline{\varphi} \circ \overline{\varphi}^{-1}$$

is a smooth map from $\mathbb{C} \setminus \{0\}$ to $\mathbb{C} \setminus \{0\}$. Observe (either by algebraic manipulations or Euclidean geometry) that

$$\overline{\varphi} \circ \underline{\varphi}^{-1}(z) = \underline{\varphi} \circ \overline{\varphi}^{-1}(z) = \frac{z}{|z|^2} = \frac{1}{\overline{z}}\,.$$

Thus,

$$\overline{\varphi} \circ f \circ \overline{\varphi}^{-1}(z) = \frac{1}{\overline{p\left(\overline{z}^{-1}\right)}} = \frac{z^n}{\overline{a}_0 + \overline{a}_1 z + \cdots + \overline{a}_n z^n}\,,$$

which is clearly a smooth map.

Since p is non-constant, Dp has finitely many zeroes. That is, f has only finitely many critical points and hence only finitely many critical values, $y_1, \ldots, y_N$. It follows that $\mathrm{reg}f = S^2 \setminus \{y_1, \ldots, y_N\}$ is connected. The preceding lemma now implies that $\#f^{-1} : \mathrm{reg}f \to \mathbb{N}$ is constant. Since p is non-constant, $\mathrm{reg}f$ is nonempty, so the constant $\#f^{-1}$ must be positive. We conclude that f is surjective. In particular, $0 \in p(\mathbb{C})$. $\square$

Theorem 4.6 (Weierstrass approximation theorem)**.** *The space of smooth functions on a compact smooth manifold-with-boundary is dense in the space of continuous functions (with respect to the uniform topology).*

Proof. We will prove the claim when $M = \overline{B}^n$. The proof for the general case is not that much harder (but we have to work in charts). First, we extend f to a continuous function on $\mathbb{R}^n$ by

$$f(x) \doteqdot \begin{cases} f(x) & \text{if } x \in \overline{B}^n \\ f\left(\dfrac{x}{\|x\|}\right)(\|x\| - 2) & \text{if } x \in \overline{B}^n_2 \setminus \overline{B}^n \\ 0 & \text{if } x \notin \overline{B}^n_2 \,. \end{cases}$$

Next define, for any $t > 0$,

$$f_t(x) \doteqdot (4\pi t)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \mathrm{e}^{\frac{\|x-y\|^2}{-2t}} f(y)\, dy\,. \tag{4.1}$$

By the bounded convergence theorem, we can differentiate through the integral; so f_t is smooth for all $t > 0$. Moreover, $f_t \to f$ uniformly as $t \to 0$. $\square$

Remark 4.7. *The function $(x, t) \mapsto f_t(x)$ defined by (4.1) is the solution to the heat equation with initial condition $f_0 = f$.*

Lemma 4.8 (Hirsch)**.** *No smooth, compact manifold-with-boundary M admits a smooth retraction $f : M \to \partial M$ to its boundary ∂M.*

Proof. Suppose, to the contrary, that there is a smooth retraction $f : M \to \partial M$. By Sard's theorem (see [**6**, §3]), there exists some $y \in \mathrm{reg} f$. Since $f|_{\partial M}$ is the identity map, $y \in \mathrm{reg} f|_{\partial M}$ and the regular value theorem implies that $f^{-1}(\{y\})$ is a smooth, compact 1-submanifold-with-boundary of M and $\partial f^{-1}(\{y\}) = f^{-1}(\{y\}) \cap \partial M = \{y\}$. But this is impossible: any compact 1-manifold-with-boundary is a finite disjoint union of circles and compact intervals, and hence possesses an *even* number of boundary points, a contradiction. $\square$

Corollary 4.9 (Smooth Brouwer)**.** *Every smooth map $f : \overline{B}^n \to \overline{B}^n$ has a fixed point.*

Proof. Suppose, to the contrary, that there exists a smooth map $g : \overline{B}^n \to \overline{B}^n$ with no fixed points. Then the map f defined by sending a point $x \in \overline{B}^n$ to the intersection of ∂B^n with the ray from $f(x)$ through x is a smooth retraction of $\overline{B}^n$ onto ∂B^n. $\square$

Corollary 4.10 (Brouwer)**.** *Every continuous map $f : \overline{B}^n \to \overline{B}^n$ has a fixed point.*

Proof. Suppose, to the contrary, that there exists a continuous map $g : \overline{B}^n \to \overline{B}^n$ with no fixed points. Set

$$\mu \doteqdot \min_{\overline{B}^n} \|g(x) - x\| > 0\,.$$

By the Weierstrass approximation theorem, there exists a smooth function $f:\overline{B}^n\to\overline{B}^n$ such that

$$\|f-g\|_\infty<\frac{\mu}{2}\,.$$

Set

$$g_\mu \doteqdot \frac{f}{1+\mu}\,.$$

Then

$$\|g_\mu-g\|_\infty=\frac{1}{1+\frac{\mu}{2}}\|p-\left(1+\tfrac{\mu}{2}\right)g\|_\infty\le\|p-g\|_\infty+\tfrac{\mu}{2}\|g\|_\infty<\mu\,.$$

It follows that g_μ has no fixed points, contradicting the smooth Brouwer fixed point theorem. Indeed, if $g_\mu(z)=z$ for some $z\in\overline{B}^n$ then

$$\mu=\min_{\overline{B}^n}\|g(x)-x\|\le\|g(z)-z\|=\|g(z)-g_\mu(z)\|\le\|g_\mu-g\|_\infty<\mu\,,$$

a contradiction. □

The following theorem, known colloquially as the "you can't comb a hairy ball" theorem, is usually proved by methods of algebraic topology. The following beautiful proof (due to Peter McGrath) is distinctly differentio-topological.

Theorem 4.11 (Hairy ball theorem)**.** *The unit two sphere $S^2\subset\mathbb{R}^3$ admits no continuous nowhere-vanishing vector field.*

In the statement of the theorem, a (**continuous**) **vector field on** S^2 is taken to be a (continuous) map $V:S^2\to\mathbb{R}^3$ such that $\langle V_p,p\rangle\equiv0$.

The subspace $L_p\doteqdot\{v\in\mathbb{R}^3:\langle v,p\rangle=0\}$ of $\mathbb{R}^3$ clearly has something to do with the tangent space T_pS^2 to S^2 at p. Indeed, we can (canonically) identify the two spaces as follows. First note that the inclusion map $\iota:S^2\to\mathbb{R}^3$ induces an injective linear homomorphism $d\iota_p:T_pS^2\to T_p\mathbb{R}^3$ for each $p\in S^2$. On the other hand, we may identify $T_p\mathbb{R}^3$ with $\mathbb{R}^3$ via the isomorphism

$$u^i\partial_i|_p\mapsto u^ie_i\,,$$

where $\{e_1,e_2,e_3\}$ the standard oriented basis for $\mathbb{R}^3$ and $\{\partial_1|_p,\partial_2|_p,\partial_3|_p\}$ is the coordinate basis for $T_p\mathbb{R}^3$ with respect to the identity chart. Now, if $\gamma:I\to S^2$ is a regular curve, then $|\iota\circ\gamma|^2\equiv1$, and hence

$$0\equiv\left.\frac{d}{dt}\right|_{t=0}\frac{1}{2}|\iota\circ\gamma|^2=\big\langle(\iota\circ\gamma)'(0),\iota\circ\gamma(0)\big\rangle\,.$$

That is, $(\iota\circ\gamma)'(0)\in L_{(\iota\circ\gamma)(0)}$. The map $[\gamma]\in T_pS^2\mapsto(\iota\circ\gamma)'(0)\in L_p$ coincides with $d\iota_p$ after identifying $T_p\mathbb{R}^3\cong\mathbb{R}^3$, and is therefore an isomorphism.

Observe also that we can equip each tangent space $T_p\mathbb{R}^3$ to $\mathbb{R}^3$ (and hence also each tangent space T_pS^2 to S^2) with a canonical inner product $\langle\cdot,\cdot\rangle_p$ using the canonical identification of $T_p\mathbb{R}^3$ and $\mathbb{R}^3$. That is,

$$\left\langle u^i\partial_i|_p, v^i\partial_i|_p\right\rangle_p \doteqdot \sum_{i=1}^{3} u^i v^i\,.$$

Proof of Theorem 4.11 (following Peter McGrath). Contrary to the claim, suppose that S^2 admits a continuous nowhere-vanishing vector field V. Without loss of generality, $|V|\equiv 1$ (since else we may replace V with $V/|V|$).

At each point $p\in S^2$, we may equip $T_pS^2\subset T_p\mathbb{R}^3$ with the orthonormal basis $\{V_p, V_p^\perp\}$, where $V_p^\perp$ is the unit vector orthogonal to both V_p and $X_p\doteqdot p^i\partial_i|_p$ such that $\{V_p, V_p^\perp, X_p\}$ has the same orientation as the coordinate basis.

For each $p\in S^2$, denote by Φ_p the isometry of $\mathbb{R}^3$ that maps p to the origin and whose derivative at p maps $\{V_p, V_p^\perp, X_p\}$ to $\{e_1,e_2,e_3\}$ after identifying $T_0\mathbb{R}^3$ with $\mathbb{R}^3$. Note that this map $F_p\doteqdot(d\Phi_p)_p : T_p\mathbb{R}^3\to T_0\mathbb{R}^3\cong\mathbb{R}^3$ maps a vector tangent to S^2 at p into the oriented plane $e_1\wedge e_2$, which we identify with $\mathbb{R}^2$ in the obvious way. We may now define the **rotation number** (with respect to V) of a regular curve $\gamma : S^1\to S^2$ as the winding number of the curve $s\mapsto F_{\gamma(s)}\gamma'(s)$.

Given $p\in S^2$ and $s\in(-1,1)$, consider the **parallel** at height s in the direction p,

$$C_{p,s}\doteqdot\{q\in S^2 : \langle q,p\rangle = s\}\,.$$

We may choose a regular parametrization $c_{p,s} : S^1\to S^2$ of each $C_{p,s}$ such that $|c_{p,s}'|\equiv 1$ and the orientation of $\{F_{c_{p,s}}c_{p,s}', c_{p,s}\}$ agrees with that of $\{e_1,e_3\}$. These curves are all regularly homotopic and so have the same rotation number, say $r\in\mathbb{Z}$. Now, when $s=0$, the two curves $c_{p,0}$ and $c_{-p,0}$ parametrize the same parallel but with opposite orientations. Thus, $r=-r$, and hence $r=0$. On the other hand, for s close to 1, the rotation number of $c_{p,s}$ is close to the rotation number of a circle in the plane because V is close to V_p on $c_{p,s}$. So $r\in\{-1,1\}$, which is impossible. $\square$

5. The tensor algebra of a linear space

> "*A comathematician is a machine for turning cotheorems into fee.*" — Paul Erdös*.

Associated with any finite dimensional real linear space V is its **dual space** V^* — the real linear space of real linear maps from V to $\mathbb{R}$. The dual $(V^*)^*$ of V^* is isomorphic to V via the obvious identification[6]:

$$v(\alpha) \doteqdot \alpha(v) \text{ for any } v \in V \text{ and } \alpha \in V^*.$$

Associated with any pair of (not necessarily distinct) finite dimensional real linear spaces spaces U and V are the **direct sum** $U \oplus V$ and **tensor product** $U \otimes V$.

The **direct sum** $U \oplus V$ is simply the real linear space formed by equipping the set of ordered pairs $(u_1, v_1) \in U \times V$ with the obvious linear structure:

$$(u_1, v_1) + \lambda(u_2, v_2) \doteqdot (u_1 + \lambda u_2, v_1 + \lambda v_2)$$

for $u_1, u_2 \in U$, $v_1, v_2 \in V$ and $\lambda \in \mathbb{R}$. If we identify U and V with $U \oplus \{0\}$ and $\{0\} \oplus V$, respectively, we may write $u+v$ instead of (u, v) without ambiguity. The direct sum is associative and Abelian, in the sense that $(U \oplus V) \oplus W$ is canonically isomorphic to $U \oplus (V \oplus W)$ and $U \oplus V$ is canonically isomorphic to $V \oplus U$. The **infinite direct sum**[7] $\oplus_{i=1}^{\infty} V_i$ is the real linear space generated by all *finite* formal linear combinations of elements $v_i \in V_i$.

Given finite dimensional linear spaces U and V, the **tensor product** $U^* \otimes V$ is *defined to be*[8] the linear space $\mathrm{Hom}(U, V)$ of linear maps from U to V. Given $\alpha \in U^*$ and $v \in V$, we can define a "tensor" $\alpha \otimes v \in U^* \otimes V$, called the **tensor product** of α and v, by

$$(\alpha \otimes v)(u) \doteqdot \alpha(u)v \text{ for } u \in U.$$

Since any finite dimensional linear space may be treated as the dual of its dual, this defines the tensor product of any pair of finite dimensional linear spaces.

Note that $\mathrm{Hom}(U, V)$ is isomorphic to the space $\mathrm{Hom}(U, V^*; \mathbb{R})$ of *bilinear* maps from $U \times V^*$ to $\mathbb{R}$ via the obvious rule

$$T(u, \vartheta) \doteqdot \vartheta(T(u)).$$

[6] Of course, *any* two linear spaces of the same finite dimension are isomorphic. The point here is that we are presented with a "canonical" choice of isomorhism, which we may henceforth employ without further mention, whereas, in general, such an isomorphism depends on making some arbitrary choice (of basis, say).

[7] Note that this is not the same as equipping the infinite Cartesian product with the canonical linear structure. Compare the box and product topologies on $\prod_{\mathbb{N}} \mathbb{R}$.

[8] There are other definitions (which agree up to "canonical" isomorphism). See, for example, [**8**, Chapter 2].

Elements of a tensor product space $U \otimes V$ which can be written as the tensor product $u \otimes v$ of elements u of U and v of V are called **simple**. Each element of $U \otimes V$ can be written as a linear combination of simple elements. Indeed, if $\{e_i\}_{i=1}^m$ is a basis for U and $\{f_i\}_{i=1}^n$ is a basis for V, then $\{e_i \otimes f_j\}_{i,j=1}^{m,n}$ is a basis for $U \otimes V$. So $U \otimes V$ is a real linear space of dimension $m \times n$.

The keen eyed observer will have noticed that we have just opened Pandora's box — we can generate an endless number of natural spaces from any given real linear space V: applying the dual, direct sum and tensor product operations yields myriad "tensor spaces" over V. These spaces are also linear, as are the spaces of linear maps between them, and also the tensor spaces over these new spaces, etc, etc. The number of algebraically distinct spaces thus produced is greatly reduced by canonical isomorphisms, however. This is formalized by the **universal property**.

Proposition 5.1 (The universal property). *Let U, V and W be finite dimensional real linear spaces. Given any bilinear map $U \times V \to W$ there exists a unique linear map $U \otimes V \to W$ such that the following diagram commutes:*

$$
\begin{array}{ccc}
U \times V & \longrightarrow & W \\
\pi \downarrow & \nearrow & \\
U \otimes V & &
\end{array}
$$

where $\pi(u, v) \doteqdot u \otimes v$.

Proof. See [**8**, Chapter 2] or [**5**, Chapter 12]. □

Modulo canonical isomorphism, the tensor product of linear spaces is associative and commutative. Note, however, that the tensor product of elements of these spaces is *not* commutative.

The **tensor algebra** $T(V)$ over a finite dimensional real linear space V is the algebra generated by finitely many direct sums and tensor products of elements of V^*:

$$T(V) \doteqdot \bigoplus_{k\in\mathbb{N}_0} \bigotimes_{i=0}^{k} V^* .$$

The tensor algebra is $\mathbb{Z}$-**graded**: it is a direct sum of **homogeneous subspaces**

$$T^k(V) \doteqdot \bigotimes_{i=0}^{k} V^*$$

satisfying $T^k(V) \otimes T^l(V) \subset T^{k+l}(V)$. The **degree** of a homogeneous tensor $T \in T^k(V)$ is defined as $\deg(T) = k$.

The tensor algebra is also $\mathbb{Z}_2$-graded: if we define the ($\mathbb{Z}_2$-**homogeneous**) subspaces $H^0(V)$ and $H^1(V)$ of direct sums of tensors of even and odd degree, respectively, then

$$T(V) = H^0(V) \oplus H^1(V)$$

and $H^i(V) \otimes H^j(V) \subset H^{i+j}(V)$, where addition of indices is understood modulo 2. Define a map $\eta : T(V) \to T(V)$ by

$$\eta(T) \doteqdot (-1)^{\deg(T)} T,$$

for homogeneous elements T, extended to all elements by distributing over direct sums. Observe that η is an involutory automorphism. Indeed, if S and T are homogeneous tensors, then

$$\eta(S \otimes T) = (-1)^{\deg(S)+\deg(T)} S \otimes T = (-1)^{\deg(S)} S \otimes (-1)^{\deg(T)} T = \eta(S) \otimes \eta(T).$$

Involutivity is clear. The even and odd subspaces are eigenspaces of η with eigenvalue 1 and -1, respectively.

The **mixed tensor algebra** $T(V)$ over a finite dimensional real linear space V is the $\mathbb{Z} \times \mathbb{Z}$-graded algebra generated by finitely many direct sums and tensor products of elements of V and V^*, modulo canonical isompor-phisms:

$$T(V) \doteqdot \bigoplus_{(k,l) \in \mathbb{N}_0 \times \mathbb{N}_0} T^{(k,l)}(V),$$

where the **homogeneous** tensor spaces $T^{(k,l)}(V)$ are defined by

$$T^{(k,l)}(V) \doteqdot \bigotimes_{i=0}^{k} V^* \otimes \bigotimes_{j=0}^{l} V.$$

By the universal property, we can identify $T^{(k,l)}(V)$ with the space

$$T^{(k,l)}(V) = \mathrm{Hom}\left(\underbrace{V, \ldots, V}_{k\text{-times}}; \bigotimes_{j=1}^{l} V\right)$$

of multilinear maps from $\prod_{i=1}^{k} V$ to $\bigotimes_{j=1}^{l} V$, where $\prod$ denotes the Cartesian product.

5.1. Traces. Let V be a finite dimensional real linear space. Observe that the homogeneous subspace $T^{(1,1)}(V)$ is equipped with a **trace** operator $\mathrm{tr} : T^{(1,1)}(V) \to \mathbb{R}$ defined by

$$\mathrm{tr}(T) \doteqdot \sum_{i=1}^{n} \vartheta^i(T(e_i)) = \sum_{i=1}^{n} T_i^i\,,$$

where $\{e_i\}_{i=1}^n$ is some basis for V and $\{\vartheta^i\}_{i=1}^n$ is the corresponding **dual basis** for V^*; that is,

$$\vartheta^j(e_i) \doteq \delta_i^j \doteq \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \,. \end{cases}$$

In order to make things more manageable, we henceforth make use of the **Einstein summation convention**. That is, unless otherwise stated (or clear from context), indices which appear twice, once in an "upper" position and once in a "lower" position, are implicitly summed over.

Note that the trace is independent of the choice of basis for V. Indeed, if $\{\hat{e}_i\}_{i=1}^n$ is some other basis for V and $\{\hat{\vartheta}^i\}_{i=1}^n$ is the corresponding dual basis for V^*, then we can find some $L = L_i^j \vartheta^i \otimes e_j \in GL(V)$ such that

$$\hat{e}_i = L(e_i) = L_i^j e_j$$

for each i. But then

$$\hat{\vartheta}^j(e_i) = (L^{-1})_i^k \hat{\vartheta}^j(\hat{e}_k) = (L^{-1})_i^j \quad \Longrightarrow \quad \hat{\vartheta}^j = (L^{-1})_i^j \vartheta^i \,,$$

where $(L^{-1})_i^j$ are the components of L^{-1}. Note that

$$\begin{aligned} (L^{-1})_i^k L_k^j &= (L^{-1})_i^k \vartheta^j(L(e_k)) \\ &= \vartheta^j(L((L^{-1})_i^k e_k)) \\ &= \vartheta^j(L(L^{-1}(e_i))) \\ &= \vartheta^j(e_i) \\ &= \delta_i^j \,, \end{aligned}$$

so

$$\begin{aligned} \hat{\vartheta}^i(T(\hat{e}_i)) &= (L^{-1})_k^i \vartheta^k(T(L_i^j e_j)) \\ &= (L^{-1})_k^i L_i^j \vartheta^k(T(e_j)) \\ &= \delta_k^j \vartheta^k(T(e_j)) \\ &= \vartheta^j(T(e_j)) \,. \end{aligned}$$

When k and l are both at least one, the homogeneous subspace $T^{(k,l)}(V)$ is equipped with a *family* of trace operators $\operatorname{tr}_{(i,j)} : T^{(k,l)}(V) \to T^{(k-1,l-1)}(V)$ defined on each $T \in T^{(k,l)}(V)$ by

$$\begin{aligned} &(\operatorname{tr}_{(i,j)} T)(u_1, \dots, u_{k-1}, \vartheta^1, \dots, \vartheta^l) \\ &\quad \doteq \operatorname{tr}\Big((u, \vartheta) \mapsto T(u_1, \dots, u_{i-1}, u, u_{i+1}, \dots, u_{k-1}, \vartheta^1, \dots, \vartheta^{j-1}, \vartheta, \vartheta^{j+1}, \vartheta^l)\Big) \,. \end{aligned}$$

for any $(k-1)$-tuple of vectors $u_1, \dots, u_{k-1}$ and any $(l-1)$-tuple of covectors $\vartheta_1, \dots, \vartheta_{l-1}$. That is, we "freeze" $k-1$ covariant factors and $l-1$ contravariant factors and take the trace of the resulting $(1,1)$-tensor.

5.2. Derivations. Let V be a finite dimensional real linear space. A linear map D from the (covariant or mixed) tensor algebra $T(V)$ to itself is called a **derivation** if it satisfies the Leibniz rule:

$$D(S \otimes T) = DS \otimes T + S \otimes DT\,.$$

A derivation on the covariant tensor algebra is **$\mathbb{Z}$-graded of degree** p if $D(T^k(V)) \subset T^{k+p}(V)$; $\mathbb{Z}\times\mathbb{Z}$-graded derivations on the mixed tensor algebra are defined similarly, as are and $\mathbb{Z}_2$- and $\mathbb{Z}_2 \times \mathbb{Z}_2$-graded derivations.

Observe that any derivation on the covariant tensor algebra is uniquely determined by its action on V^*. Indeed,

$$\begin{aligned} DT &= D(T_{i_1\dots i_k}\vartheta^{i_1} \otimes \cdots \otimes \vartheta^{i_k}) \\ &= T_{i_1\dots i_k} D(\vartheta^{i_1} \otimes \cdots \otimes \vartheta^{i_k}) \\ &= T_{i_1\dots i_k} \sum_{j=1}^{k} \vartheta^{i_1} \otimes \cdots \otimes D\vartheta^{i_j} \otimes \cdots \otimes \vartheta^{i_k}\,. \end{aligned}$$

Similarly, any derivation on the mixed tensor algebra is uniquely determined by its actions on V and V^*.

A graded derivation of degree zero on the mixed tensor algebra **commutes with traces** if

$$D(\operatorname{tr} T) = \operatorname{tr}(DT)$$

for any trace operator tr. Such derivations are uniquely determined by their action on V since

$$\begin{aligned} 0 &= D(\vartheta(u)) \\ &= D(\operatorname{tr}(\vartheta \otimes u)) \\ &= \operatorname{tr}(D(\vartheta \otimes u))) \\ &= \operatorname{tr}(D\vartheta \otimes u + \vartheta \otimes Du)) \\ &= D\vartheta(u) + \vartheta(Du) \end{aligned}$$

and hence

$$D\vartheta(u) = -\vartheta(Du)$$

for any $u \in V$ and any $\vartheta \in V^*$.

5.3. Symmetric and skew-symmetric tensors. Let V be a finite dimensional linear space. Given $k \in \mathbb{N}$ and $i, j \in \{1, \dots, k\}$, denote by π_{ij} the isomorphism of $\prod_{i=1}^k V$ which interchanges the i-th and j-th factors. A tensor $T \in T^k(V)$ is **symmetric** with respect to its i-th and j-th components if

$$T(\pi_{ij}(v)) = T(v)$$

for any $v \in \prod_{i=1}^k V$. T is **skew-symmetric** with respect to its i and j-th components if

$$T(\pi_{ij}(v)) = -T(v)$$

for any $v \in \prod_{i=1}^k V$. T is **totally symmetric** (resp. **totally skew-symmetric**) if it is symmetric (resp. skew-symmetric) with respect to every pair of components. Denote by $\Sigma^k(V)$ and $\Lambda^k(V)$ the subspaces of $T^k(V)$ consisting of its totally symmetric and skew-symmetric tensors, respectively.

The linear maps $\mathrm{Sym} : T^k(V) \to T^k(V)$ and $\mathrm{Alt} : T^k(V) \to T^k(V)$ defined by

$$\mathrm{Sym}(T)(v_1, \ldots, v_k) \doteqdot \frac{1}{k!} \sum_{\sigma \in S_k} T(v_{\sigma_1}, \ldots, v_{\sigma(k)})$$

and

$$\mathrm{Alt}(T)(v_1, \ldots, v_k) \doteqdot \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) T(v_{\sigma_1}, \ldots, v_{\sigma(k)})$$

respectively, are projections onto $\Sigma^k(V)$ and $\Lambda^k(V)$ (this is clear since they are linear, involutive and map into the corresponding subspace). If $\{\theta^i\}_{i=1}^n$ is a basis for V^*, then

$$\{\mathrm{Sym}(\theta^{i_1} \otimes \cdots \otimes \theta^{i_k})\}_{i_1, \ldots i_k = 1}^n \quad \text{and} \quad \{\mathrm{Alt}(\theta^{i_1} \otimes \cdots \otimes \theta^{i_k})\}_{i_1, \ldots i_k = 1}^n$$

are bases for $\Sigma^k(V)$ and $\Lambda^k(V)$, respectively. Some combinatorics then yield

$$\dim(\Sigma^k(V)) = \binom{n+k-1}{k} \quad \text{and} \quad \dim(\Lambda^k(V)) = \binom{n}{k},$$

where n is the dimension of V. In particular, $\dim(\Lambda^k(V)) = 0$ for $k > n$.

The tensor product of two totally symmetric (resp. totally skew-symmetric) tensors is not necessarily totally symmetric (resp. totally skew-symmetric). Define the **symmetric product** $\odot$ and **skew-symmetric product** $\wedge$ of two homogeneous tensors $\alpha, \beta \in T^k(V)$ by

$$\alpha \odot \beta \doteqdot \mathrm{Sym}(\alpha \otimes \beta)$$

and

$$\alpha \wedge \beta \doteqdot \mathrm{Alt}(\alpha \otimes \beta)$$

respectively, and extended to arbitrary tensors by distributing over addition.

Example 5.2. *When $\alpha, \beta \in V^*$, we obtain the familiar formulae*

$$\alpha \odot \beta = \frac{1}{2}(\alpha \otimes \beta + \beta \otimes \alpha) \tag{5.1}$$

and

$$\alpha \wedge \beta = \frac{1}{2}(\alpha \otimes \beta - \beta \otimes \alpha). \tag{5.2}$$

Another fact you should be familiar with is the decomposition

$$V^* \otimes V^* = \Sigma^2(V) \oplus \Lambda^2(V),$$

which follows from the fact that every simple element $\alpha \otimes \beta$ can be written as

$$\alpha \otimes \beta = \alpha \odot \beta + \alpha \wedge \beta.$$

Clearly, the symmetric (resp. skew-symmetric) product produces a symmetric (resp. skew-symmetric) tensor from a pair of symmetric (resp. skew-symmetric) tensors. The ($\mathbb{Z}$-graded) algebra of symmetric tensors is obtained by equipping the subspace

$$\Sigma(V) \doteqdot \bigoplus_{k=0}^{\infty} \Sigma^k(V)$$

of the tensor algebra with the symmetric product. The ($\mathbb{Z}$-graded) algebra of skew-symmetric tensors is obtained by equipping the subspace

$$\Lambda(V) \doteqdot \bigoplus_{k=0}^{\infty} \Lambda^k(V)$$

of the tensor algebra with the skew-symmetric product.

From now on, we will use the common names **exterior algebra** and **wedge product** for $\Lambda(V)$ and $\wedge$ respectively, and (**exterior**) k-**form** for elements of $\Lambda^k(V)$.

Proposition 5.3. *The wedge product is associative and anti-commutative:*

$$\alpha \wedge \beta = (-1)^{k\ell} \beta \wedge \alpha$$

for any two homogeneous elements $\alpha \in \Lambda^k(V)$ and $\beta \in \Lambda^\ell(V)$.

Proof. To prove associativity, first note that

$$\begin{aligned}
(\alpha \wedge \beta) \wedge \gamma &= \mathrm{Alt}(\mathrm{Alt}(\alpha \otimes \beta) \otimes \gamma) \\
&= \mathrm{Alt}((\alpha \otimes \beta) \otimes \gamma - (1 - \mathrm{Alt})(\alpha \otimes \beta) \otimes \gamma) \\
&= \mathrm{Alt}((\alpha \otimes \beta) \otimes \gamma) - \mathrm{Alt}((1 - \mathrm{Alt})(\alpha \otimes \beta) \otimes \gamma)).
\end{aligned}$$

By Exercise 5.5,

$$\mathrm{Alt}((1 - \mathrm{Alt})(\alpha \otimes \beta) \otimes \gamma) = 0.$$

It follows that

$$(\alpha \wedge \beta) \wedge \gamma = \mathrm{Alt}((\alpha \otimes \beta) \otimes \gamma).$$

Similarly

$$\alpha \wedge (\beta \wedge \gamma) = \mathrm{Alt}(\alpha \otimes (\beta \otimes \gamma)).$$

Associativity now follows from associativity of $\otimes$.

Since $\wedge$ distributes over addition, it suffices to prove the claim when α and β are simple. The claim follows from (5.2) when $k = 1$. Indeed, in that case,

$$\begin{aligned}\alpha \wedge \beta &= \alpha \wedge \beta^1 \wedge \cdots \wedge \beta^\ell \\ &= -\beta_1 \wedge \alpha \wedge \beta^2 \wedge \cdots \wedge \beta^\ell \\ &= \dots \\ &= (-1)^\ell \beta^1 \wedge \cdots \wedge \beta^\ell \wedge \alpha .\end{aligned}$$

Suppose, then, that the claim holds for some fixed k and any ℓ. Let $\alpha = \alpha^0 \wedge \alpha^k$ be a simple element of degree $k+1$ (where α^k has degree k) and β any simple element (of degree ℓ, say). Then

$$\begin{aligned}\alpha \wedge \beta &= \alpha^0 \wedge \alpha^k \wedge \beta \\ &= (-1)^{k\ell} \alpha^0 \wedge \beta \wedge \alpha^k \\ &= (-1)^{k\ell+\ell} \beta \wedge \alpha^0 \wedge \alpha^k \\ &= (-1)^{(k+1)\ell} \alpha \wedge \beta .\end{aligned}$$

Anti-commutativity now follows by induction. □

The **interior product** (a.k.a. **contraction**) of a k-form $\alpha \in \Lambda^k(V)$ by a vector $v \in V$ is the $(k-1)$-form $\iota_v \alpha$ defined by

$$\iota_v \alpha(v_1, \dots, v_{k-1}) \doteqdot \alpha(v, v_1, \dots, v_{k-1}) .$$

The interior product ι_v extends to $\Lambda(V)$ by distributing over addition. Observe that ι_v is an **anti-derivation**. This means that it satisfies the **skew-Leibniz rule**:

$$\iota_v(\alpha \wedge \beta) = \iota_v \alpha \wedge \beta + (-1)^k \alpha \wedge \iota_v \beta$$

for any k-form α, any $\beta \in \Lambda$ and any $v \in V$ (Exercise 5.8). We say that an anti-derivation $D : \Lambda(V) \to \Lambda(V)$ is **graded** if there is some k such that $D(\Lambda^p(V)) \subset \Lambda^{p+k}(V)$ for each p. The number k is its **degree**.

By the following proposition, ι_v is the unique graded anti-derivation of degree -1 on the exterior algebra which satisfies

$$\iota_v \alpha = \alpha(v)$$

for one-forms α.

Proposition 5.4. *Let $D_i : \Lambda(V) \to \Lambda(V)$, $i = 1, 2$, be anti-derivations. If*

$$D_1 \alpha = D_2 \alpha$$

for all $\alpha \in \Lambda^1(V)$, then $D_1 = D_2$.

Proof. Since $D_2 - D_1$ is an anti-derivation, it suffices to show that the only anti-derivation which is zero on one-forms is the zero derivation. This follows immediately from the formula

$$D(\alpha^0 \wedge \cdots \wedge \alpha^k) = \sum_{i=0}^{k} (-1)^i \alpha^0 \wedge \cdots \wedge \alpha^{i-1} \wedge D\alpha^i \wedge \alpha^{i+1} \wedge \cdots \wedge \alpha^k.$$

for any anti-derivation D and one-forms α^i, $i = 0, \ldots, k$. □

Proposition 5.5. *The interior product is closed:*

$$\iota_v \iota_v \alpha = 0$$

for every $v \in V$ and $\alpha \in \Lambda(V)$, and skew-symmetric,

$$(\iota_u \iota_v + \iota_v \iota_u)\alpha = 0$$

for every $u, v \in V$ and $\alpha \in \Lambda(V)$.

Proof. Observe that

$$\iota_u \iota_v (\alpha \wedge \beta) = \iota_u \iota_v \alpha \wedge \beta + (-1)^{k-1} \iota_v \alpha \wedge \iota_u \beta + (-1)^k \iota_u \alpha \wedge \iota_v \beta + \alpha \wedge \iota_u \iota_v \beta$$

and hence

$$(\iota_u \iota_v + \iota_v \iota_u)(\alpha \wedge \beta) = (\iota_u \iota_v + \iota_v \iota_u)\alpha \wedge \beta + \alpha \wedge (\iota_u \iota_v + \iota_v \iota_u)\beta.$$

That is, $\iota_u \iota_v + \iota_v \iota_u$ is a **derivation**. The claim follows since $\iota_u \iota_v + \iota_v \iota_u$ vanishes on 1-forms and any derivation D satisfies

$$D(\alpha^0 \wedge \cdots \wedge \alpha^k) = \sum_{i=0}^{k} \alpha^0 \wedge \cdots \wedge \alpha^{i-1} \wedge D\alpha^i \wedge \alpha^{i+1} \wedge \cdots \wedge \alpha^k.$$

on a product of 1-forms $\alpha^1, \ldots, \alpha^k$. □

Recall that

$$\dim(\Lambda^k(V)) = \binom{n}{k},$$

where n is the dimension of V. In particular, $\Lambda(V)$ is finite dimensional:

$$\dim(\Lambda(V)) = \sum_{k=0}^{\infty} \dim(\Lambda^k(V)) = \sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

Moreover, $\Lambda^k(V)$ is isomorphic to $\Lambda^{n-k}(V)$ and, in particular, $\Lambda^n(V)$ is 1-dimensional. Each nonzero element Ω of $\Lambda^n(V)$ induces an isomorphism $*_\Omega : \Lambda^k(V) \to \Lambda^{n-k}(V)$, called the **Hodge star map**, defined on basis elements $\theta^1 \wedge \cdots \wedge \theta^k$ by

$$*_\Omega(\theta^1 \wedge \cdots \wedge \theta^k) \doteqdot \iota_{e_k} \ldots \iota_{e_1} \Omega,$$

where $\{e_i\}_{i=1}^n$ is the basis for V dual to the basis $\{\theta^i\}_{i=1}^n$ for V^*. This definition extends to all of $\Lambda^k(V)$ by linearity.

If V^* is equipped with an inner product, then it is natural to equip it with an oriented orthonormal basis. Since the determinant of any special orthogonal transformation is 1, we obtain, by Exercise 5.9, a canonical choice of n-form

$$*1 \doteqdot e^1 \wedge \cdots \wedge e^n .$$

That is, $*1$ does not depend on the choice of oriented orthonormal basis (although it will clearly change sign if the orientation is reversed).

5.4. The tensor algebra of the tangent space. Now, we have seen that to every point p of a differentiable manifold M is associated a finite dimensional linear space — the tangent space T_pM to M at p. The space $T_p^*M \doteqdot (T_pM)^*$ dual to T_pM is called the **cotangent space** at p. Its elements are called **cotangent vectors** or **covectors** (at p).

A (**mixed**) **tensor** at p is any element of the (mixed) tensor algebra of T_pM. That is, any finite direct sum of finitely many tensor products of elements of T_p^*M (and T_pM). In practice, it will suffice to consider the **homogeneous** tensor spaces $T_p^{(k,l)}M \doteqdot T^{(k,l)}(T_pM)$.

Exercises.

Exercise 5.1. *Let V be a finite dimensional linear space. Show that the map $L : V \to (V^*)^*$, $v \mapsto L_v$ defined by*

$$L_v(\alpha) \doteqdot \alpha(v) \;\; \textit{for any} \;\; \alpha \in V^*$$

is an isomorphism.

Exercise 5.2. *Let U, V and W be finite dimensional linear spaces. Write down isomorphisms from $(U \oplus V) \oplus W$ to $U \oplus (V \oplus W)$ and from $U \oplus V$ to $V \oplus U$.*

Exercise 5.3. *Let U and V be finite dimensional linear spaces. If $\{e_i\}_{i=1}^m$ is a basis for U and $\{f_i\}_{i=1}^n$ is a basis for V, show that $\{e_i \otimes f_j\}_{i,j=1}^{m,n}$ is a basis for $U \otimes V$. Deduce that* $\dim(U \otimes V) = m \times n$.

Exercise 5.4. *Let U and V be finite dimensional linear spaces. Write down an isomorphism from $U \otimes V$ to $V \otimes U$. Show (by way of a counterexample) that, in general, $u \otimes v \neq v \otimes u$ when $U = V$.*

Exercise 5.5. *Let V be a real linear space and μ and α elements of its tensor algebra $T(V)$. Suppose that*

$$\mathrm{Alt}(\mu) = 0 .$$

Show that

$$\mathrm{Alt}(\alpha \otimes \mu) = 0 = \mathrm{Alt}(\mu \otimes \alpha) .$$

Exercise 5.6. *Let $\alpha_1, \dots, \alpha_k$ be 1-forms over a linear space V. Show that*

$$\alpha^1 \wedge \cdots \wedge \alpha^k = \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) \alpha^{\sigma(1)} \otimes \cdots \otimes \alpha^{\sigma(k)}.$$

Exercise 5.7. *Let α be a 1-form and β a k-form over a linear space V. Show that $\alpha \wedge \beta$ is the $(k+1)$-form defined by*

$$\alpha \wedge \beta(u_0, \dots, u_k) = \frac{1}{k+1} \sum_{i=1}^{k} (-1)^i \alpha(u_i) \beta(u_1, \dots, u_{i-1}, u_0, u_{i+1}, \dots, u_k).$$

Exercise 5.8. *Show that the interior product is a graded derivation of degree -1.*

Exercise 5.9. *Let V be a real linear space equipped with a basis $\{e_i\}_{i=1}^n$. Given $M \in \mathrm{GL}(V)$, we obtain a second basis $\{f_i\}_{i=1}^n$, where $f_i \doteqdot M(e_i)$. Show that*

$$f_1 \wedge \cdots \wedge f_n = \det(M) e_1 \wedge \cdots \wedge e_n.$$

Exercise 5.10. *Let $\varphi = (x^1, \dots, x^n) : U \to \mathbb{R}^n$ be a coordinate chart for a smooth manifold M. Show that the differentials dx^i_p at p of the component functions $x^i : U \to \mathbb{R}$ form a basis for T^*_pM which, moreover, is dual to the corresponding coordinate basis $\{\partial_{x^i}|_p\}_{i=1}^n$ for T_pM in the sense that*

$$dx^j_p(\partial_{x^i}|_p) = \delta^j_i \doteqdot \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$

6. The tangent bundle and its tensor algebra

We have constructed a tangent space T_pM at each point p of a manifold M. When we put all of these spaces together, we obtain the **tangent bundle** TM of M:

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

The tangent bundle is equipped with a natural surjection $\pi : TM \to M$ which sends a pair (p,v) to the "base point" p. (Note that we do *not* equip TM with the disjoint union topology. The topology for TM is constructed below).

If M has dimension n, then we can endow TM with the structure of a $2n$-dimensional manifold. Indeed, given a chart $\varphi : U \to V$ for M, we can define a "chart" Φ for TM on the set $\pi^{-1}(U) = \{(p,v) \in TM : p \in U\}$ by

$$\Phi(p,v) \doteqdot (\varphi^1(p), \dots, \varphi^n(p), v(\varphi^1), \dots, v(\varphi^n)) \in \mathbb{R}^{2n}\,.$$

The first n coordinates describe the point p, and the remaining n give the components v^i of the vector v with respect to the coordinate basis $\{\partial_i|_p\}_{i=1}^n$. We will often write the coordinates on M as $\varphi = (x^1, \dots, x^n)$ and the corresponding coordinates on TM as $\Phi = (x^1, \dots, x^n, \dot{x}^1, \dots, \dot{x}^n)$.

We declare a subset $V \subset TM$ open if and only if $\Phi(V \cap \pi^{-1}(U)) \subset \mathbb{R}^{2n}$ is open for each chart $\varphi : U \to \mathbb{R}^n$, where Φ is the induced chart on TM. To check that these charts give TM the structure of a differentiable manifold, we need to compute the transition maps. So consider two non-trivially overlapping charts $\varphi : U \to V$ and $\psi : W \to Z$ for M. Then the transition map $\Psi \circ \Phi^{-1}$ first takes a $2n$-tuple $(x^1, \dots, x^n, \dot{x}^1, \dots, \dot{x}^n)$ to the element $(\varphi^{-1}(x^1, \dots, x^n), \dot{x}^i \partial_i^{(\varphi)})$ of TM, then maps this to $\mathbb{R}^{2n}$ by Ψ. (We add the superscripts (φ) or (ψ) to distinguish the coordinate tangent vectors given by the chart φ from those given by the chart ψ). The first n coordinates of the result are then just $\psi \circ \varphi^{-1}(x^1, \dots, x^n)$. To compute the remaining n coordinates, we need to write $\partial_i^{(\varphi)}$ in terms of the coordinate tangent vectors $\partial_i^{(\psi)}$. So consider, for any $f \in C(M)$,

$$\begin{aligned}
\partial_i^{(\varphi)}\big|_p f &= D(f \circ \varphi^{-1})_{\varphi(p)} e_i \\
&= D((f \circ \psi^{-1}) \circ (\psi \circ \varphi^{-1}))_{\varphi(p)} e_i \\
&= D(f \circ \psi^{-1})_{\psi(p)} \circ D(\psi \circ \varphi^{-1})_{\varphi(p)} e_i \\
&= D(f \circ \psi^{-1})_{\psi(p)} (G_i^j(p) e_j) \\
&= G_i^j(p) D(f \circ \psi^{-1})_{\psi(p)} e_j \\
&= G_i^j(p) \partial_j^{(\psi)}|_p f\,,
\end{aligned}$$

where $G(p) \doteqdot D(\psi \circ \varphi^{-1})_{\varphi(p)}$. Therefore,

$$\dot{x}^i \partial_i^{(\varphi)}|_p = \dot{x}^i G_i^j \partial_j^{(\psi)}|_p = D(\psi \circ \varphi^{-1})_{\varphi(p)} \dot{x}$$

and hence

$$\Psi \circ \Phi^{-1}(x, \dot{x}) = (\psi \circ \varphi^{-1}(x), D(\psi \circ \varphi^{-1})_x \dot{x}) .$$

This is a smooth map since $\psi \circ \varphi^{-1}$ is smooth.

The tangent bundle gives rise to a natural notion of a vector field on M.

Definition 6.1. *A* ***vector field*** *on a manifold M is a smooth* ***section*** *of the tangent bundle TM. That is, a smooth map $V : M \to TM$ such that $\pi(V(p)) = p$. We will often conflate $V(p) = (p, V_p) \in TM$ with $V_p \in T_pM$*

The space of vector fields on M (smooth sections of TM) is denoted $\Gamma(TM)$. It is naturally a module over the ring $C(M)$ of smooth functions when equipped with the obvious linear structure.

In order to check whether a section $V : M \to TM$ is smooth (and hence defines a vector field), we can work locally. In a chart $\varphi : U \to \mathbb{R}^n$, V can be written as $V = V^i \partial_i$, which gives us n functions $V^1, \ldots, V^n$. It is then easy to show that V is a smooth vector field if and only if these component functions are always smooth as functions on M. That is, the vector field is smooth if, when viewed through a chart, it is smooth in the usual sense of an n-tuple of smooth functions.

Over a small region of a manifold (such as a chart) the space of smooth vector fields is in 1 to 1 correspondence with n-tuples of smooth functions. However, when looked at over the whole manifold, things are not so simple. Indeed, the hairy ball theorem says that there are no continuous vector fields on the sphere S^2 which are everywhere non-zero. On the other hand there are certainly non-zero functions on S^2 (constants, for example).

Our notion of a tangent vector as a derivation allows us to think of a vector field in another way.

Proposition 6.2. *The module $\Gamma(TM)$ of smooth vector fields over a manifold M is isomorphic to the module of derivations on $C(M)$. That is, the $\mathbb{R}$-linear operators $V : C(M) \to C(M)$ satisfying the Leibniz rule*

$$V(fg) = (Vf)g + f(Vg) \quad f, g \in C(M) .$$

Proof. Let V be a derivation on $C(M)$. Given $p \in M$, the map $V_p : C(M) \to \mathbb{R}$ given by $V_p f = (Vf)(p)$ is a derivation at p, and hence defines an element of T_pM. The map $p \mapsto V_p$ from M to TM is therefore a vector field. The identification is clearly linear with respect to $\mathbb{R}$. We need to check smoothness. The i-th component of V_p with respect to the coordinate tangent basis $\{\partial_i\}_{i=1}^n$ for a chart $\varphi : U \to \mathbb{R}^n$ is given by

$$V_p^i = V_p\varphi^i\,.$$

This is, by assumption, a smooth function of p for each i (since φ^i is a smooth function and V maps smooth functions to smooth functions – here one should really multiply φ^i by a smooth cut-off function to convert it to a smooth function on the whole of M just as we did for derivations at $p \in M$). It follows that the vector field $p \mapsto V_p$ is smooth.

Conversely, given a smooth vector field $p \mapsto V_p \in T_pM$, the map $f \to Vf$ defined by $Vf(p) = V_pf = V_p^i\partial_i|_pf$ satisfies the two conditions in the proposition and takes a smooth function to a smooth function. Linearity of the identification is again clear. It is also easy to check that composition of the two identifications is the identity map. □

6.1. The tangent bundle as a vector bundle. The tangent bundle is also equipped with a natural *vector bundle* structure.

Definition 6.3. *A (smooth)* ***vector bundle*** *over a differentiable manifold M with* ***fibre*** *$\mathbb{R}^k$ is a differentiable manifold E equipped with a smooth surjection (called a* ***bundle projection****) $\pi : E \to M$ satisfying the following conditions:*

(i) For every $p \in M$ the ***fibre*** *$E_p \doteqdot \pi^{-1}(p)$ is a real linear space of dimension k*

(ii) Every $p \in M$ admits a neighbourhood $U \subset M$ and a diffeomorphism (a ***local trivialization****) $\phi : \pi^{-1}(U) \to U \times \mathbb{R}^k$ such that*

  (a) $\pi_U(\phi(\pi^{-1}(p))) = p$, where $\pi_U :\in U \times \mathbb{R}^k \to U$ is defined by

$$\pi_U(z, v) \doteqdot z,$$

  and

  (b) $\phi|_{E_p} : E_p \to \{p\}\times\mathbb{R}^k$ is a linear isomorphism for every $p \in M$.

A (smooth) ***section*** *of a vector bundle $\pi : E \to M$ is a (smooth) map $V : M \to E$ such that $\pi(V(p)) = p$. The set of (smooth) sections of E is denoted by $\Gamma(E)$.*

In short, a vector bundle is a manifold which locally has the structure of a Cartesian product of a manifold with a vector space. Of course, this is in general not true globally. It is common to think of vector bundles as "twisted" products of a manifold with a vector space.

Example 6.4.

(i) The cylinder $S^n \times \mathbb{R}^k$ is a trivial vector bundle over S^n with fibre $\mathbb{R}^k$.

(ii) The Möbius strip $E \doteqdot (\mathbb{R} \times \mathbb{R})/\sim$, *where*[9]

$$(s,u) \sim (t,v) \quad \textit{iff} \quad t \in s + \mathbb{Z} \;\; \textit{and} \;\; v = \begin{cases} u & \textit{if } t - s = 0 \mod 2 \\ -u & \textit{if } t - s = 1 \mod 2 \end{cases}$$

is a non-trivial vector bundle over $S^1 = \mathbb{R}/\mathbb{Z}$ *with fibre* $\mathbb{R}$.

Proposition 6.5. *The tangent bundle* TM *of a manifold* M *equipped with its natural projection* $\pi : TM \to M$ *is a vector bundle.*

Proof. The projection π is clearly a surjection. To see that it is smooth, we choose a chart $\varphi : U \to \mathbb{R}^n$ for M and a corresponding chart $\Phi : \pi^{-1}(U) \to \mathbb{R}^{2n}$ for TM and observe that

$$\varphi \circ \pi \circ \Phi^{-1}(x, \dot{x}) = \varphi \circ \pi(\varphi^{-1}(x), \dot{x}^i \partial_i) = \varphi(\varphi^{-1}(x)) = x \,,$$

which is clearly smooth. Next, we define, for any chart $\varphi : U \to \mathbb{R}^n$ for M, a local trivialization $\phi : \pi^{-1}(U) \to U \times \mathbb{R}^n$ via

$$\phi(p, v) \doteqdot (p, v^i e_i) \,,$$

where $v^i = v\varphi^i$. Then condition (ii) (a) is certainly satisfied since

$$\phi(E_p) = \{p\} \times \mathbb{R}^n \,.$$

Condition (ii) (b) is also clear. Finally, we need to check that ϕ is a diffeomorphism. To see this, denote by $\widetilde{\varphi}$ the chart $\varphi \times \mathrm{Id}$ for $U \times \mathbb{R}^n$ and observe that

$$\widetilde{\varphi} \circ \phi \circ \Phi^{-1}(x, \dot{x}) = \widetilde{\varphi} \circ \phi(\varphi^{-1}(x), \dot{x}^i \partial_i) = \widetilde{\varphi}(\varphi^{-1}(x), \dot{x}^i e_i) = (x, \dot{x}) \,.$$

Hence $\widetilde{\varphi} \circ \varphi \circ \Phi^{-1}$ is the identity map. The claim follows. $\square$

We now briefly collect some basic vector bundle theory.

A smooth map $f : E_1 \to E_2$ from the total space of one vector bundle $\pi_1 : E_1 \to M_1$ to that of another $\pi_2 : E_2 \to M_2$ is called a (**vector bundle**) **homomorphism** (a.k.a. a **bundle map**) if

(1) there exists a smooth map $g : M_1 \to M_2$ such that the diagram

$$\begin{array}{ccc} E_1 & \xrightarrow{f} & E_2 \\ \pi_1 \downarrow & & \downarrow \pi_2 \\ M_1 & \xrightarrow[g]{} & M_2 \end{array}$$

commutes, and

(2) $f|_{\pi_1^{-1}(\{p\})}$ is a linear map from $\pi_1^{-1}(\{p\})$ to $\pi_2^{-1}(\{g(p)\})$ for each $p \in M_1$.

[9] The first component generates the circle $S^1 = \mathbb{R}/\mathbb{Z}$ and the second component gains a "half twist" every time the first component is traversed.

We say that the homomorphism f **covers** the map g.

A vector bundle homomorphism $f : E_1 \to E_2$ is called a (**vector bundle**) **isomoprhism** if f is a diffeomorphism and $f|_{\pi_1^{-1}(\{p\})}$ is an isomorphism for each $p \in M_1$. (Equivalently, there exists a vector bundle homomorphism $f^{-1} : E_2 \to E_1$ which is inverse to f.)

The **differential** $(p, v) \mapsto (f(p), df_p v)$ (denoted df) of a smooth map $f : M \to N$ between manifolds M and N is a homomorphism of their tangent bundles TM and TN (see Exercise 6.7).

A **subbundle** of a vector bundle $\pi : E \to M$ is a vector bundle $\pi_F : F \to M$ equipped with a vector bundle homomorphism $\iota : F \to E$ covering the identity map on M. Typically, F is a submanifold of E and ι is the inclusion map.

The *dual*, *direct sum* and *tensor product* constructions extend to vector bundles "fibrewise" — the idea is to perform the operations on the fibres and then choose the charts and trivializations accordingly. These constructions are particularly straightforward for TM. For example, the **tensor product** $TM \otimes TM$ is constructed by equipping the set

$$TM \otimes TM \doteqdot \sqcup_{p\in M} T_pM \otimes T_pM$$

with the projection $\pi : TM \otimes TM \to M$ defined by

$$\pi(p, T) \doteqdot p\,,$$

the charts $\Phi : TM \otimes TM \to \mathbb{R}^{3n}$ defined by

$$\Phi(p, T^{ij}\partial_i^{\varphi}|_p \otimes \partial_j^{\varphi}|_p) \doteqdot (\varphi(p), T^{ij} e_i \otimes e_j)\,,$$

where φ is any chart for M, and the local trivializations $\phi : \pi^{-1}(U) \to U \times \mathbb{R}^n \otimes \mathbb{R}^n$ defined[10] by

$$\phi(p, T^{ij}\partial_i^{\varphi}|_p \otimes \partial_j^{\varphi}|_p) \doteqdot (p, T^{ij} e_i \otimes e_j)\,.$$

The bundle $T^*M \doteqdot (TM)^*$ dual to TM is called the **cotangent bundle** (see Exercise 6.6).

A **covector field** on M is a smooth section $\vartheta : M \to T^*M$ of the cotangent bundle T^*M. A **tensor field** on M is a smooth section $T : M \to E$ of any bundle E obtained from TM and T^*M by finitely many direct sums and tensor products. We typically conflate $\vartheta(p) = (p, \vartheta_p) \in T^*M$ with $\vartheta_p \in T_pM$ and similarly for $T(p) = (p, T_p) \in E$ and $T_p \in E_p$. As for vector fields, smoothness of a tensor field is equivalent to smoothness of its components with respect to coordinate bases vectors ∂_{x^i} and one forms dx^i.

[10] Note that $\mathbb{R}^n \otimes \mathbb{R}^n$ is canonically isomorphic to $\mathbb{R}^{n^2}$.

In order to define vector fields on (immersed) submanifolds which are not necessarily tangential to the submanifold, it is useful to define the notion of *restriction* and *pullback* bundles.

The **restriction** $\pi_M : E_M \to M$ of a vector bundle $\pi : E \to N$ to a submanifold $M \subset N$ is defined by equipping $E_M \doteqdot \pi^{-1}(M)$ with the restriction π_M of π to E_M and the "obvious" smooth structure and local trivializations.

Similarly, given a vector bundle $\pi : E \to N$ and a smooth map $f : M \to N$, there is a vector bundle $f^*\pi : f^*E \to M$ (called the **pullback bundle**) whose fibres are $(f^*E)_p = E_{f(p)}$. One way to construct the pullback bundle is to equip the disjoint union

$$f^*E \doteqdot \{(p, e) \in M \times E : f(p) = \pi(e)\}$$

with the submanifold differentiable structure, the projection $f^*\pi$ defined by $f^*\pi(p, e) \doteqdot p$, and the local trivializations $f^*\phi : (f^*\pi)^{-1}(f^{-1}(U)) \to f^{-1}(U) \times \mathbb{R}^k$ defined by

$$\pi_{\mathbb{R}^k}(f^*\phi(p, e)) = \pi_{\mathbb{R}^k}(\phi(e)) \,,$$

where $\phi : \pi^{-1}(U) \to U \times \mathbb{R}^k$ is a local trivialization for $\pi : E \to N$ and $\pi_{\mathbb{R}^k}$ is the projection onto the linear factor.

An important example is the bundle $\gamma^*(TM)$ of tangent vectors to a manifold M along a regular curve $\gamma : I \to M$.

Given a smooth map $f : M \to N$, observe that any element ϑ of the bundle T^*N induces an element of T^*M by composing with the homomorphism $df : TM \to TN$. That is,

$$u \mapsto \vartheta_{f(p)}(df_p u) \,.$$

This construction extends to covariant tensors of any degree in the obvious way. Somewhat confusingly, both forms induced by $\vartheta \in TN$ (the element of f^*TN and the element of T^*M) are called the **pullback** of ϑ and denoted by $f^*\vartheta$.

6.2. The tensor algebra of the tangent bundle. Any vector bundle E formed from finitely many tensor products and direct sums of the bundles TM and T^*M is called a **tensor bundle**. In practice, it will suffice to consider **homogeneous** tensor bundles, which are finite tensor products of the bundles TM and T^*M. Cotangent factors are traditionally referred to as *covariant*, while tangent factors are traditionally referred to as *contravariant*. Up to a canonical isomorphism, we can always assume that the covariant factors precede the contravariant ones.

Note that the full tensor algebra of a vector bundle is no longer well-defined as a vector bundle (since it cannot be locally homeomorphic to any finite dimensional Euclidean space). On the other hand, the sections $\Gamma(E)$ of a vector bundle E over a differentiable manifold M form a finite dimensional module over the ring $C(M)$ of smooth functions on M. Duals, direct sums and tensor products of modules are defined in the same way as for linear spaces. As you might expect, there is a correspondence between the two points of view. Indeed, a tensor field $T \in \Gamma(E^* \otimes F)$ (i.e. a field of linear maps $T_p : E_p \to F_p$) induces a $C(M)$-linear map $\tilde{T} : \Gamma(E) \to \Gamma(F)$ via

$$\tilde{T}(V)(p) \doteqdot T_p(V_p)\,.$$

Conversely, a $C(M)$-linear map $\tilde{T} : \Gamma(E) \to \Gamma(F)$ induces a tensor field $T \in \Gamma(E^* \otimes F)$ via

$$T_p(v) \doteqdot \tilde{T}(V)(p)\,,$$

where $V \in \Gamma(E)$ is some extension of $v \in T_pM$. We claim that the result is independent of the choice of extension. To see this, first note that

$$\tilde{T}(V)(p) = (f\tilde{T}(V))(p) = \tilde{T}(fV)(p)\,,$$

where f is a function on M which is 1 at p and whose support lies inside the domain U of a trivialization $\phi : \pi^{-1}(U) \to U \times \mathbb{R}^k$ of E about p. On the other hand, if we write V locally in U as $V = V^i e_i$, then we may write fV *globally* as

$$fV = f\tilde{V}^i\tilde{e}_i\,,$$

where $\tilde{V}^i$ and $\tilde{e}_i$ are any smooth extensions to M of V^i and e_i which are unchanged in $\operatorname{spt} f$. Thus,

$$\begin{aligned}
\tilde{T}(V)(p) &= \tilde{T}(f\tilde{V}^i\tilde{e}_i)(p)\\
&= (f\tilde{V}^i)(p)\tilde{T}(\tilde{e}_i)(p)\\
&= V^i(p)\tilde{T}(\tilde{e}_i)(p)\\
&= v^i\tilde{T}(\tilde{e}_i)(p)\,,
\end{aligned}$$

which implies the claim.

Since these two identifications are linear and inverse to each other, we conclude that

$$\Gamma(E^* \otimes F) \cong \Gamma(E)^* \otimes \Gamma(F)\,.$$

Similar arguments show that duals and direct sums also "commute with Γ". In particular,

$$\Gamma(T^{(k,\ell)}M) \cong \bigotimes_{i=0}^{k}\Gamma(T^*M) \otimes \bigotimes_{j=0}^{\ell}\Gamma(TM)\,.$$

We define the (covariant) **tensor algebra** of M by

$$\bigoplus_{k\in\mathbb{N}_0}\bigotimes_{i=0}^{k}\Gamma(T^*M)$$

and the (mixed) **tensor algebra** of M by

$$\bigoplus_{(k,\ell)\in\mathbb{N}_0\times\mathbb{N}_0}\left(\bigotimes_{i=0}^{k}\Gamma(T^*M)\otimes\bigotimes_{j=0}^{\ell}\Gamma(TM)\right).$$

Note that, although these spaces of tensor fields may not admit bases (since they are modules, not linear spaces), it is a consequence of the universal property that every tensor field is a linear combination of simple tensor fields.

Trace operators may be defined as for the mixed tensor algebra over a linear space.

6.3. Derivations. Let M be a smooth manifold. An $\mathbb{R}$-linear map D from the (covariant or mixed) tensor algebra to itself is called a **derivation** if it satisfies the Leibniz rule:

$$D(S\otimes T)=DS\otimes T+S\otimes DT\,.$$

Graded derivations are defined as for derivations on the tensor algebra of a linear space.

Since every tensor field is a linear combinatiom of simple tensor fields, any derivation on the covariant tensor algebra is uniquely determined by its action on covector fields and functions and any derivation on the mixed tensor algebra is uniquely determined by its actions on functions, vector fields and covector fields.

A graded derivation of degree zero on the mixed tensor algebra which commutes with traces is uniquely determined by its action on functions and vector fields.

Exercises.

Exercise 6.1. *Recall that a subset $V\subset TM$ is declared to be open if and only if $\Phi(V\cap\pi^{-1}(U))\subset\mathbb{R}^{2n}$ is open for each chart $\varphi:U\to\mathbb{R}^n$, where Φ is the induced chart on TM, defined by $\Phi(x,v^i\partial_i^\varphi)\doteqdot(\varphi(x),v^ie_i)$. Show that TM is paracompact and Hausdorff.*

Exercise 6.2. *Show that the projection $\pi:TM\to M$ is a submersion.*

Exercise 6.3. *Let $V\in\Gamma(TM)$ be a vector field on a smooth manifold M. Given any chart $\varphi:U\to\mathbb{R}^n$, define the n functions $V_p^1,\dots,V_p^n$ via*

$V_p = V_p^i \partial_i|_p$. Show that V is a smooth vector field if and only if these component functions are always smooth as functions on M.

Exercise 6.4. *Show that the set $\Gamma(TM)$ of smooth vector fields (when equipped with the natural addition and scalar multiplication) forms a module over the ring $C(M)$ of smooth functions.*

Exercise 6.5. *The* ***Lie bracket*** *$[U, V]$ of two vector fields U and V is the operator defined on smooth functions $f \in C(M)$ by*

$$f \mapsto [U, V]f \doteqdot UVf - VUf\,,$$

where U and V are interpreted as derivations. Show that $[U, V]$ is a derivation (and hence defines a vector field).

Exercise 6.6. *Equip the* ***cotangent bundle***

$$T^*M \doteqdot \sqcup_{p\in M} T_p^*M \doteqdot \{(p, \vartheta) : p \in M, \vartheta \in T_p^*M\}$$

with the projection $\pi : T^*M \to M$ defined by $\pi(p, \vartheta) \doteqdot p$. Construct charts and local trivializations which make it a vector bundle.

Exercise 6.7. *Let $f : M \to N$ be a smooth map between differentiable manifolds. Show that $df : TM \to TN$ is a vector bundle homomorphism.*

7. The Lie derivative and Lie algebras

Any vector field on a differentiable manifold M is naturally associated with an ordinary differential equation: if $V \in \Gamma(TM)$ and $p \in M$, then a basic problem is to find an *integral curve* of V through the point p; that is, a smooth map $\gamma : I \to M$ for some interval I containing 0 such that

$$\begin{cases} \gamma'(t) = V_{\gamma(t)} \quad \text{for all} \quad t \in I\,, \\ \gamma(0) = p \end{cases} \tag{7.1}$$

By writing the integral curve equation with respect to a chart as a system of ODE and applying the the Picard–Lindelöf theorem, we can always find a unique solution, at least for small values of t.

Theorem 7.1. *Given $V \in \Gamma(TM)$ and $p \in M$ there exists $\delta > 0$, a neighborhood U of p in M and a unique smooth map (called a* ***local flow*** *of V) $\Psi : U \times (-\delta, \delta) \to M$ which satisfies*

$$\begin{cases} \partial_t \Psi(q,t) = V_{\Psi(q,t)} \quad \textit{for all} \quad (q,t) \in U \times (-\delta,\delta)\,, \\ \Psi(q,0) = q\,. \end{cases} \tag{7.2}$$

For each $t \in (-\delta, \delta)$ the map $\Psi_t : U \to M$ defined by $\Psi_t(y) \doteqdot \Psi(y,t)$ is a local diffeomorphism (onto its image) and

$$\Psi_s \circ \Psi_t = \Psi_{s+t}$$

whenever s, t and $s+t$ are in $(-\delta, \delta)$.

Remark 7.2. *The vector field ∂_t is the natural vector field on $U \times (-\delta, \delta)$ defined by*

$$(\partial_t f)(p,t) = \frac{d}{ds}\Big|_{s=0} f(p, t+s)\,.$$

The smoothness of Ψ as a function of q amounts to smooth dependence of solutions on their initial conditions, and we get the added bonus that the maps Ψ_t are local diffeomorphisms, and they admit a group structure: $\Psi_t \circ \Psi_s = \Psi_{s+t}$ but only for small s and t.

We note that, in general, the local flow does not extend to a "global flow" on $M \times \mathbb{R}$, except in nice cases, e.g. when M is compact (or when V is compactly supported).

The flow Ψ_t of a vector field X "flows" points of M along the integral curves of X for time t. Its differential provides an identification of the tangent spaces of M along the flow. This allows us to differentiate other vector fields: given a second vector field Y, the inverse of the isomorphism $(d\Psi_t)_p : T_pM \to T_{\Psi_t(p)}M$ brings $Y_{\Psi_t(p)}$ back to T_pM. This gives us a family of vectors in the *same* linear space, so we can compare them by differentiation.

Definition 7.3. *Let X and Y be vector fields on a manifold M. The* ***Lie derivative*** *of Y in the direction of X is the vector field $\mathcal{L}_X Y$ defined by*

$$(\mathcal{L}_X Y)(p) \doteqdot \frac{d}{dt}\Big|_{t=0}\big((d\Psi_t)_p^{-1} Y_{\Psi_t(p)}\big)\,,$$

where $\Psi : U \times I \to M$ is the local flow of X in a neighborhood U of the point p.

Note that, by the group property of the flow, $\Psi_{-t} \circ \Psi_t$ is the identity diffeomorphism and hence, by differentiating $\Psi_{-t} \circ \Psi_t$, we find that

$$(d\Psi_t)_p^{-1} = d(\Psi_{-t})_{\Psi_t(p)}\,.$$

Proposition 7.4. *Given any two vector fields $X, Y \in \Gamma(TM)$ on a manifold M,*

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

where $[\cdot,\cdot]$ denotes the Lie bracket (see Exercise 6.5).

Proof. Fix a point $p \in M$. Suppose first that $X_p \neq 0$. We will compute the Lie derivative by constructing a special chart about p in which the flow of X is particularly simple. First, take any chart $\varphi : U \to \mathbb{R}^n$ about p. Up to composition with an affine linear map, we can assume that $\varphi(p) = 0$ and $X_p = \partial_n|_p$. Set

$$\Sigma \doteqdot \varphi^{-1}\left(\{(w^1, \ldots, w^n) \in \varphi(U) : w^n = 0\}\right)\,.$$

This is a smooth $(n-1)$-dimensional submanifold of M which passes through p and is transverse to the vector field X on some neighborhood O of p (i.e. $T_q\Sigma \oplus \mathbb{R}X_q = T_qM$ for all $q \in \Sigma \cap O$).

Now consider the map $\Psi_\Sigma : \Sigma \times I \to M$ given by restricting the flow Ψ of X to $\Sigma \times I$. With respect to the natural bases $\{\partial_1, \ldots, \partial_{n-1}, \partial_t\}$ for $T_{(p,0)}(\Sigma \times \mathbb{R})$ and $\{\partial_1, \ldots, \partial_{n-1}, \partial_n = X_p\}$ for T_pM,

$$(d\Psi_\Sigma)_{(p,0)} = \begin{bmatrix} \mathrm{I}_{n-1} & 0 \\ 0 & 1 \end{bmatrix}\,.$$

In particular, $(d\Psi_\Sigma)_{(p,0)}$ is non-singular and hence, by the inverse function theorem, there is a neighborhood V of $(p,0)$ in $\Sigma \times \mathbb{R}$ on which Ψ_Σ is a diffeomorphism. So the map $\varphi_\Sigma : V \to \mathbb{R}^n$ given by

$$\varphi_\Sigma(\Psi_\Sigma(q,t)) \doteqdot \varphi(q) + te_n$$

is well-defined and forms a chart for M. The special feature of this chart is that the flow of X takes the particularly simple form

$$\begin{aligned}\varphi_\Sigma \circ \Psi_t \circ \varphi_\Sigma^{-1}(x^1, \ldots, x^n) &= \varphi_\Sigma \circ \Psi_t \circ \Psi_{x^n}\big(\varphi^{-1}(x^1, \ldots, x^{n-1}, 0)\big) \\ &= \varphi_\Sigma \circ \Psi_{x^n+t}\big(\varphi^{-1}(x^1, \ldots, x^{n-1}, 0)\big) \\ &= (x^1, \ldots, x^{n-1}, x^n + t)\,.\end{aligned}$$

In particular, $\partial_n = X$ and $d\Psi_t|_p \partial_i|_p = \partial_i|_{\Psi_t(p)}$. Thus, writing $Y = Y^i\partial_i$, we see that

$$\left((d\Psi_t)_p^{-1} Y_{\Psi_t(p)}\right)^i = Y^i(\Psi_t(p)) \,.$$

Since $X^n = 1$ and $X^i = 0$ for $i \neq n$, we find that

$$\begin{aligned}
\left.\frac{d}{dt}\right|_{t=0} \left((d\Psi_t)_p^{-1} Y_{\Psi_t(p)}\right)^i &= \left.\frac{d}{dt}\right|_{t=0} Y^i(\Psi_t(p)) \\
&= X_p Y^i \\
&= [X, Y]_p^i \,.
\end{aligned}$$

This proves the claim in case $X_p \neq 0$.

If $X_p = 0$, then $\Psi_t(p) = p$ for all t. In particular, this implies that $d\Psi_t|_p$ maps T_pM to itself and $(d\Psi_t)_p^{-1} = (d\Psi_{-t})_p$. Employing a chart $\varphi : U \to \mathbb{R}^n$ near p, we find that

$$\begin{aligned}
\mathcal{L}_X Y|_p &\doteqdot \left.\frac{d}{dt}\right|_{t=0} \left((d\Psi_t)_p^{-1} Y_{\Psi_t(p)}\right) \\
&= \left.\frac{d}{dt}\right|_{t=0} (d\Psi_{-t})_p Y_p \\
&= Y^i(p) \left.\frac{d}{dt}\right|_{t=0} (d\Psi_{-t})_p \partial_i|_p \\
&= -\, Y^i(p) \left.\frac{d}{dt}\right|_{t=0} (d\Psi_t)_p \partial_i|_p \\
&= -\, Y^i(p) \left.\frac{d}{dt}\right|_{t=0} (d\Psi_t|_p)_i^j \partial_j|_p \,.
\end{aligned}$$

The components of $d\Psi_t|_p$ may be computed as

$$\begin{aligned}
(d\Psi_t|_p)_i^j &= D(\varphi^j \circ \Psi_t \circ \varphi^{-1})|_{\varphi(p)} e_i \\
&= \left.\frac{d}{ds}\right|_{s=0} (\varphi^j \circ \Psi_t \circ \varphi^{-1})_{\varphi(p)+se_i} \,.
\end{aligned}$$

Thus, by Clairaut's theorem,

$$\begin{aligned}
\left.\frac{d}{dt}\right|_{t=0} (d\Psi_t|_p)_i^j &= \left.\frac{d}{dt}\right|_{t=0} \left.\frac{d}{ds}\right|_{s=0} (\varphi^j \circ \Psi_t \circ \varphi^{-1})_{\varphi(p)+se_i} \\
&= \left.\frac{d}{ds}\right|_{s=0} \left.\frac{d}{dt}\right|_{t=0} (\varphi^j \circ \Psi_t \circ \varphi^{-1})_{\varphi(p)+se_i} \\
&= \left.\frac{d}{ds}\right|_{s=0} X^j(\varphi^{-1}(\varphi(p) + se_i)) \\
&= \partial_i|_p X^j \,.
\end{aligned}$$

Since $X^i_p = 0$ for each $i = 1, \dots, n$, it follows that

$$\begin{aligned}\mathcal{L}_X Y|_p &= Y^i_p \partial_i|_p X^j \partial_j|_p \\ &= Y^i_p \partial_i|_p X^j \partial_j|_p - X^i_p \partial_i|_p Y^j \partial_j|_p \\ &= [X, Y]_p \,.\end{aligned}$$

This completes the proof. □

The next proposition shows that the Lie derivative behaves naturally under smooth maps. But first we need a definition

Definition 7.5. *Let $F : M \to N$ be a smooth map between differentiable manifolds M and N and let X and Y be vector fields on M. A vector field U on N is* (locally) F-related to *a vector field X on M* (near $x \in M$) *if $U_{F(y)} = dF_y(X_y)$ for all y in M (in some neighborhood of x).*

Proposition 7.6. *Let $F : M \to N$ be a smooth map between manifolds M and N and let U and V be vector fields on M. Then, for any vector fields $\tilde{U}$ and $\tilde{V}$ on N which are F-related to U and V resp.,*

$$[\tilde{U}, \tilde{V}]_{F(x)} = dF_x([U, V]_x) \,.$$

Proof. Recalling the chain rule $(dF(U))f \doteqdot U(f \circ F)$, we find, for any function $f \in C(N)$,

$$\begin{aligned}dF([U, V])f &= [U, V]f \circ F \\ &= U(Vf \circ F) - V(Uf \circ F) \\ &= U(dF(V)f) - V(dF(U)f) \\ &= U(\tilde{V}f \circ F) - V(\tilde{U}f \circ F) \\ &= dF(U)\tilde{V}f - dF(V)\tilde{U}f \\ &= \tilde{U}\tilde{V}f - \tilde{V}\tilde{U}f = [\tilde{U}, \tilde{V}]f \,.\end{aligned}$$ □

Proposition 7.7. *If $[X, Y] = 0$, then the flows of X and Y commute:*

$$\Psi_{X,t} \circ \Psi_{Y,s} = \Psi_{Y,s} \circ \Psi_{X,t} \,,$$

where $\Psi_{X,t}$ is the flow of X for time t, and similarly for Y.

Proof. Given any function $f \in C(M)$,

$$0 = [X, Y]f = \frac{\partial^2}{\partial s \partial t}(f \circ \Phi_{Y,s} \circ \Phi_{X,t} - f \circ \Phi_{X,t} \circ \Phi_{Y,s}) .$$

Since $\Psi_{X,0}(p) = \Psi_{Y,0}(p) = p$, integrating yields.

$$f \circ \Phi_{Y,s} \circ \Phi_{X,t} = f \circ \Phi_{X,t} \circ \Phi_{Y,s} \,.$$

The claim follows since f is arbitrary (and, given any two points $p, q \in M$ a function f can be found with $f(p) = 1$ and $f(q) = -1$). □

Proposition 7.8. *Every triple of vector fields X, Y and Z on a manifold M satisfies the* ***Jacobi identity***

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

Proof. See Exercise 7.2. □

Definition 7.9. *A* ***Lie algebra*** *is a (real) linear space E equipped with a skew-symmetric, (real) bilinear map $[\cdot,\cdot]:E\times E\to E$ which satisfies the Jacobi identity.*

Example 7.10 (Left invariant vector fields)**.** *Let G be a Lie group. Given $g\in G$, denote by $L_g:G\to G$ left multiplication by g. A vector field $X\in\Gamma(TG)$ is called* ***left invariant*** *if $(dL_g)_xX_x=X_{gx}$.*

Denote by $\mathfrak{g}$ the linear space of of left invariant vector fields on G. By definition, for any $U,V\in\mathfrak{g}$,

$$[U,V]_{gx}=[(dL_g)U,(dL_g)V]_x=(dL_g)_x[U,V]_x$$

and hence $[U,V]\in\mathfrak{g}$. It follows that $\mathfrak{g}$ is a Lie algebra. It is called the ***Lie algebra*** *of G.*

The Lie derivative (in the direction of a vector field X) may also be defined on tensor fields. First consider the case of covector fields. Given $p\in M$, we can relate the covector field to a family of covectors at p by exploiting the derivative of the flow of X in a manner dual to the above construction for vector fields: given a diffeomorphism $\phi:M\to N$, define the **pullback** $\phi^*\vartheta\in T_pM$ of $\vartheta T_{\phi(p)}N$ by setting

$$\phi^*\omega(u)\doteqdot\vartheta(d\phi_pu)\;\text{ for all }\;u\in T_pM\,.$$

Thus, if Ψ is the flow of X at p, then $\Psi_t^*\vartheta_{\Psi(p,t)}$ is a family of covector fields at p. So we may differentiate it to obtain the **Lie derivative**

$$(\mathcal{L}_X\vartheta)_p\doteqdot\left.\frac{d}{dt}\right|_{t=0}\Psi_t^*\vartheta\,.$$

This construction extends immediately to any covariant tensor field if we define the pullback $\phi^*T\in T_p^kM$ of $T\in T^k_{\phi(p)}N$ (by a diffeomorphism $\phi:M\to N$) in the analogous manner:

$$(\phi^*T)(u_1,\ldots,u_k)\doteqdot T(d\phi_pu_1,\ldots,d\phi_pu_k)\;\text{ for all }\;u_1,\ldots,u_k\in T_pM\,.$$

Equivalently,

$$\phi^*T=T_{i_1\ldots i_k}\phi^*\vartheta^{i_1}\otimes\cdots\otimes\phi^*\vartheta^{i_k}$$

for any basis $\{\vartheta^i\}_{i=1}^k$ for T_p^*M.

For a general homogeneous mixed tensor $T \in T^{(k,l)}_{\phi(p)}N$, we define $T \in T^{(k,l)}_p M$ by

$$T^{\phi} = T^{j_1 \dots j_l}_{i_1 \dots i_k} \phi^* \vartheta^{i_1} \otimes \cdots \otimes \phi^* \vartheta^{i_k} \otimes d\phi_p e_{j_1} \otimes \cdots \otimes d\phi_p e_{j_l}$$

for any bases $\{e_i\}_{i=1}^n$ for T_pM and $\{\vartheta^i\}_{i=1}^k$ for T^*_pM. The Lie derivative of $T \in \Gamma(T^{(k,l)}M)$ is then

$$(\mathcal{L}_X T)_p \doteqdot \left.\frac{d}{dt}\right|_{t=0} T^{\Psi_t}.$$

Of course, this extends to a general tensor by distributing over linear combinations.

Proposition 7.11. *Given any $X \in \Gamma(TM)$, the Lie derivative $\mathcal{L}_X$ is a graded derivation of degree zero which commutes with traces.*

Proof. Denote by Ψ the flow of X about a point p. Since

$$(S \otimes T)^{\Psi_t} = S^{\Psi_t} \otimes T^{\Psi_t}$$

we immediately obtain

$$\mathcal{L}_X(S \otimes T) = \mathcal{L}_X S \otimes T + S \otimes \mathcal{L}_X T.$$

So, since $\mathcal{L}_X$ is certainly $\mathbb{R}$-linear, it is a derivation. It is also clear that it is graded of degree zero. We leave the final step as an easy exercise. □

Exercises.

Exercise 7.1. *Show that*

$$[\partial_i, \partial_j] = 0$$

for any two coordinate vector fields ∂_i and ∂_j. Deduce that

$$[X, Y] = (X^i \partial_i Y^j - Y^i \partial_i X^j)\partial_j$$

for any two vector fields X and Y.

Exercise 7.2. *. Let X, Y and Z be vector fields on a manifold M. Show that*

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0.$$

Hint: By Proposition 7.6,

$$d\Psi_t([X, Y]) = [d\Psi_t(X), d\Psi_t(Y)],$$

where Ψ is the flow of Z.

Exercise 7.3. *Let G be a Lie group. Denote by $\mathfrak{g}$ its Lie algebra. Show that the map*

$$v \in T_e M \mapsto V \in \mathfrak{g}$$

defined by

$$V_g \doteqdot (dL_g)_e v$$

is an isomorphism.

Exercise 7.4. *Let G and H be Lie groups, with respective Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. Let $\phi : G \to H$ be a smooth group homomorphism. Prove the following statements.*

(a) *The induced map $\phi_* : \mathfrak{g} \to \mathfrak{h}$ is an algebra homomorphism.*

(b) *If ϕ is a diffeomorphism, then ϕ_* is an isomorphism.*

(c) *If ϕ is a submersion, then $\ker \phi_*$ is a Lie subalgebra of $\mathfrak{g}$*

Exercise 7.5. *Suppose that H is a* ***Lie subgroup*** *of G (i.e. a subgroup which is also a submanifold). Show that $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$.*

8. Frobenius' theorem

We have seen that a vector field X on a manifold M corresponds to first order differential equation. By Theorem 7.1, X is tangent to a unique integral curve through each point $p \in M$. If $X(p) \neq 0$, there is a neighbourhood U of p such that each integrable curve is an embedded, 1-dimensional submanifold.

Given a set of k vector fields $E_1, \ldots, E_k$ on M and a point $p \in M$, it is natural to seek a map $F : \mathbb{R}^k \to M$ for which $F(0) = p$ and $dF(\partial_i) = E_i$ for each $i = 1, \ldots, k$, where $\{\partial_i\}_{i=1}^k$ denote the canonical coordinate vector fields on $\mathbb{R}^k$. We attempt to construct such a map naïvely as follows. By Theorem 7.1, we can arrange that $dF(\partial_k) = E_k$ by integrating E_k; that is, by setting $F(te_k) \doteqdot \Psi_k(x,t)$ for t sufficiently small, where Ψ_i is the local flow of E_i. Next, we follow the integral curves of the vector field E_{k-1} from $F(te_k)$ for time s to obtain $F(te_k + se_{k-1})$; that is, we set $F(te_k + se_{k-1}) \doteqdot \Psi_{k-1}(\Psi_k(x,t), s)$, and so on following the flows of $E_{k-2}, \ldots, E_2$, and finally E_1. Unfortunately, that doesn't always work.

Example 8.1. *Consider the vector fields $E_1 \doteqdot \partial_1$ and $E_2(x_1, x_2) \doteqdot (1 + x_1)\partial_2$ on $\mathbb{R}^2$. Following the recipe outlined above, we set $F(0,0) \doteqdot (0,0)$, follow the flow of E_2 to get $F(0,t) = (0,t)$, and then follow the flow of E_1 to get $F(s,t) = (s,t)$. This gives $dF(\partial_1) = E_1$ but not $dF(\partial_2) = E_2$. Note that if we change our procedure by first integrating along the vector field E_1, and then the vector field E_2, then we don't get the same result: Instead, we get $F(s,t) = (s, t + st)$. In fact, there is an easy way to tell that this could not have worked: If there was such a map, then we would have*

$$[E_1, E_2] = [dF(\partial_1), dF(\partial_2)] = dF([\partial_1, \partial_2]) = 0\,.$$

But in the example we have $[E_1, E_2] = \partial_2 \neq 0$. The following proposition shows that this is the only local obstruction to constructing such a map.

Proposition 8.2. *Suppose that the vector fields $E_1, ..., E_k \in \Gamma(M)$ commute pairwise: $[E_i, E_j] = 0$ for $i, j = 1, \ldots, k$. Then for each $p \in M$ there exists a neighbourhood U of 0 in $\mathbb{R}^k$ and a unique smooth map $F : U \to M$ satisfying $F(0) = p$ and $(dF)_x(\partial_i) = (E_i)_{F(x)}$ for every $x \in U$ and $i \in \{1, ..., k\}$.*

Proof. We construct the map F exactly as outlined in the example above. In other words, we set

$$F(x^1, \ldots, y^k) = \Psi_{1,y^1} \circ \cdots \circ \Psi_{1,y^1}(p)\,,$$

where $\Psi_{i,t}$ is the flow of the vector field E_i for time t. This gives immediately that $dF(\partial_1) = E_1$ everywhere. But since, by Proposition 7.7, the flows commute, we also obtain $dF(\partial_i) = E_i$ for each $i = 2, \ldots, k$. $\square$

More generally, we may consider subbundles of the tangent bundle.

Definition 8.3. *A* ***distribution*** *on M is a subbundle D of the tangent bundle TM. A distribution D is* ***involutive*** *if $[X,Y] \in \Gamma(D)$ whenever $X, Y \in D$ and* ***integrable*** *if for each $p \in M$ there exists a submanifold $\Sigma \subset M$ such that $D_x = T_x\Sigma$.*

Clearly, an integrable distribution is involutive. Frobenius' theorem states that the converse is true.

Theorem 8.4 (Frobenius' Theorem, first version)**.** *A distribution is integrable if and only if it is involutive.*

Proof. We need only prove the sufficiency of involutivity. Let $D \subset TM$ be an involutive distribution (of dimension $m \leq n$, say). Given $p \in M$, we want to construct a submanifold through p tangent to the distribution. Choose a chart $\phi : U \to \mathbb{R}^n$ for M about p such that D_p is the subspace of T_pM generated by the first m coordinate tangent vectors. By the implicit function theorem, passing to a smaller neighbourhood if necessary, we can find smooth functions $a_i^j : U \to \mathbb{R}$, $i = 1, \ldots, m$, $j = m+1, \ldots, n$, such that

$$D_q = \left\{ \sum_{i=1}^m c^i \left(\partial_i|_q + \sum_{j=m+1}^n a_i^j(q)\partial_j|_q \right) : (c^1, \ldots, c^m) \in \mathbb{R}^m \right\}$$

for each $q \in U$.

Define, for each $i = 1, \ldots, m$,

$$E_i \doteqdot \partial_i + a_i^j \partial_j \,.$$

Then $E_i(q) \in D_q$ for each $q \in U$ and

$$\begin{aligned} [E_i, E_j] &= \left[\partial_i + \sum_{k=m+1}^n a_i^k \partial_k ,\, \partial_j + \sum_{k=m+1}^n a_j^k \partial_k \right] \\ &= \sum_{k=m+1}^n \left((\partial_i a_j^k - \partial_j a_i^k) + \sum_{\ell=m+1}^n (a_i^\ell \partial_\ell a_j^k - a_j^\ell \partial_\ell a_i^k) \right) \partial_k \end{aligned}$$

On the other hand, $[E_i, E_j] = C_{ij}^k E_k$ for some functions C_{ij}^k, $i, j, k = 1, \ldots, m$ since, by hypothesis, $[E_i, E_j] \in \Gamma(D)$. Since $[E_i, E_j]$ has no component in the direction ∂_k for $k = 1, \ldots, m$, we conclude that $C_{ij}^k \equiv 0$ for each i, j, k.

Proposition 8.2 now gives the existence of a map F from a region of $\mathbb{R}^m$ into M with $dF(\partial_i) = E_i$ for $i = 1, \ldots, m$. In particular, dF is of full rank, and hence locally an embedding. So passing to a smaller neighbourhood if necessary completes the proof. $\square$

9. Differential forms and the exterior calculus

The exterior algebra $\Lambda(M)$ of a differentiable manifold M is the graded algebra of totally skew-symmetric covariant tensor fields. That is, $\Lambda(M) = \Lambda(\Gamma(T^*M))$ equipped with the wedge product. Elements of $\Lambda(M)$ are called **differential forms**.

The subspace $\Lambda^1(M)$ is canonically isomorphic to $\Gamma(T^*M)$. Its elements, the covector fields, are in this context called **differential one-forms**.

The **differential** $df \in \Lambda^1(M)$ of a smooth function $f \in C(M)$ is defined by

$$df(V) \doteqdot Vf$$

for any $V \in \Gamma(TM)$. The differential at p is the one-form $df_p \in \Lambda^1(T_pM) \cong T_p^*M$ defined by

$$df_p(V_p) \doteqdot V_pf = df(V)(p).$$

Note that the differential df_p is well-defined even if f is only defined on a neighbourhood of p. In particular, given a chart $\varphi = (x^1, \dots, x^n) : U \to \mathbb{R}^n$ for M, we can define the differentials $\{dx^i\}_{i=1}^n$. At each $p \in U$, $\{dx_p^i\}_{i=1}^n$ is the basis dual to the coordinate basis $\{\partial_i|_p\}_{i=1}^n$. Indeed,

$$dx^j(\partial_i) \doteqdot \partial_i x^j \doteqdot \frac{\partial}{\partial x^i}(x^j \circ \varphi^{-1}) = \delta_i^j.$$

Observe that

$$df(V) = Vf = V^i\partial_i f = dx^i(V)\partial_i f.$$

So

$$df = \partial_i f dx^i.$$

The differential d can be (uniquely) extended to a **closed**, **graded antiderivation** $d : \Lambda(M) \to \Lambda(M)$ of **degree** $+1$.

Proposition 9.1. *There exists a unique graded antiderivation $d : \Lambda(M) \to \Lambda(M)$ of degree $+1$ which is closed (i.e. $d \circ d = 0$) and agrees with the differential df when $f \in \Lambda^0(M)$. This derivation is called the **exterior derivative**.*

Proof. We first determine (uniquely) the action of d on 1-forms $\omega \in \Lambda^1(M)$. Given local coordinates $x : U \to \mathbb{R}^n$ for M^n, we have $\omega|_U = \omega_i dx^i$. Then for any $p \in U$ we obtain, by asserting $\mathbb{R}$-linearity, the Leibniz rule, and closedness of d, the formula

$$\begin{aligned}
(d\omega)_p &= [d(\omega_i dx^i)]_p \\
&= (\partial_j\omega_i dx^j \wedge dx^i)|_p - (d\omega_i)_p \wedge (ddx^i)_p \\
&= (\partial_j\omega_i dx^j \wedge dx^i)|_p.
\end{aligned}$$

This gives the expression

$$d\omega = \sum_{i<j} (\partial_j \omega_i - \partial_i \omega_j) dx^j \wedge dx^i$$

in terms of the local basis $\{dx^i \wedge dx^j\}_{i<j}$ for $\Lambda^2(M)$, which uniquely determines $d : \Lambda^1(M) \to \Lambda^2(M)$.

The Leibniz rule and $\mathbb{R}$-linearity then uniquely determine $d\omega$ for any $\omega \in \Lambda(M)$: If $\omega = \omega^1 \wedge \cdots \wedge \omega^k \in \Lambda^k(M)$ is a homogeneous product of 1-forms ω^i, then the Leibniz rule yields the formula

$$d\omega = \sum_{i=1}^{k} (-1)^{i+1} \omega^1 \wedge \cdots \wedge \omega^{i-1} \wedge d\omega^i \wedge \omega^{i+1} \wedge \cdots \wedge \omega^k .$$

Since d is uniquely determined on one-forms, this uniquely determines d on homogeneous forms. The $\mathbb{R}$-linearity then uniquely determines d on every $\omega \in \Lambda(M)$. □

A differential form $\omega \in \Lambda(M)$ satisfying $d\omega = 0$ is called **closed** and a differential form $\omega \in \Lambda(M)$ satisfying $\omega = d\phi$ for some $\phi \in \Lambda(M)$ is called **exact**. Evidently, exact forms are closed. That the converse is not necessarily true is a fundamental observation of *differential topology*.

Proposition 9.2. *The exterior derivative commutes with pullbacks: given a smooth map $F : M \to N$,*

$$d(F^*\omega) = F^* d\omega .$$

Proof. First note that, for any function $f \in C(N)$, the chain rule yields

$$d(F^*f) = d(f \circ F) = df \circ dF = F^* df .$$

For a general homogeneous $\omega \in \Lambda(M)$, we compute in local coordinates

$$\begin{aligned} d(F^*\omega) &= d\left[F^*(\omega_{i_1 \ldots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k})\right] \\ &= d\left(F^*\omega_{i_1 \ldots i_k} F^* dx^{i_1} \wedge \cdots \wedge F^* dx^{i_k}\right) \\ &= d(F^*\omega_{i_1 \ldots i_k}) \wedge F^* dx^{i_1} \wedge \cdots \wedge F^* dx^{i_k} \\ &\quad + F^*\omega_{i_1 \ldots i_k} d\left(d(F^* x^{i_1}) \wedge \cdots \wedge d(F^* x^{i_k})\right) \\ &= (F^* d\omega_{i_1 \ldots i_k}) \wedge F^* dx^{i_1} \wedge \cdots \wedge F^* dx^{i_k} \\ &= F^*(d\omega_{i_1 \ldots i_k} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}) \\ &= F^* d\omega . \end{aligned}$$

This completes the proof. □

Proposition 9.3 (Cartan's formula)**.** *Given a vector field $U \in \Gamma(TM)$, the restriction of the Lie derivative $\mathcal{L}_U$ to the exterior algebra $\Lambda(M)$ satisfies*

$$\mathcal{L}_U = \iota_U \circ d + d \circ \iota_U .$$

Proof. Since the Lie derivative satisfies the Leibniz rule with respect to tensor products, this is also the case for wedge products:

$$\mathcal{L}_U(\omega \wedge \phi) = \mathcal{L}_U\omega \wedge \phi + \omega \wedge \mathcal{L}_U\phi\,.$$

That is, $\mathcal{L}_U$ is a graded derivation of degree zero. So it suffices to prove the stated identity on smooth functions and one forms. The claim is easily checked for functions. For a one-fom ω, Exercise 9.1 yields

$$\begin{aligned}\iota_V(\mathcal{L}_U\omega) = (\mathcal{L}_U\omega)(V) &= \mathcal{L}_U(\omega(V)) - \omega(\mathcal{L}_U V)\\ &= U(\omega(V)) - \omega([U,V])\\ &= d\omega(U,V) + V(\omega(U))\\ &= \iota_V\iota_U d\omega + \iota_V(d(\omega(U)))\\ &= \iota_V\left(\iota_U d\omega + d\iota_U\omega\right)\,.\end{aligned}$$

The claim follows since V is arbitrary. □

Corollary 9.4. *The Lie derivative commutes with the exterior derivative.*

Proof. Since d is closed, Cartan's formula yields

$$\mathcal{L}_U \circ d = d \circ \iota_U \circ d = d \circ \mathcal{L}_U.$$ □

Differential forms allow an alternative formulation of Frobenius' theorem. In order to state it, we associate, to a given a distribution D on M, the subspace $\Lambda_0(D) \subset \Lambda(M)$ consisting of differential forms which vanish when restricted to D. This subspace is closed under multiplication by smooth functions and under wedge products.

Theorem 9.5 (Frobenius' theorem, second version)**.** *A distribution D is integrable if and only if $\Lambda_0(D)$ is closed under exterior differentiation.*

Proof. If D is integrable, then it is involutive. The formula for the exterior derivative in Exercise 9.2 then implies that $\Lambda_0(D)$ is closed under exterior differentiation.

On the other hand, if D is not integrable, then, by Frobenius' theorem, we can find vector fields X and Y in $\Gamma(D)$ and $p \in M$ such that $[X,Y]_p \notin D_p$. Then there exists a 1-form $\omega \in \Lambda_0(D)$ satisfying $\omega_p([X,Y]_p) = 1$. But ω satisfies

$$\begin{aligned}d\omega(X,Y) &= X\omega(Y) - Y\omega(X) - \omega([X,Y])\\ &= -\,\omega([X,Y])\\ &\neq 0\end{aligned}$$

at p. So $\Lambda_0(D)$ is not closed under exterior differentiation. □

Exercises.

Exercise 9.1. *Let $\omega \in \Lambda(M)$ be a differential one-form. Show that*

$$d\omega(U,V) = U(\omega(V)) - V(\omega(U)) - \omega([U,V])\,. \tag{9.1}$$

for any $U, V \in \Gamma(TM)$.

Exercise 9.2. *Let $\omega \in \Lambda^k(M)$ be a differential k-form. Show that*

$$\begin{aligned} d\omega(U_0, U_1, \ldots, U_k) = &\sum_{i=0}^{k} (-1)^i U_i \omega\left(U_0, \ldots, U_{i-1}, U_{i+1}, \ldots, U_k\right) \\ &+ \sum_{0 \le i < j \le k} (-1)^{i+j} \omega\left([U_i, U_j], U_0, \ldots, U_{i-1}, U_{i+1}, \ldots\right. \\ &\qquad\qquad \left. \ldots, U_{j-1}, U_{j+1}, \ldots, U_k\right) \end{aligned} \tag{9.2}$$

for any $U_0, \ldots, U_k \in \Gamma(TM)$.

Observe that the Hodge star map $*$ on the *orthogonal space* $\mathbb{R}^n$ induces a map (which we also call the **Hodge star**) on the exterior algebra of the *manifold* $\mathbb{R}^n$ by setting

$$*1|_p \doteqdot dx^1|_p \wedge \cdots \wedge dx^n|_p\,.$$

Exercise 9.3. *Recall that the gradient* $\operatorname{grad} f$ *of a function* $f : \mathbb{R}^3 \to \mathbb{R}$ *and the divergence* $\operatorname{div} U$ *and curl* $\operatorname{curl} U$ *of a vector field* $U : \mathbb{R}^3 \to \mathbb{R}^3$ *are defined by*

$$\operatorname{grad} f_p \cdot u \doteqdot D_u f_p \;\text{ for any }\; u \in \mathbb{R}^3,$$

$$\operatorname{div} U \doteqdot \sum_{i=1}^{n} \frac{\partial U^i}{\partial x^i}$$

and

$$\operatorname{curl} U \doteqdot \left(\frac{\partial U^2}{\partial x^3} - \frac{\partial U^3}{\partial x^2}, \frac{\partial U^3}{\partial x^1} - \frac{\partial U^1}{\partial x^3}, \frac{\partial U^1}{\partial x^2} - \frac{\partial U^2}{\partial x^1}\right).$$

Recall also that the cross product $\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ is defined by

$$u \times v \doteqdot (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1)\,.$$

By relating each of these operations to the exterior derivative $d : \Lambda(\mathbb{R}^3) \to \Lambda(\mathbb{R}^3)$, the wedge product $\wedge : \Lambda(\mathbb{R}^3) \times \Lambda(\mathbb{R}^3) \to \Lambda(\mathbb{R}^3)$, and the Hodge star $* : \Lambda^k(\mathbb{R}^3) \to \Lambda^{n-k}(\mathbb{R}^3)$, prove the following identities:

$$\operatorname{curl} \operatorname{grad} f = 0$$

$$\operatorname{div}(\mathbf{u} \times \mathbf{v}) = \mathbf{v} \cdot \operatorname{curl} \mathbf{u} - \mathbf{u} \cdot \operatorname{curl} \mathbf{v}$$

$$\operatorname{div} \operatorname{curl} \mathbf{v} = 0$$

$$\operatorname{curl}(f\mathbf{v}) = \operatorname{grad} f \times \mathbf{v} + f \operatorname{curl} \mathbf{v}$$

$$\operatorname{div}(\operatorname{grad} f \times \operatorname{grad} g) = 0\,.$$

$$\operatorname{div}(g \operatorname{grad} f \times f \operatorname{grad} g) = 0 \, .$$
$$\operatorname{curl}(\operatorname{curl} \mathbf{u}) = \operatorname{grad}(\operatorname{div} \mathbf{v}) - \operatorname{div}(\operatorname{grad} \mathbf{v}) \, .$$

Show that any solution $\mathbf{A}$ *to the* ***Hemholtz equation***

$$\operatorname{curl} \operatorname{curl} \mathbf{A} = \mathbf{A}$$

automatically satisfies the ***vector Hemholtz equation***

$$-\Delta \mathbf{A} = \mathbf{A}$$

and the ***solenoidal condition***

$$\operatorname{div} \mathbf{A} = 0 \, .$$

10. Orientability, integration, and Stokes' Theorem

Definition 10.1. *An atlas $\mathcal{A}$ for a differentiable manifold M is* ***oriented*** *if, for each pair of charts $\phi : U \to \mathbb{R}^n$ and $\eta : V \to \mathbb{R}^n$ in $\mathcal{A}$ satisfying $U \cap V \neq 0$, the Jacobian determinant $\det d(\eta \circ \phi^{-1})$ is positive. A differentiable manifold is* ***orientable*** *if it admits such an atlas. An* ***orientation*** *on an orientable manifold is an equivalence class of oriented atlases, where two oriented atlases are equivalent if their union is an oriented atlas.*

Example 10.2. *A submanifold of dimension n in $\mathbb{R}^{n+1}$ is called a hypersurface. An orientation on a hypersurface M is equivalent to the choice of a unit normal vector continuously over the whole of M: Given an orientation on the hypersurface, choose the unit normal ν such that for any chart ϕ in the oriented atlas for M,*

$$\det[\partial_1, \ldots, \partial_n, \nu] > 0 . \tag{10.1}$$

This is continuous on M since it is continuous on overlaps of charts. Conversely, given ν chosen continuously over all of ν, we choose an atlas for M consisting of all those charts for which (10.1) *holds.*

There is a useful relationship between orientability of a differentiable manifold M^n and the space of n-forms $\Lambda^n(M)$:

Proposition 10.3. *A differentiable manifold M^n is orientable if and only if there exists an n-form $\omega \in \Lambda^n(M^n)$ which is nowhere vanishing on M^n.*

Proof. Suppose there exists a nowhere vanishing n-form ω. Let $\mathcal{A}$ be the set of charts ϕ for M for which $\omega(\partial_1, \ldots, \partial_n) > 0$. Then $\mathcal{A}$ is an atlas for M, since any chart for M is either in $\mathcal{A}$ or has its composition with a reflection in $\mathcal{A}$. We claim that $\mathcal{A}$ is oriented: For any pair of charts ϕ and η in $\mathcal{A}$ (with non-trivial overlap),

$$\partial_{\eta^i} = d(\eta^{-1} \circ \phi)(\partial_{\phi^j})$$

and hence, by the linearity and skew-symmetry of ω,

$$\omega(\partial_{\eta^1}, \ldots, \partial_{\eta^n}) = \det\left[d(\eta \circ \phi^{-1})\right]\omega(\partial_{\phi^1}, \ldots, \partial_{\phi^n}) . \tag{10.2}$$

Thus,

$$\det\left[d(\eta \circ \phi^{-1})\right] > 0 .$$

Conversely, suppose that M admits an oriented atlas $\mathcal{A} = \{\phi_\alpha : U_\alpha \to V_\alpha\}_{\alpha \in I}$. Let $\{\rho_\beta\}_{\beta \in J}$ be a partition of unity subordinate to the cover $\{U_\alpha\}_{\alpha \in I}$ so that, for each $\beta \in J$, there exists $\alpha(\beta) \in I$ such that $\mathrm{supp}\rho_\beta \subset U_{\alpha(\beta)}$. Define

$$\omega \doteqdot \sum_{\beta \in J} \rho_\beta \, d\phi^1_{\alpha(\beta)} \wedge \cdots \wedge d\phi^n_{\alpha(\beta)} .$$

Then ω is nowhere vanishing. $\square$

Orientations for manifolds-with-boundary are defined in the same way as for manifolds. An orientation on a manifold-with-boundary induces a canonical orientation on the boundary.

Proposition 10.4. *Let M be a manifold-with-boundary. If* $\operatorname{int}(M)$ *is oriented, then so is ∂M.*

Proof. Let $\mathcal{A}$ be an oriented boundary atlas for M. Then the corresponding atlas for ∂M is automatically oriented: Any pair of overlapping oriented boundary charts for M map $\mathbb{R}^n_+$ to $\mathbb{R}^n_+$, and the derivative of a transition map on the boundary must have the form

$$d(\eta \circ \phi^{-1}) = \begin{bmatrix} d(\eta_0 \circ \phi_0^{-1}) & * \\ 0 & a \end{bmatrix}$$

where η_0 and ϕ_0 are the restrictions of η and ϕ, respectively, to the boundary and $a \doteqdot \left\langle d(\eta \circ \phi^{-1})(e_{n+1}), e_{n+1} \right\rangle > 0$. The claim follows. $\square$

In the proof of the Proposition above, we ignored the case $n = 1$ — the boundary of a 1-dimensional manifold is a 0-dimensional manifold (i.e. a collection of points). What does it mean to define an orientation on a zero-dimensional manifold? Our original definition clearly makes no sense in that case. However, the equivalent definition in terms of non-vanishing n-forms does make sense: We will say that a 0-manifold N is oriented if it is equipped with a function (i.e. a 0-form) from N to $\mathbb{Z}_2 = \{-1, 1\}$. In this case we also have to allow boundary charts for 1-manifolds which map to $(-\infty, 0]$ as well as $[0, \infty)$ (in higher dimensions, we can always transform charts into any half-plane via an orientation-preserving map to map into the upper half-plane, but not if $n = 1$).

Next, we shall introduce a notion of integration using differential forms. A key point to keep in mind here is that none of our definitions depend on us having any notion of volume, surface area or length. Nevertheless, the structure of differential forms is exactly what is required to produce a well-defined notion of integration.

Let M^n be a compact, oriented differentiable manifold-with-boundary. We define the integral $\int_M \omega$ of any $\omega \in \Lambda^n(M)$ as follows: Let $\{\rho_\alpha : \alpha \in I\}$ be a partition of unity subordinate to an oriented boundary atlas for M, so that for each α there exists an oriented chart (either a regular chart or a boundary chart) $\phi_\alpha : U_\alpha \to \mathbb{R}^n$ for M, such that $\operatorname{supp}\rho_\alpha \subset U_\alpha$. Then

$$\int_M \omega \doteqdot \sum_{\alpha \in I} \int_{\phi_\alpha(U_\alpha)} (\phi_\alpha^{-1})^*(\rho_\alpha \omega)(e_1, \ldots, e_n)\, dx^1 \ldots dx^n.$$

By (10.2) and the area formula (i.e. the change of variables formula for integrals), the integral is well-defined.

Now we are in a position to prove the fundamental result concerning integration of forms on manifolds. This will also give us a new geometric interpretation of the exterior derivative.

Theorem 10.5. *Let M^n be a compact oriented differentiable manifold-with-boundary. Then, for any $(n-1)$-form $\omega \in \Lambda^{n-1}(M)$,*

$$\int_M d\omega = \int_{\partial M} \omega\,,$$

where the integral on the right-hand side is taken using the induced orientation on ∂M, integrating the restriction of ω to ∂M (i.e. the pull-back of ω by the inclusion map).

In particular, if M is a compact manifold (without boundary), then the integral of the exterior derivative of any $(n-1)$-form is zero.

Proof. Let $\{\rho_\alpha\}_\alpha$ be a partition of unity on M with each ρ_α supported in a chart $\phi_\alpha : U_\alpha \to \mathbb{R}^n$. Then, for any $(n-1)$-form ω,

$$\begin{aligned}
\int_M d\omega &= \int_M d\left(\sum_\alpha \rho_\alpha \omega\right) \\
&= \sum_\alpha \int_{\phi_\alpha(U_\alpha)} \big[(\phi_\alpha^{-1})^* d(\rho_\alpha\omega)\big](e_1,\ldots,e_n)dx^1\ldots dx^n \\
&= \sum_\alpha \int_{\phi_\alpha(U_\alpha)} d\big[(\phi_\alpha^{-1})^*(\rho_\alpha\omega)\big](e_1,\ldots,e_n)dx^1\ldots dx^n\,.
\end{aligned}$$

Writing ω locally as

$$\omega = \sum_{j=1}^n \omega_j dx^1 \wedge \cdots \wedge dx^{j-1} \wedge dx^{j+1} \wedge \cdots \wedge dx^n\,,$$

we obtain

$$d\big[(\phi_\alpha^{-1})^*(\rho_\alpha\omega)\big](e_1,\ldots,e_n) = \sum_{j=1}^n \frac{\partial(\rho_\alpha\omega_j)}{\partial x^j}\,.$$

If ϕ_α is an interior chart, then $\phi_\alpha(U_\alpha) \subset \mathbb{R}^n$ is open. Since $\rho_\alpha = 0$ on the boundary of $\phi_\alpha(U_\alpha)$, Fubini's theorem and the fundamental theorem of calculus imply that

$$\sum_{j=1}^n \int_{\phi_\alpha(U_\alpha)} \frac{\partial(\rho_\alpha\omega_j)}{\partial x^j}\,dx^1\ldots dx^n = 0.$$

If ϕ_α is a boundary chart, then we instead obtain

$$\sum_{j=1}^{n}\int_{\phi_\alpha(U_\alpha)}\frac{\partial(\rho_\alpha\omega_j)}{\partial x^j}\,dx^1\ldots dx^n = \int_{\mathbb{R}^{n-1}\times\{0\}}\int_0^\infty \frac{\partial(\rho_\alpha\omega_n)}{\partial x^n}\,dx^n dx^1\ldots dx^{n-1}$$
$$= \int_{\mathbb{R}^{n-1}\times\{0\}}\rho_\alpha\omega_n\,dx^1\ldots dx^{n-1}.$$

The claim follows. □

Exercises.

Exercise 10.1. *Show that every one-dimensional manifold is orientable.*

Exercise 10.2. *Show that every connected manifold has either zero or two orientations.*

Exercise 10.3. *Suppose that $F:\mathbb{R}^{n+1}\to\mathbb{R}$ has non-zero derivative everywhere on $M \doteqdot F^{-1}(0)$. Show that M is orientable.*

Exercise 10.4. *Use* (10.2) *and the area formula (i.e. the change of variables formula for integrals) to prove that integration of n-forms is well-defined.*

11. Connections

We have seen that the differentiable structure of a manifold gives rise to a natural notion of differentiation of smooth functions, encoded in the structure of the tangent bundle. On $\mathbb{R}^n$ (viewed as a manifold with the differentiable structure induced by the identity chart) it is possible to introduce a notion of *directional derivative* of vector fields $V \in \Gamma(T\mathbb{R}^n)$. However, this is only because $\mathbb{R}^n$ admits an additional canonical structure — namely, $\mathbb{R}^n$ admits a global *parallelism*, which allows us to identify different tangent spaces $T_p\mathbb{R}^n$. This is the family $\tau : \mathbb{R}^n \times \mathbb{R}^n \to T\mathbb{R}^n$ of *parallel translation* maps $\tau_p \doteqdot \tau(p,\cdot) : \mathbb{R}^n \to T_pM$ defined by

$$(p, v^i e_i) \mapsto \tau_p(v^i e_i) \doteqdot v^i \partial_i|_p \,, \tag{11.1}$$

where $\{e_i\}_{i=1}^n$ are the standard basis vectors for $\mathbb{R}^n$ and, for each $p \in \mathbb{R}^n$, $\{\partial_j|_p\}_{i=1}^n$ is the coordinate basis for $T_p\mathbb{R}^n$ with respect to the identity chart. The *directional derivative* D_uV of a vector field $V : \mathbb{R}^n \to T\mathbb{R}^n$ in the direction of a vector $v \in T_p\mathbb{R}^n$ is then defined by

$$D_uV \doteqdot \tau_p\left[\frac{d}{dt}\Big|_{t=0}\left(\tau^{-1}_{\omega(t)}V_{\omega(t)}\right)\right] , \tag{11.2}$$

where ω is any curve through p such that $\omega'(0) = u$ (that is, $u = [\omega]$) and the subtraction is with respect to the affine structure of $T_p\mathbb{R}^n$. Thus, writing V with respect to the local field of bases $\{\partial_i\}_{i=1}^n$ as $V^i\partial_i$,

$$D_uV = (D_uV^i)\partial_i|_p \,,$$

where D_uV^i is the directional derivative of the function V^i in the direction of the vector u.

In an abstract manifold, there is no such canonical identification of tangent spaces at different points and hence no canonical notion of directional derivative of a vector field[11] — different choices for the identification will in general give rise to distinct directional derivatives. Instead of first introducing such a "parallelism" and using it to define a directional derivative as in (11.2), we cut straight to the chase and introduce an abstract notion of differentiation of vector fields directly. As it turns out, the two points of view are equivalent.

Definition 11.1. *A* ***connection*** *on a differentiable manifold M is a map $\nabla : TM \times \Gamma(TM) \to TM$, which we write as ∇_uV instead of $\nabla(u, V)$, which satisfies the following properties:*

[11] Note that the Lie derivative does *not* induce a suitable directional derivative: since $\mathcal{L}_UV|_p$ depends on U in a neighborhood of p, the "directional derivative" $\mathcal{L}_uV \doteqdot \mathcal{L}_UV|_p$ defined via some local extension U of $u \in T_pM$, say, depends on the extension.

(1) *$u \mapsto \nabla_u V$ is a bundle map (covering the identity): $\pi(\nabla_u V) = \pi(u)$ and*
$$\nabla_{\mu u+v} V = \mu \nabla_u V + \nabla_v V$$
for any $V \in \Gamma(TM)$, $u, v \in TM$ and $\mu \in \mathbb{R}$.

(2) *$V \mapsto \nabla_u V$ is a derivation:*
$$\nabla_u(\mu V + W) = \mu \nabla_u V + \nabla_u W\,,$$
$$\nabla_u(fV) = (uf)V_p + f(p)\nabla_u V$$
for any $u \in T_pM$, $\mu \in \mathbb{R}$, $V, W \in \Gamma(TM)$ and $f \in C(M)$.

(3) *∇ is smooth: If $U, V \in \Gamma(TM)$, then $\nabla_U V \in \Gamma(TM)$, where $(\nabla_U V)_p \doteqdot \nabla_{U_p} V$.*

Given a tangent vector $u \in TM$, the corresponding map $\nabla_u : \Gamma(TM) \to TM$, $V \mapsto \nabla_u V$, is called the **covariant derivative** in the direction u. Given a vector field $V \in \Gamma(TM)$, the map $\nabla V : TM \to TM$, $u \mapsto \nabla_u V$, is a tensor of type $(1, 1)$, called the **covariant differential** of V.

More generally, we can define a connection $\nabla : TM \times \Gamma(E) \to E$ on any vector bundle E over M by replacing $V, W \in \Gamma(TM)$ with $V, W \in \Gamma(E)$ in the above definition (in the third part of the definition, U will still be a vector field).

We first observe that, for $u \in T_pM$, $\nabla_u Y$ is uniquely determined by the restriction of Y to an open neighborhood of p.

Lemma 11.2. *Let M be a manifold with connection ∇ and let X and Y be vector fields such that $X|_U = Y|_U$ for some open set U of M. Then $\nabla_u X = \nabla_u Y$ for every $u \in \pi^{-1}(U)$. In particular, given any open $U \subset M$, $\nabla_u Y$ is well defined for any $Y \in \Gamma(TU)$ and $u \in \pi^{-1}(U)$.*

Proof. To prove the first claim, it suffices to prove that $\nabla_u W = 0$ for any W such that $W|_U \equiv 0$. The claim follows by applying this to the vector field $W \doteqdot X - Y$. To see that $\nabla_u W = 0$, let f be a smooth function satisfying $f(p) = 0$ and $f|_{M\setminus U} \equiv 1$. Then $fW = W$ and hence
$$\nabla_u W = \nabla_u(fW) = (uf)W_p + f(p)\nabla_u W = 0\,.$$

To prove the second claim, we extend Y to a smooth vector field $\overline{Y}$ on M and set $\nabla_u Y \doteqdot \nabla_u \overline{Y}$. This is well defined since, by the previous claim, $\nabla_u \overline{Y}$ is independent of the extension. □

In fact, we can say more: In order to differentiate a vector field in the direction of $\omega'(0)$, we only need to know the values of Y near $\omega(0)$ along ω.

Lemma 11.3. *Let $\omega : I \to M$ be a curve through $\omega(0) = p \in M$ and $X, Y \in \Gamma(TM)$ vector fields which agree on ω; that is, $\omega^* X = \omega^* Y$. Then*

$\nabla_u Y = \nabla_u X$, *where* $u = \omega'(0)$. *In particular,* $\nabla_u Y$ *is well defined for any* $Y \in \Gamma(\omega^* TM)$.

Proof. Let $\phi : U \to V$ be a chart about $p \in M$. Given $u = u^i \partial_i|_p \in T_p M$ and $Y = Y^i \partial_i \in TU$ we compute, using the definition of ∇ and Lemma 11.2,

$$\begin{aligned}\nabla_u Y = u^i \nabla_{\partial_i|_p}(Y^j \partial_j) &= u^i (\partial_i|_p Y^j)\partial_j|_p + u^i Y^j(p) \nabla_{\partial_i|_p}\partial_j \\ &= \Big(uY^k + u^i Y^j(p) {\Gamma_{ij}}^k(p)\Big)\, \partial_k|_p\,,\end{aligned}$$

where we have defined the n^3 functions ${\Gamma_{ij}}^k : U \to \mathbb{R}$ via

$$ {\Gamma_{ij}}^k \partial_k = \nabla_{\partial_i}\partial_j\,. \tag{11.3}$$

Thus,

$$\nabla_u Y = \Big((Y^k \circ \omega)'(0) + u^i Y^j(p) {\Gamma_{ij}}^k(p)\Big)\, \partial_k|_p\,. \tag{11.4}$$

The first claim follows as in the proof of the preceding lemma. To prove the second claim, we extend Y smoothly to a vector field $\overline{Y}$ defined on a neighborhood U of p (by making it constant in $x^1, \dots, x^{n-1}$ with respect to some chart $x : U \to \mathbb{R}^n$ for which $x^n \circ \omega \equiv 0$) and set $\nabla_u Y \doteqdot \nabla_u \overline{Y}$. This is well defined since, by the previous claim, $\nabla_u \overline{Y}$ is independent of the extension. □

The functions ${\Gamma_{ij}}^k$ defined by (11.3) are called the **coefficients** (or **Christoffel symbol**) of the connection ∇.

We will often need to differentiate vector fields along paths or other smooth maps. This is achieved most naturally by introducing the pullback connection.

Definition 11.4. *Let N be a manifold equipped with a connection ∇ on TN and let $F : M \to N$ be a smooth map of a manifold M into N. The* ***pullback connection*** ${}^F\nabla$ *is the unique connection on F^*TNsatisfying*

$$ {}^F\nabla_u F^* V \doteqdot F^* \nabla_{dF(u)} V \tag{11.5}$$

for any $V \in \Gamma(TN)$, where $(F^*V)_p \doteqdot V_{F(p)}$ defines the ***pullback*** *of a vector field $V \in \Gamma(TN)$ and $F^*(F(p), v) = (p, v)$ defines the* ***pullback*** *of a vector $(F(p), v) \in T_{F(p)}N$.*

We need to check that ${}^F\nabla$ is well defined on all sections of F^*TN (i.e. not just the pullback sections F^*V for $V \in \Gamma(TN)$) and uniquely determined by (11.5). To see this, let $\varphi : U \to \mathbb{R}^n$ be a chart for N. Then the vector fields $F^*\partial_i|_p$ form a basis for $(F^*TN)_p$ for all $p \in F^{-1}(U)$ and hence, imposing the Leibniz rule,

$$\begin{aligned}
{}^F\nabla_u V = {}^F\nabla_u(V^j F^*\partial_j) &= (uV^j)F^*\partial_j|_p + V^j(p){}^F\nabla_u F^*\partial_j \\
&= (uV^k)\partial_k|_{F(p)} + V^j(p)\nabla_{dF_p(u)}\partial_j \\
&= (uV^k)\partial_k|_{F(p)} + u^i V^j(p)(dF_p)_i{}^l \nabla_{\partial_l|_{F(p)}}\partial_j \\
&= (uV^k)\partial_k|_{F(p)} + u^i V^j(p)(dF_p)_i{}^l \left(\Gamma_{il}{}^k \circ F\right)(p)\partial_k|_p \\
&= \left(uV^k + u^i V^j(p)\,{}^F\Gamma_{ij}{}^k(p)\right)\partial_k|_p\,,
\end{aligned}$$

where we have defined the coefficients ${}^F\Gamma_{ij}{}^k(p) \doteqdot (dF_p)_i{}^l \left(\Gamma_{il}{}^k \circ F\right)(p)$ and identified the fibres of F^*TN with the corresponding fibres of TN. The claims follow.

We will regularly consider the pullback connection along a curve $\omega : I \to M$. In this case, we will generally write ∇_t for ${}^\omega\nabla_{\partial_t}$, where ∂_t is the canonical tangent vector field to I:

$$\partial_t|_s f = \left.\frac{d}{dt}\right|_{t=0} f(s+t)\,.$$

Given $u \in TM$, the covariant derivative ∇_u extends to an operator $\nabla_u : TM \times \Gamma(E) \to E$ on any homogeneous tensor bundle E by setting $\nabla_u f \doteqdot uf$ for smooth functions $f \in C(M)$ and asserting that ∇_u satisfy the Leibniz rule with respect to tensor products and commute with contractions. That is,

$$\nabla_u(S \otimes T) = \nabla_u S \otimes T + S \otimes \nabla_u T$$

for each $u \in TM$ and homogeneous tensor fields S and T, and

$$\nabla_u \operatorname{tr}(T) = \operatorname{tr}(\nabla_u T)$$

for any homogeneous tensor field T and any trace tr. Indeed, given any covector field $\vartheta \in \Gamma(T^*M)$ and vector fields $U, V \in \Gamma(TM)$, these assertions imply that

$$\begin{aligned}
U\left(\vartheta(V)\right) &= \nabla_U\left(\vartheta(V)\right) \\
&= \nabla_U \operatorname{tr}\left(\vartheta \otimes V\right) \\
&= \operatorname{tr}\left(\nabla_U\vartheta \otimes V + \vartheta \otimes \nabla_U V\right) \\
&= \nabla_U\vartheta(V) + \vartheta(\nabla_U V)
\end{aligned}$$

so that

$$\nabla_U\vartheta(V) = U\left(\vartheta(V)\right) - \vartheta(\nabla_U V)\,,$$

which uniquely determines the tensor $\nabla\vartheta$. Similarly, given any tensor field $T \in \Gamma\left(\bigotimes^k T^*M \otimes \bigotimes^\ell TM\right)$, vector fields $U, U_1, \ldots, U_k \in \Gamma(TM)$ and covector fields $\vartheta^1, \ldots, \vartheta^\ell \in \Gamma(T^*M)$,

$$\begin{aligned} U\left(T(U_1, \ldots, U_k, \vartheta^1, \ldots, \vartheta^\ell)\right) = {} & \nabla_U T(U_1, \ldots, U_k, \vartheta^1, \ldots, \vartheta^\ell) \\ & + T(\nabla_U U_1, \ldots, U_k, \vartheta^1, \ldots, \vartheta^\ell) \\ & + \cdots + T(U_1, \ldots, \nabla_U U_k, \vartheta^1, \ldots, \vartheta^\ell) \\ & + T(U_1, \ldots, U_k, \nabla_U \vartheta^1, \ldots, \vartheta^\ell) \\ & + \cdots + T(U_1, \ldots, U_k, \vartheta^1, \ldots, \nabla_U \vartheta^\ell)\,, \end{aligned} \tag{11.6}$$

which uniquely determines the tensor ∇T.

11.1. Parallel translation. As we alluded to in the introduction to the previous section, a connection provides a notion of *parallelism* for our manifold M, at least along curves.

Definition 11.5. *Let M be a manifold with connection ∇. A vector field X on M is **parallel** if*

$$\nabla X \equiv 0\,.$$

*Given a curve $\omega : I \to M$, X is **parallel along** ω if*

$$\nabla_t X \equiv 0\,.$$

Since ∇_t is a linear operator on $\Gamma(\omega^*TM)$ (the vector fields along ω), the set of parallel vector fields along ω is a vector space over $\mathbb{R}$. Applying results from the theory of linear ordinary differential equations to (11.4) we find, for each $t_0 \in I$ and $v \in T_{\omega(t_0)}M$, a unique parallel vector field V along ω satisfying $V(t_0) = v$.

Proposition 11.6. *Let $\omega : I \to M$ be a piecewise smooth curve. Given $v \in T_{\omega(0)}M$, there exists a unique parallel vector field V along ω.*

Proof. Let $\varphi : U \to \mathbb{R}^n$ be a chart containing $p \doteqdot \omega(0)$. We wish to solve

$$0 = \nabla_t V = \left(\frac{dV^k}{dt} + (\omega')^i V^j {\Gamma_{ij}}^k \circ \omega\right) \partial_k \circ \omega\,,$$

where we use the shorthand

$$\frac{dV^k}{dt} \doteqdot (V^k \circ \omega)' \;\text{ and }\; \omega' \doteqdot (\varphi \circ \omega)'\,.$$

Since the coordinate basis vectors are never zero, this is simply

$$0 = \frac{dV^k}{dt} + (\omega')^i V^j {\Gamma_{ij}}^k \circ \omega$$

for every $k = 1, \dots, n$. This is a linear system of n first order ordinary differential equations for the n functions $V^1, \dots, V^n$ along I, and so (since the coefficient functions $(\omega')^i {\Gamma_{ij}}^k \circ \omega$ are bounded and piecewise continuous) there exist unique solutions with given initial values $V^1(0), \dots, V^n(0)$. We can then extend V to all of ω by solving the corresponding equation in overlapping charts. By uniqueness of solutions, the resulting vector fields must agree on overlaps. $\square$

In particular, the space of parallel vector fields along ω is finite dimensional and has dimension equal to that of M. Thus, we can construct canonical isomorphisms between the tangent spaces to M at different points of ω as follows: Given $t \in I$ and $v \in T_{\omega(0)}M$, let V be the parallel vector field along ω satisfying $V(0) = v$. Then we define the linear isomorphism $\tau_t : T_{\omega(0)}M \to T_{\omega(t)}M$ by

$$\tau_t(v) \doteqdot V(t) \,.$$

We refer to these isomorphisms as **parallel translation along** ω.

As you might expect, the covariant derivative is indeed the differential operator determined by the parallel translation operators.

Theorem 11.7. *Let $\omega : I \to M$ be a smooth curve and V a vector field along ω. Then*

$$\nabla_t V(0) = \frac{d}{dt}\Big|_{t=0} \left(\tau_t^{-1}(V)\right) \,.$$

Proof. Let $E_1(t), \dots, E_n(t)$ be parallel vector fields along ω which are pointwise linearly independent, where $n = \dim M$. Then there exist n functions $V^j : I \to \mathbb{R}, j = 1, \dots, n$, such that

$$V(t) = V^j(t) E_j(t) \,.$$

On the one hand,

$$\nabla_t V = \nabla_t (V^j E_j) = \partial_t V^j E_j \,.$$

On the other hand,

$$\frac{d}{dt}\Big|_{t=0} \left(\tau_t^{-1}(V)\right) = \frac{d}{dt}\Big|_{t=0} \left(V^j(t) E_j\big|_{t=0}\right) = (\partial_t|_{t=0} V^j) E_j\big|_{t=0} \,. \qquad \square$$

The parallel translation operators give a convenient way to identify the tangent spaces to M at different points along a smooth curve. However, it is important to note that the parallel translation operators depend on the choice of curve. In particular, the parallel translation operators cannot be extended to give canonical identifications of all the tangent spaces to each other (indeed, if we could do this, we could construct non-vanishing vector fields, which, as we have mentioned, is topologically impossible in some situations). Also, it does not really make sense to think of a parallel vector field as being "constant", as the following example illustrates.

Example 11.8. *Consider the manifold S^2 equipped with the submanifold connection defined in Exercise 11.1. Then a vector field along a curve in S^2 is parallel if and only if its rate of change (as a vector in $\mathbb{R}^3$) is always normal to the surface of S^2.*

Consider the path γ on S^2 which starts at the north pole, follows a line of longitude to the equator, follows the equator for some distance (say, a quarter of the way around) and then follows another line of longitude back to the north pole. Say,

$$\gamma(t) \doteqdot \begin{cases} (0, \sin t, \cos t) & \text{if} \quad t \in (0, \frac{\pi}{2})\,, \\ (\sin t, -\cos t, 0) & \text{if} \quad t \in (\frac{\pi}{2}, \pi)\,, \\ (-\cos t, 0, -\sin t) & \text{if} \quad t \in (\pi, \frac{3\pi}{2})\,. \end{cases}$$

Note that, on each of the three segments,

$$D_t \gamma' = -\gamma \perp T_\gamma S^2$$

and hence

$$\nabla_t \gamma' = 0\,.$$

We compute the vector field given by parallel translation along γ of a vector which is orthogonal to the initial velocity vector at the north pole.

On the first segment, parallel translation keeps the vector constant as a vector in $\mathbb{R}^3$ (this must be the parallel translation since it remains tangent to S^2, and has zero rate of change, so certainly the tangential component of its rate of change is zero).

On the segment around the equator, we start with a vector tangent to the equator. Parallel translation will give us the tangent vector to the equator of the same length as $V(\frac{\pi}{2})$ as we move around the equator. Indeed, the vector $W(t) \doteqdot |V(\frac{\pi}{2})|\gamma'(t)$ is parallel along γ and satisfies $W(\frac{\pi}{2}) = V(\frac{\pi}{2})$ and hence $V(t) = W(t)$.

On the final segment, the situation is the same as the first segment: We can take V to be constant as a vector in $\mathbb{R}^3$.

We conclude that parallel translation around the entire loop has the effect of rotating the vector through the angle $\frac{\pi}{2}$. Indeed, by choosing a different angle between the two lines of longitude, we can generate arbitrary rotations by parallel translation around the loop.

11.2. Vertical projections on the tangent bundle. There is another way of looking at connections: Let us return to the original problem — to define the directional derivative of a vector field. Recall that a vector field $V \in \Gamma(TM)$ is a smooth map $V : M \to TM$ satisfying $V_p \in T_pM$ for every $p \in M$. Since this is just a smooth map between manifolds, we can differentiate it! Thinking of TM purely as a manifold, this gives the derivative map $(dV)_p : T_pM \to T_{V_p}(TM)$, where $T(TM)$ is the tangent bundle of TM. In other words, we can think of the derivative of a vector field on a manifold M (of dimension n) as a vector tangent to the ($2n$-dimensional) manifold TM. Each fibre T_pM is a submanifold of TM, so a tangent vector to TM at a point $\xi = (p, v) \in TM$ will have some component tangent to the fibre T_pM and some component transverse to it. In local coordinates $x^1, \ldots, x^n, \dot{x}^1, \ldots, \dot{x}^n$ for TM, the tangent vectors $\partial_1, \ldots, \partial_n$ corresponding to the first n-coordinates represent change in position in M, which means that motion in these directions amounts to moving across a family of fibres in TM; the tangent vectors $\dot{\partial}_1, \ldots, \dot{\partial}_n$ corresponding to the remaining n coordinates are tangent to the fibres. Writing the vector field $V \in \Gamma(TM)$ in the coordinates for TM as $V(x^1, \ldots, x^n) = (x^1, \ldots, x^n, V^1, \ldots, V^n)$, we compute

$$(dV)_p u = u^i \partial_i\big|_{V_p} + (uV^i)\dot{\partial}_i\big|_{V_p} .$$

Another way of thinking of a connection is as a projection of this onto the subspace of $T_\xi(TM)$ tangent to the fibre, which we can naturally identify with the fibre itself (the fibre is a vector space, so it can be canonically identified with its tangent space at each point as in 11.1). We will denote by $\mathcal{V}_\xi$ the subspace of $T_\xi(TM)$ spanned by the vectors $\dot{\partial}_1\big|_\xi, \ldots, \dot{\partial}_n\big|_\xi$. Note that this space is independent of the choice of local coordinates:

Proposition 11.9. *$\mathcal{V}_\xi = \ker(d\pi_\xi)$.*

Proof. See exercise 11.4 □

Given $\xi = (p, v) \in TM$, we call $\mathcal{V}_\xi$ the *vertical subspace* of $T_\xi(TM)$. The vertical subspace is naturally identified with T_pM by the map $\iota : \mathcal{V}_\xi \to T_pM$ which sends $v^i\dot{\partial}_i\big|_\xi$ to $v^i\partial_i\big|_p$. Roughly speaking, a connection corresponds to an extension of ι to the whole space $T_\xi(TM)$.

Definition 11.10. *A* vertical projection *on TM is a map $\Pi : \xi \to \Pi_\xi$, which assigns to each $\xi = (p, v) \in TM$ a linear map Π_ξ from $T_\xi(TM)$ to T_pM satisfying*

(1) $\Pi_\xi = \iota$ on $\mathcal{V}_\xi$ and

(2) Π is consistent with the additive structure on TM: If we take ξ_1 and ξ_2 to be paths in TM of the form $\xi_i(t) = (p(t), v_i(t))$, then

$$\Pi_{\xi_1}(\xi_1') + \Pi_{\xi_2}(\xi_2') = \Pi_\xi(\xi') ,$$

where $\xi \doteqdot \xi_1 + \xi_2 \doteqdot (p, v_1 + v_2)$.

Given a vertical projection Π, we can produce a connection ∇ by setting

$$\nabla_v X \doteqdot \Pi_{(p,X_p)}(dX|_p v) .$$

for any $X \in \Gamma(TM)$ and any $v \in T_pM$. Conversely, given a connection ∇, we can produce a vertical projection Π by taking

$$\Pi_\xi \left(u^i \partial_i\big|_\xi + \dot{u}^i \dot{\partial}_i\big|_\xi \right) \doteqdot \nabla_t X \tag{11.7}$$

for any $\xi = (p, v) \in TM$, where the right-hand side is the covariant derivative of the vector field $X(t) \doteqdot \left(v^i + t\dot{u}^i\right) \partial_i$ along the curve $\gamma(t) \doteqdot \varphi^{-1}\left(\varphi(p) + tu^i e_i\right)$.

It is instructive to consider the parallel transport operators in terms of the vertical projections: since Π_ξ maps a $2n$-dimensional vector space to an n-dimensional vector space, and is non-degenerate on the n-dimensional vertical subspace $\mathcal{V}_\xi$ tangent to the fibre T_pM, the kernel of Π_ξ is an n-dimensional subspace of $T_\xi(TM)$ which is complementary to the vertical subspace. We call this the horizontal subspace $\mathcal{H}_\xi$ of $T_\xi(TM)$. A vector field $X \in \Gamma(TM)$ is parallel along a curve γ if and only if $\nabla_t X$ lies in $\mathcal{H}$ at every point. A vertical projection is uniquely determined by the choice of a horizontal subspace at each point (complementary to the vertical subspace and consistent with the linear structure).

11.3. Existence and non-uniqueness of connections. We will show that every smooth manifold can be equipped with a connection. In fact, there are many connections on any manifold, and no preferred or canonical one (later, when we introduce Riemannian metrics, we will have a way of producing a canonical connection).

Proposition 11.11. *Every smooth manifold admits a connection (in fact, many).*

Proof. Choose a locally finite cover of M by coordinate charts $\{\varphi_\alpha : U_\alpha \to V_\alpha\}$ and a subordinate partition of unity $\{\rho_\alpha : M \to \mathbb{R}\}$. We will use this partition of unity to patch together connections on each chart defined by coordinate differentiation: On each coordinate patch U_α we define, for each $u \in TU_\alpha$, an operator $\nabla^{(\alpha)}_u$ on vector fields with support contained inside U_α via

$$\nabla^{(\alpha)}_u X \doteqdot (uX^i)\partial^{(\alpha)}_i .$$

We then set, for any $X \in \Gamma(TM)$ and $u \in TM$,

$$\nabla_u X \doteqdot \sum_\alpha \nabla^{(\alpha)}_u (\rho_\alpha X) .$$

This makes sense because the sum is actually finite at each point of M. The result is clearly $\mathbb{R}$-linear in both u and X. The Leibniz rule and smoothness are also easily checked. $\square$

Note that, although connections are $\mathbb{R}$-linear in both variables, they (because of the Leibniz rule) are not "tensors": Given $u \in T_pM$, $\nabla_u X$ is not $\mathbb{R}$-linear in X and, indeed, depends on the values of X in a neighborhood of p, not just at p. However, given two connections, their *difference* is a tensor:

Proposition 11.12. *Let D and ∇ be two connections on a manifold M. For any $p \in M$ we define a map $A_p : T_pM \times \Gamma(TM) \to T_pM$ by $A_p(u, \cdot) \doteqdot D_u - \nabla_u$. Given $u \in T_pM$,*

$$A(u, X) = A(u, Y)$$

for any two vector fields X and Y satisfying $X_p = Y_p$.

Thus, given $p \in M$, the map $A_p : T_pM \times T_pM \to T_pM$ defined by $A_p(u, v) \doteqdot A(u, V)$ for any $V \in \Gamma(TM)$ such that $V_p = v$ is well defined.

Moreover, A depends smoothly on p: If X and Y are smooth vector fields then $p \mapsto A_p(X_p, Y_p)$ is a smooth function.

Proof. See Exercise 11.2. □

Exercises.

Exercise 11.1. *Let M be a submanifold of $\mathbb{R}^N$. Define $\nabla : TM \times \Gamma(TM) \to TM$ by*

$$\nabla_u V \doteqdot (D_u V)^\top ,$$

where D is the directional derivative on $T\mathbb{R}^N$ and, for any $p \in M$ and $u \in T_p\mathbb{R}^N$, $(u)^\top$ denotes the orthogonal projection (with respect to the inner product on $\mathbb{R}^N$) of u onto T_pM. Show that ∇ is a connection.

Exercise 11.2. *Let D and ∇ be two connections on a manifold M. For any $p \in M$ we define a map $A_p : T_pM \times \Gamma(TM) \to T_pM$ by $A_p(u, \cdot) \doteqdot D_u - \nabla_u$. Given $u \in T_pM$, show that*

$$A(u, X) = A(u, Y)$$

for any two vector fields X and Y satisfying $X_p = Y_p$.

Thus, given $p \in M$, the map $A_p : T_pM \times T_pM \to T_pM$ defined by $A_p(u, v) \doteqdot A(u, V)$ for any $V \in \Gamma(TM)$ such that $V_p = v$ is well defined.

Show that A depends smoothly on p: If X and Y are smooth vector fields then $p \mapsto A_p(X_p, Y_p)$ is a smooth function.

Exercise 11.3. *Let M be a manifold with connection ∇. Suppose that $A_p : T_pM \times T_pM \to T_pM$ is bilinear for each $p \in M$ and varies smoothly over M (in the sense that A applied to two smooth vector fields gives a smooth vector field). Show that*

$$D_u V \doteqdot \nabla_u V + A(u, V)$$

defines another connection on M.

Exercise 11.4. *Show that* $\mathcal{V}_\xi = \ker(D_\xi \pi)$, *where* $\pi : TM \to M$ *is the bundle projection.*

Exercise 11.5. *Check that the vertical projection* Π_ξ *defined by* (11.7) *is independent of the chosen coordinates.*

12. Geodesics and the exponential map

Parametrized straight lines $\gamma : t \mapsto p + tv$ in Euclidean space are characterized dynamically by the fact that they have no acceleration. That is, $D_t\gamma' = 0$. In Newtonian physics, particles accelerate proportionally to force exerted upon them and hence straight lines are the paths taken by (non-stationary) particles undergoing no forces. More generally, a particle moving subject to a constraint that it lies in a submanifold M (with no other external forces) moves as it would in free space, except that any component of its acceleration in directions normal to the surface are automatically cancelled out by the constraint forces. In other words, the motion of the particle is determined by the equation

$$\nabla_t\gamma' = \big(D_t\gamma'\big)^\top = 0\,,$$

where $\top$ is the projection onto the tangent bundle of M.

Definition 12.1. *Let M be a manifold with connection ∇. A* ***geodesic*** *is a path $\gamma : I \to M$ having no covariant acceleration; i.e.*

$$\nabla_t\gamma' \equiv 0\,.$$

We can obtain existence and uniqueness of geodesics with given initial data $(\gamma(0), \gamma'(0)) = (p, v) \in TM$ by writing the geodesic equation locally as a second-order ODE in coordinates $\varphi : U \to \mathbb{R}^n$ containing p. Indeed,

$$\nabla_t\gamma' = \Big((\gamma'')^k + (\gamma')^i(\gamma')^j{\Gamma_{ij}}^k \circ \gamma\Big)\,\partial_k \circ \gamma\,.$$

This vector vanishes only if

$$(\gamma'')^k + (\gamma')^i(\gamma')^j{\Gamma_{ij}}^k \circ \gamma = 0 \tag{12.1}$$

for every $k = 1, \ldots, n$, where $(\gamma'')^k \doteqdot \partial_t(\gamma')^k$. Existence and uniqueness of solutions with given initial data then follow from the theory of second-order ODE. By uniqueness, the solutions can be extended across coordinate transitions. We *could* do it this way, but it will also be useful to exhibit the second order geodesic equation as a first order system on the manifold TM. First observe that, introducing a new variable $v^i \doteqdot (\gamma')^i$, the second order system (12.1) becomes the first order system

$$\begin{aligned} \frac{d}{dt}\gamma^i &= v^i\,; \\ \frac{d}{dt}v^k &= -\,v^iv^j{\Gamma_{ij}}^k \circ \gamma\,. \end{aligned} \tag{12.2}$$

We want to identify the solutions of this system as integral curves of a vector field on TM. Extending the coordinate chart $\varphi : U \to \mathbb{R}^n$ to a chart $\Phi : \pi^{-1}(U) \to \mathbb{R}^{2n}$ containing $(p, v) \in TM$, consider the (local) vector field

$$\mathcal{G} \doteqdot \dot{x}^i\partial_i - \dot{x}^i\dot{x}^j{\Gamma_{ij}}^k\dot{\partial}_k\,. \tag{12.3}$$

Observe that $\mathcal{G}$ does not depend on the choice of coordinates and hence extends uniquely to a vector field on TM.

Let $\Omega : I \to TM$ be an integral curve of $\mathcal{G}$. Set $\omega \doteqdot \pi \circ \Omega : I \to M$. Observe that $d\pi_{(p,v)}(\partial_i) = \partial_i|_p$ and $d\pi_{(p,v)}(\dot{\partial}_i) = 0$. Then, abusing notation,

$$\omega' = d\pi_\Omega(\mathcal{G}) = \Omega\,.$$

It follows from (12.2) that ω is geodesic. Conversely, suppose that $\omega : I \to M$ is geodesic and define $\Omega : I \to TM$ by

$$\omega' = \Omega\,.$$

Then, by (12.2), Ω is an integral curve of $\mathcal{G}$.

The vector field $\mathcal{G}$ is called the **geodesic flow**.

Theorem 12.2. *Given $\xi = (p, v) \in TM$ there exists a unique maximal geodesic $\gamma_\xi : I \to M$ satisfying $(\gamma_\xi(0), \gamma_\xi'(0)) = (p, v)$ and γ_ξ depends smoothly on ξ and t. Indeed, γ_ξ is given by*

$$\gamma_\xi(t) = (\pi \circ \Phi)(\xi, t)$$

and γ_ξ' is given by

$$\gamma_\xi'(t) = \Phi(\xi, t)\,,$$

where $\Phi : TM \times I \to TM$ is the maximal flow of the geodesic flow $\mathcal{G}$.

Proof. Given the preceding discussion, this follows from the existence and uniqueness of the maximal flow of $\mathcal{G}$, which follows from Theorem 7.1. □

Proposition 12.3. *Given $\xi \in TM$, let $I_\xi \subset \mathbb{R}$ be the domain of γ_ξ, the maximal geodesic satisfying $(\gamma_\xi(0), \gamma_\xi'(0)) = \xi$. Show that*

$$I_{-\xi} = -I_\xi \quad \textit{and} \quad I_{r\xi} = \tfrac{1}{r} I_\xi \quad \textit{for any} \quad r > 0\,.$$

Proof. See Exercise 12.2 □

Definition 12.4. *Let M be a manifold with connection ∇. The **domain of the exponential map** $\mathcal{T}M$ is the subset of TM given by*

$$\mathcal{T}M \doteqdot \{\xi \in TM : 1 \in I_\xi\}\,,$$

where I_ξ is the domain of the maximal geodesic $\gamma_\xi : I_\xi \to M$ satisfying $(\gamma_\xi(0), \gamma_\xi'(0)) = \xi$.

*The **exponential map** is the map $\exp : \mathcal{T}M \to M$ defined by*

$$\exp \xi = \gamma_\xi(1)\,.$$

*Given any $p \in M$, the **exponential map at** p is the map $\exp_p : \mathcal{T}M \cap T_pM \to M$ given by*

$$\exp_p v \doteqdot \gamma_{(p,v)}(1)\,.$$

That is, $\exp_p \doteqdot \exp|_{\mathcal{T}M \cap T_pM}$.

*The **zero section** of TM is the vector field $0 \in \Gamma(TM)$ defined by $0_p = 0$ (the zero vector in T_pM).*

Theorem 12.5. *The domain of the exponential map is open and **star-shaped with respect to the zero section**; that is, for any $\xi \in \mathcal{T}M$ and any $r \in [0,1]$, $r\xi \in \mathcal{T}M$.*

For every $p \in M$ the exponential map at p is differentiable and $(d\exp_p)_{0_p}$ has maximal rank (equal to the dimension of M).

The map $\pi \times \exp : \mathcal{T}M \to M \times M$ defined by

$$(\pi \times \exp)(\xi) \doteqdot (\pi(\xi), \exp \xi)$$

is differentiable on the zero section $0(M) \subset \mathcal{T}M$ and $d(\pi \times \exp)$ has maximal rank (equal to twice the dimension of M) on $0(M)$. In particular, $\pi \times \exp$ is a local diffeomorphism near $(p, 0) \in TM$ for any $p \in M$.

Proof. That $\mathcal{T}M$ is open is immediate since the domain of the maximal flow of $\mathcal{G}$ is open. That $\mathcal{T}M$ is starshaped follows from Exercise 12.2. To prove the remaining claims, we need to calculate the derivative of $\pi \times \exp$.

Fix $p \in M$ and choose a coordinate chart $\varphi : U \to \mathbb{R}^n$ around p and extend it to a corresponding coordinate chart $\Phi : \pi^{-1}(U) \to \mathbb{R}^{2n}$ around T_pM. Let $\eta : U \times U \to \mathbb{R}^{2n}$ be the chart for $M \times M$ near (p,p) given by $\varphi \times \varphi$. Since $\pi \times \exp$ is continuous, the set $(\pi \times \exp)^{-1}(U \times U)$ is open. We need to calculate the derivative of $\eta \circ (\pi \times \exp) \circ \Phi^{-1}$ on $\Phi(0(U))$. Observe that

$$\begin{aligned}\eta \circ (\pi \times \exp) \circ \Phi^{-1}(x, \dot{x}) &= \eta\left((\pi \times \exp)(\varphi^{-1}(x), \dot{x}^i \partial_i|_{\varphi^{-1}(x)})\right)\\ &= (x, \varphi(\exp_{\varphi^{-1}(x)}(\dot{x}^i \partial_i|_{\varphi^{-1}(x)}))\,.\end{aligned}$$

The derivative of the first factor with respect to x is the identity matrix and is zero with respect to $\dot{x}$. So it suffices to compute the derivative of the map

$$(x, \dot{x}) \mapsto \varphi \circ \exp_{\varphi^{-1}(x)}(\dot{x}^i \partial_i)$$

with respect to $\dot{x}$ at zero. But this is also of maximal rank: Identifying the tangent space of T_pM at the zero vector with T_pM in the usual way, we compute

$$(d\exp_p)_0(v) = \frac{d}{dt}\Big|_{t=0} \exp_p(tv) = \gamma'_{(p,v)}(0) = v\,.$$

We have proved that both $\exp_p$ and $\pi \times \exp$ are of maximal rank. The final claim follows from the inverse function theorem. ☐

Corollary 12.6. *For any $p \in M$ there exists a neighborhood $U \subset M$ of p and a neighborhood $\mathcal{O} \subset \mathcal{T}M$ of the zero section $0|_U$ such that, given any pair of points q and r in U, there exists a unique geodesic $\gamma_{qr} : [0,1] \to M$ with $\gamma_{qr}(0) = q$, $\gamma_{qr}(1) = r$ and $\gamma'_{qr}(0) \in \mathcal{O}$. Moreover, $\mapsto \gamma_{qr}$ depends smoothly on q and r.*

Proof. Choose a neighborhood $\mathcal{O}$ of $(p, 0)$ in $\mathcal{T}M$ on which $\pi \times \exp$ is a diffeomorphism, and a neighborhood U of p in M sufficiently small that $U \times U \subset (\pi \times \exp)(\mathcal{O})$. These neighborhoods satisfy our requirements. $\square$

Exercises.

Exercise 12.1. *Show that the vector field $\mathcal{G}$, defined in local coordinates by* (12.3)*, does not depend on the choice of coordinates and hence extends uniquely to a vector field on TM.*

Exercise 12.2. *Given $\xi \in TM$, let $I_\xi \subset \mathbb{R}$ be the domain of γ_ξ, the maximal geodesic satisfying $(\gamma_\xi(0), \gamma'_\xi(0)) = \xi$. Show that*

$$I_{-\xi} = -I_\xi \quad \textit{and} \quad I_{r\xi} = \tfrac{1}{r} I_\xi \quad \textit{for any} \quad r > 0\,.$$

13. Torsion and curvature

Given two vector fields U and V, we can generate a third vector field $\mathrm{Tor}(U,V)$, called the **torsion** of U and V, via antisymmetrization:

$$\mathrm{Tor}(U,V) \doteqdot \nabla_U V - \nabla_V U - [U,V]\,.$$

Observe that $\mathrm{Tor}(U,V)_p$ depends only on the values of U and V at p.

Lemma 13.1. *Let M be a manifold with connection ∇. The torsion operator* Tor *of ∇ is bilinear over the ring $C(M)$ of smooth functions. Equivalently, there is a unique tensor* $\mathrm{Tor} \in \Gamma(T^*M \otimes T^*M \otimes TM)$ *such that*

$$\mathrm{Tor}(U,V)|_p = \mathrm{Tor}_p(U_p, V_p) \ \ \textit{for all}\ \ U,V \in \Gamma(TM) \ \ \textit{and}\ \ p \in M\,.$$

Proof. Since Tor is antisymmetric and $\mathbb{R}$-bilinear, to prove bilinearity over $C(M)$ it suffices to show that $\mathrm{Tor}(fU,V) = f\,\mathrm{Tor}(U,V)$ for all $f \in C(M)$. Indeed, for any smooth function f on M,

$$\begin{aligned}
\mathrm{Tor}(U,fV) &= \nabla_U(fV) - \nabla_{fV}U - [U,fV] \\
&= (Uf)V + f\nabla_U V - f\nabla_V U - ((Uf)V + fUV - fVU) \\
&= f\nabla_U V - f\nabla_V U - f[U,V] \\
&= f\,\mathrm{Tor}(U,V)\,.
\end{aligned}$$

The second claim is a consequence of the identification of $\Gamma(T^*M \otimes T^*M \otimes TM)$ with $\Gamma(T^*M) \otimes \Gamma(T^*M) \otimes \Gamma(TM)$ described in §6.2. $\square$

Definition 13.2. *A connection is* ***symmetric*** *(a.k.a.,* ***torsion-free****) if its torsion tensor is identically zero.*

Observe that, for any symmetric connection ∇, covariant differentiation of coordinate vector fields commutes:

$$\nabla_i \partial_j = \nabla_j \partial_i\,.$$

Given two vector fields U, $V \in \Gamma(TM)$, the **curvature operator** $\mathrm{Rm}(U,V) : \Gamma(TM) \to \Gamma(TM)$ is the second order covariant differential operator defined by

$$\mathrm{Rm}(U,V)W \doteqdot \nabla_V(\nabla_U W) - \nabla_U(\nabla_V W) - \nabla_{[V,U]}W\,.$$

Lemma 13.3. *Let M be a manifold with connection ∇. The map*

$$(U,V,W) \mapsto \mathrm{Rm}(U,V)W$$

is trilinear over the ring $C(M)$ of smooth functions. Equivalently, there is a unique tensor $\mathrm{Rm} \in \Gamma(T^*M \otimes T^*M \otimes T^*M \otimes TM)$ *such that*

$$(\mathrm{Rm}(U,V)W)|_p = \mathrm{Rm}_p(U_p,V_p)W_p \ \ \textit{for all}\ \ U,V,W \in \Gamma(TM) \ \ \textit{and}\ \ p \in M\,.$$

Proof. It suffices to prove that $\mathrm{Rm}(U,V)W$ is $C(M)$-linear in each vector field. Since it is anti-symmetric in U and V, we only have to check this for U and W. To see this for U, observe that

$$\begin{aligned}\nabla_V(\nabla_{fU}W) - \nabla_{[V,fU]}W &= \nabla_V(f\nabla_U W) - \nabla_{(Vf)U+f[V,U]}W \\ &= (Vf)\nabla_U W + f\nabla_V(\nabla_U W) - Vf\nabla_U W - f\nabla_{[V,U]}W \\ &= f\left(\nabla_V(\nabla_U W) - \nabla_{[V,U]}W\right) .\end{aligned}$$

The claim follows. To prove $C(M)$-linearity in W, consider

$$\begin{aligned}\mathrm{Rm}(U,V)(fW) &= \nabla_V(\nabla_U(fW)) - \nabla_U(\nabla_V(fW)) - \nabla_{[V,U]}(fW) \\ &= \nabla_V\left((Uf)W + f\nabla_U W\right) - \nabla_U\left((Vf)W + f\nabla_V W\right) \\ &\quad - ([V,U]f)W - f\nabla_{[V,U]}W \\ &= (VUf - UVf)W + (Uf)\nabla_V W - (Vf)\nabla_U W \\ &\quad + (Vf)\nabla_U W - (Uf)\nabla_V W + f\left(\nabla_V(\nabla_U W) - \nabla_U(\nabla_V W)\right) \\ &\quad - ([V,U]f)W - f\nabla_{[V,U]}W \\ &= f\,\mathrm{Rm}(U,V)W .\end{aligned}$$

The second claim is a consequence of the identification of $\Gamma(T^*M \otimes T^*M \otimes T^*M \otimes TM)$ with $\Gamma(T^*M) \otimes \Gamma(T^*M) \otimes \Gamma(T^*M) \otimes \Gamma(TM)$ described in §6.2. □

Lemma 13.4 (The first Bianchi identity)**.** *For any $u, v, w \in T_pM$,*

$$\begin{aligned}\mathrm{Rm}(u,v)w + {}& \mathrm{Rm}(v,w)u + \mathrm{Rm}(w,u)v \\ &= \mathrm{Tor}(u,\mathrm{Tor}(v,w)) + \mathrm{Tor}(v,\mathrm{Tor}(w,u)) + \mathrm{Tor}(w,\mathrm{Tor}(u,v)) \\ &\quad - \nabla_u \mathrm{Tor}(v,w) - \nabla_v \mathrm{Tor}(w,u) - \nabla_w \mathrm{Tor}(u,v) .\end{aligned}$$

In particular, if ∇ is symmetric then

$$\mathrm{Rm}(u,v)w + \mathrm{Rm}(v,w)u + \mathrm{Rm}(w,u)v = 0 .$$

Proof. By the trilinearity of Rm, it suffices to prove the claim for any coordinate vectors ∂_i, ∂_j and ∂_k. This simplifies the calculation since the coordinate vector fields commute. Indeeed,

$$\begin{aligned}&\mathrm{Rm}(\partial_i,\partial_j)\partial_k + \mathrm{Rm}(\partial_j,\partial_k)\partial_i + \mathrm{Rm}(\partial_k,\partial_i)\partial_j \\ &\quad = \nabla_j(\nabla_i\partial_k) - \nabla_i(\nabla_j\partial_k) + \nabla_k(\nabla_j\partial_i) - \nabla_j(\nabla_k\partial_i) + \nabla_i(\nabla_k\partial_j) - \nabla_k(\nabla_i\partial_j) \\ &\quad = \nabla_j\left(\nabla_i\partial_k - \nabla_k\partial_i\right) + \nabla_k\left(\nabla_j\partial_i - \nabla_i\partial_j\right) + \nabla_i\left(\nabla_k\partial_j - \nabla_j\partial_k\right) \\ &\quad = -\nabla_i(\mathrm{Tor}(\partial_j,\partial_k)) - \nabla_j(\mathrm{Tor}(\partial_k,\partial_i)) - \nabla_k(\mathrm{Tor}(\partial_i,\partial_j)) \\ &\quad = -\nabla_i \mathrm{Tor}(\partial_j,\partial_k) - \nabla_j \mathrm{Tor}(\partial_k,\partial_i) - \nabla_k \mathrm{Tor}(\partial_i,\partial_j) \\ &\qquad + \mathrm{Tor}(\partial_i,\mathrm{Tor}(\partial_j,\partial_k)) + \mathrm{Tor}(\partial_j,\mathrm{Tor}(\partial_k,\partial_i)) + \mathrm{Tor}(\partial_k,\mathrm{Tor}(\partial_i,\partial_j)) .\end{aligned}$$

Alternatively, one could compute with arbitrariry vector fields and apply the Jacobi identity. □

Given vector fields U and V, we can extend the operator $\mathrm{Rm}(U,V)$ to a homogeneous operator on the entire tensor algebra: Given a tensor field S, we simply set

$$\mathrm{Rm}(U,V)S \doteqdot \nabla_V(\nabla_U S) - \nabla_U(\nabla_V S) - \nabla_{[V,U]}S\,.$$

By exactly the same computations as in Lemma 13.3, the map $U, V, S \mapsto \mathrm{Rm}(U,V)S$ is $C(M)$-trilinear and hence defines a tensor.

Lemma 13.5. *Let M be a manifold with connection ∇. Given vector fields U and V, the curvature operator $\mathrm{Rm}(U,V)$ satisfies the Leibniz rule with respect to the tensor product and commutes with contractions. Thus, the tensor field $\mathrm{Rm}(U,V)S$ is determined, for any homogeneous tensor field $S \in \Gamma\left(\bigotimes^k T^*M \otimes \bigotimes^\ell TM\right)$, by the formula*

$$\begin{aligned}(\mathrm{Rm}(U,V)S)&(U_1,\dots,U_k,\alpha^1,\dots,\alpha^\ell)\\ &= -\sum_{i=1}^{k} S(U_1,\dots,\mathrm{Rm}(U,V)U_i,\dots,U_k,\alpha^1,\dots,\alpha^\ell)\\ &\quad - \sum_{j=1}^{\ell} S(U_1,\dots,U_k,\alpha^1,\dots,\mathrm{Rm}(U,V)\alpha^j,\dots,\alpha^\ell)\,. \end{aligned} \tag{13.1}$$

Proof. We first show that, given $U, V \in \Gamma(TM)$, $\mathrm{Rm}(U,V)$ distributes over the tensor product. Using the fact that ∇ distributes over $\otimes$ we compute locally in a coordinate basis $\{\partial_i\}$ for any homogeneous tensor fields S and T,

$$\begin{aligned}\nabla_i(\nabla_j(S\otimes T)) &= \nabla_i\left(\nabla_j S\otimes T + S\otimes\nabla_j T\right)\\ &= \nabla_i(\nabla_j S)\otimes T + \nabla_j S\otimes\nabla_i T\\ &\quad + \nabla_i S\otimes\nabla_j T + S\otimes\nabla_i(\nabla_j T)\,.\end{aligned}$$

It follows that

$$\mathrm{Rm}(\partial_i,\partial_j)(S\otimes T) = \mathrm{Rm}(\partial_i,\partial_j)S\otimes T + S\otimes\mathrm{Rm}(\partial_i,\partial_j)T\,.$$

The claim follows since both sides are tensorial.

That $\mathrm{Rm}(U,V)$ commutes with contractions is a simple consequence of the fact that ∇ commutes with contractions: Let U and V be vector fields, S a homogeneous tensor field and tr any contraction defined on $\mathrm{Rm}(U,V)S$.

Then

$$\begin{aligned}
\operatorname{tr}(\operatorname{Rm}(U,V)S) &\doteqdot \operatorname{tr}(\nabla_V(\nabla_U S) - \nabla_U(\nabla_V S) - \nabla_{[V,U]}S) \\
&= \nabla_V \operatorname{tr}(\nabla_U S) - \nabla_U \operatorname{tr}(\nabla_V S) - \nabla_{[V,U]} \operatorname{tr} S \\
&= \nabla_V(\nabla_U \operatorname{tr} S) - \nabla_U(\nabla_V \operatorname{tr} S) - \nabla_{[V,U]} \operatorname{tr} S \\
&= \operatorname{Rm}(U,V) \operatorname{tr} S\,.
\end{aligned}$$

The remaining claim now follows as in the formula (11.6) since the curvature operator vanishes on the smooth functions. □

Example 13.6. *A useful special case of the formula* (13.1) *is the curvature formula*

$$(\operatorname{Rm}(u,v)\alpha)(w) = -\alpha(\operatorname{Rm}(u,v)w)$$

for covectors $\alpha \in T^*M$.

13.1. The linearized geodesic equation. A fundamental tool for the study of geodesics is the **linearized geodesic equation**, also known as **Jacobi's equation of geodesic deviation**.

Let $\omega : (-t_0,t_0)\times(-\varepsilon_0,\varepsilon_0) \to M$ be a smooth one-parameter family of geodesics. That is, for each $\varepsilon \in (-\varepsilon_0,\varepsilon_0)$, the curve $t \mapsto \omega(t,\varepsilon)$ is a geodesic. Since the canonical coordinate vectors ∂_t and ∂_ε for $(-t_0,t_0)\times(-\varepsilon_0,\varepsilon_0)$ commute, we have by the definition of the pullback connection (abusing notation slightly)

$$\nabla_\varepsilon \partial_t \omega - \nabla_t \partial_\varepsilon \omega = \operatorname{Tor}(\partial_\varepsilon \omega, \partial_t \omega)$$

and

$$[\nabla_\varepsilon, \nabla_t] = \operatorname{Rm}(\partial_t \omega, \partial_\varepsilon \omega)\,.$$

Theorem 13.7 (Jacobi (1836))**.** *Let $\omega : (-t_0,t_0)\times(-\varepsilon_0,\varepsilon_0) \to M$ be a smooth one-parameter family of geodesics $\omega(\cdot,\varepsilon)$. Then*

$$0 = \nabla_t(\nabla_t \partial_\varepsilon \omega) + \nabla_t(\operatorname{Tor}(\partial_t\omega, \partial_\varepsilon\omega)) + \operatorname{Rm}(\partial_t\omega,\partial_\varepsilon\omega)\partial_t\omega\,.$$

Proof. By the geodesic property of the family,

$$\begin{aligned}
0 = \nabla_\varepsilon(\nabla_t \partial_t \omega) &= \nabla_t(\nabla_\varepsilon \partial_t \omega) + \operatorname{Rm}(\partial_t\omega,\partial_\varepsilon\omega)\partial_t\omega \\
&= \nabla_t\left(\nabla_t \partial_\varepsilon\omega + \operatorname{Tor}(\partial_t\omega,\partial_\varepsilon\omega)\right) + \operatorname{Rm}(\partial_t\omega,\partial_\varepsilon\omega)\partial_t\omega \\
&= \nabla_t(\nabla_t \partial_\varepsilon\omega) + \nabla_t\left(\operatorname{Tor}(\partial_t\omega,\partial_\varepsilon\omega)\right) + \operatorname{Rm}(\partial_t\omega,\partial_\varepsilon\omega)\partial_t\omega\,. \square
\end{aligned}$$

In particular, if $J \doteqdot \partial_\varepsilon|_{\varepsilon=0}\omega$ and γ is the geodesic $t \mapsto \omega(t,0)$, then J solves

$$0 = \nabla_t\nabla_t J + \operatorname{Tor}(\gamma', \nabla_t J) + \nabla_t \operatorname{Tor}(\gamma', J) + \operatorname{Rm}(\gamma', J)\gamma'\,. \tag{13.2}$$

Solutions to (13.2) are called **Jacobi fields** (along γ).

Exercises.

Exercise 13.1. *The* ***covariant differential*** *of a homogeneous tensor field $S \in \mathscr{T}^{(k,l)}(TM)$ is the tensor field $\nabla S \in \mathscr{T}^{(k+1,l)}(TM)$ defined by*

$$(\nabla S)(X_0, X_1, \dots, X_k, \alpha^1, \dots, \alpha^l) \doteqdot (\nabla_{X_0} S)(X_1, \dots, X_k, \alpha^1, \dots, \alpha^l)\,.$$

The covariant differential of a general tensor field is defined as the sum of the covariant differentials of its homogeneous parts. Covariant differentials can be iterated: the ***second covariant differential*** (*or* ***covariant Hessian***) *of S is the tensor field $\nabla^2 S \doteqdot \nabla(\nabla S)$, the* ***third covariant differential*** *of S is the tensor field $\nabla^3 S \doteqdot \nabla(\nabla(\nabla S))$, etc. Suppose that ∇ is torsion-free and let X, Y and Z be vector fields and S a tensor field. Prove the following three* ***Ricci identities***:

$$\text{(R1)} \qquad \nabla^2 S(Y, X) - \nabla^2 S(X, Y) = \mathrm{Rm}(X, Y)S\,,$$

$$\text{(R2)} \qquad \nabla^3 S(Y, X, Z) - \nabla^3 S(X, Y, Z) = \mathrm{Rm}(X, Y)(\nabla_Z S) - \nabla_{\mathrm{Rm}(X,Y)Z} S$$

and

$$\text{(R3)} \quad \nabla^3 S(X, Z, Y) - \nabla^3 S(X, Y, Z) = (\nabla_X \mathrm{Rm})(Y, Z)S + \mathrm{Rm}(Y, Z)(\nabla_X S)\,,$$

where the tensor field $\nabla \mathrm{Rm}$ is defined (*in the usual way*) *by commuting with contractions:*

$$\begin{aligned}\nabla_X(\mathrm{Rm}(Y, Z)S) = (\nabla_X \mathrm{Rm})(Y, Z)S &+ \mathrm{Rm}(\nabla_X Y, Z)S \\ &+ \mathrm{Rm}(Y, \nabla_X Z)S + \mathrm{Rm}(Y, Z)(\nabla_X S)\,.\end{aligned}$$

Combine (R2) *and* (R3) *to obtain the* ***second Bianchi identity***: *Suppose that ∇ is torsion-free. If X, Y and Z are vector fields and S a tensor field, then*

$$\nabla_X \mathrm{Rm}(Y, Z)S + \nabla_Y \mathrm{Rm}(Z, X)S + \nabla_Z \mathrm{Rm}(X, Y)S = 0\,.$$

Exercise 13.2. *Let $\{e_i\}_{i=1}^n$ be a local field of bases for TM and let $\{\vartheta^i\}_{i=1}^n$ be the field of dual bases. Denote by ${\Gamma_{ij}}^k \doteqdot \vartheta^k(\nabla_i e_j)$ and ${C_{ij}}^k \doteqdot \vartheta^k([e_i, e_j])$ the corresponding connection coefficients and commutator relations, respectively. Show that*

$${\mathrm{Tor}_{ij}}^k = {\Gamma_{ij}}^k - {\Gamma_{ji}}^k - {C_{ij}}^k$$

and

$${\mathrm{Rm}_{ijk}}^\ell = e_j {\Gamma_{ik}}^\ell - e_i {\Gamma_{jk}}^\ell + {\Gamma_{ik}}^p {\Gamma_{jp}}^\ell - {\Gamma_{jk}}^p {\Gamma_{ip}}^\ell - {C_{ji}}^p {\Gamma_{pk}}^\ell\,.$$

Exercise 13.3. *Let M be a manifold equipped with a connection ∇. Given a local field of bases $\{e_i\}_{i=1}^n$ for TM, define the (local)* ***connection 1-forms*** *$\{{\omega_i}^k\}_{i,k=1}^n$ by*

$${\omega_j}^k \doteqdot -{\Gamma_{ij}}^k \vartheta^i\,,$$

where $\{\Gamma_{ij}{}^k\}_{i,j,k=1}^n$ *are the connection coefficients corresponding to* $\{e_i\}_{i=1}^n$, *i.e.,*

$$\nabla_i e_j = \Gamma_{ij}{}^k e_k \,,$$

and $\{\vartheta^i\}_{i=1}^n$ *is the local field of bases dual to* $\{e_i\}_{i=1}^n$.

(a) Show that

$$\omega_j{}^k(e_i) = (\nabla_i \vartheta^k)(e_j) \,.$$

Define the ***torsion 2-forms*** $\{\mathrm{Tor}^k\}_{k=1}^n$ *by*

$$\mathrm{Tor}^k \doteqdot d\vartheta^k + \vartheta^j \wedge \omega_j{}^k \,.$$

(b) Show that

$$\iota_U \iota_V \mathrm{Tor}^k = \vartheta^k(\mathrm{Tor}(U,V)) \,,$$

where the tensor on the right is the torsion tensor of ∇.

Define the ***curvature 2-forms*** $\{\mathrm{Rm}_k{}^\ell\}_{k,\ell=1}^n$ *by*

$$\mathrm{Rm}_k{}^\ell \doteqdot d\omega_k{}^\ell + \omega_k{}^p \wedge \omega_p{}^\ell \,.$$

(c) Show that

$$\iota_U \iota_V \mathrm{Rm}_k{}^\ell = \mathrm{Rm}(U, V, e_k, \vartheta^\ell) \,.$$

(d) Prove the "first Bianchi identity"

$$d\mathrm{Tor}^\ell = \omega_k{}^\ell \wedge \mathrm{Tor}^k - \mathrm{Rm}_k{}^\ell \wedge \vartheta^k \,.$$

(e) Assuming ∇ *is torsion-free, prove the "second Bianchi identity"*

$$d\,\mathrm{Rm}_k{}^\ell = \mathrm{Rm}_k{}^p \wedge \omega_p{}^\ell - \omega_k{}^p \wedge \mathrm{Rm}_p{}^\ell \,.$$

14. Riemannian metrics

Next, we want to introduce a way of defining geometric notions such as lengths of curves, angles between curves and areas of subsets of our manifold. This is achieved through the introduction of a *Riemannian metric* — a family of inner products on the tangent spaces.

Definition 14.1. *A* ***Riemannian metric*** *or* ***metric tensor*** *on a differentiable manifold M is a smooth section $g \in \Gamma(T^*M \otimes T^*M)$ such that, for each $p \in M$, the bilinear form $g_p : T_pM \times T_pM \to \mathbb{R}$ is symmetric,*

$$g_p(u,v) = g_p(v,u)\,,$$

and positive definite,

$$u \neq 0 \quad \Longrightarrow \quad g_p(u,u) > 0\,.$$

A ***Riemannian manifold*** *is a differentiable manifold equipped with a Riemannian metric.*

For each $p \in M$, the inner product g_p induces a norm $|\cdot|$ on T_pM in the usual way:

$$|u|_p \doteqdot \sqrt{g_p(u,u)}\,.$$

We define the function $|\cdot| : TM \to \mathbb{R}$ by $|(p,u)| \doteqdot |u|_p$.

As for any tensor field, a metric can be locally described in terms of its coefficients in a local chart: Let $\{dx^j\}_{j=1}^n$ be the basis of covector fields dual to the coordinate basis $\{\partial_j\}_{j=1}^n$ of some chart. Then, at points of the chart, $g = g_{ij}dx^i \otimes dx^j$, where $g_{ij} = g(\partial_i, \partial_j)$. The smoothness of g is equivalent to the smoothness of all the coefficient functions g_{ij} with respect to some (and hence any) chart. More generally, given a basis $\{e_i\}_{i=1}^{\dim M}$ for T_pM, we can write g with respect to the dual basis $\{\vartheta^i\}_{i=1}^{\dim M}$ for T_p^*M as $g = g_{ij}\vartheta^i \otimes \vartheta^j$, where $g_{ij} = g(e_i, e_j)$. A basis $\{e_i\}_{i=1}^{\dim M}$ for T_pM (or a local field of bases for TM) is called **orthonormal** if $g(e_i, e_j) = \delta_{ij}$.

A Riemannian metric induces a canonical isomorphism, $\cdot^\flat$, between TM and T^*M via the rule

$$u^\flat(v) \doteqdot g(u,v)\,.$$

The inverse of $\cdot^\flat$ is denoted by $\cdot^\sharp$ and the two isomorphisms are sometimes referred to as the **musical isomorphisms**. These isomorphisms extend in a natural way to the homogeneous tensor bundles. For example, we can identify a tensor $T \in T^{(2,0)}M$ with a tensor in the bundle $T^{(1,1)}M$ via the rule

$$T(u,\alpha) \doteqdot T(u,\alpha^\sharp)\,.$$

With respect to dual local bases $\{e_i\}_{i=1}^{\dim M}$ and $\{\vartheta^i\}_{i=1}^{\dim M}$, the musical isomorphisms correspond to **raising** and **lowering of indices** using the metric: denote by g^{ij} the components of the **dual metric** $g \in \Gamma(TM \otimes TM)$ defined by

$$g(\alpha, \beta) \doteqdot g(\alpha^\sharp, \beta^\sharp)\,.$$

Observe that $g_{ik}g^{kj} = \delta^{ij}$. So given $u \in TM$, the components of $u^\flat$ with respect to the dual basis are given by $(u^\flat)_i = u_i \doteqdot g_{ik}u^k$. The $(1, 1)$ tensor defined above is related to its $(2, 0)$ counterpart by $T_i{}^j \doteqdot g^{jk}T_{ik}$. These rules extend in the obvious way to higher degree homogeneous tensors and we will freely use them without mention from here on.

The standard inner product on Euclidean space $\mathbb{R}^n$ induces a Riemannian metric on $\mathbb{R}^n$ via the natural identification of the tangent spaces $T_p\mathbb{R}^n$ with $\mathbb{R}^n$. With respect to the identity chart, its components are δ_{ij}. In fact, the manifold $\mathbb{R}^n$ can be made a Riemannian manifold in many ways: For each $1 \leq i \leq j \leq n$, let f_{ij} be any bounded, smooth function and set $f_{ij} = f_{ji}$ for each $1 \leq j \leq i \leq n$. Then, for C sufficiently large, the matrix corresponding to the functions $g_{ij} = C\delta_{ij} + f_{ij}$ is positive definite everywhere, and hence g_{ij} defines a Riemannian metric.

Any differentiable manifold can be equipped with a metric (in many ways).

Lemma 14.2. *Every smooth manifold carries a Riemannian metric (in fact, many of them).*

Proof. The proof is essentially the same as the proof of existence of connections: Choose an atlas $\{\varphi_\alpha : U_\alpha \to V_\alpha \subset \mathbb{R}^n\}$ and a subordinate partition of unity $\{\rho_\alpha\}$. On each of the regions V_α in $\mathbb{R}^n$, choose a Riemannian metric g^α. Then define, for any $X, Y \in \Gamma(TM)$,

$$g(X, Y) \doteqdot \sum_\alpha \rho_\alpha g^\alpha(d\varphi_\alpha(X), d\varphi_\alpha(Y))\,. \qquad \square$$

Submanifolds and, more generally, immersed submanifolds of Riemannian manifolds carry a natural Riemannian metric induced by the "ambient" Riemannian metric.

Definition 14.3. *Let $F : M \to N$ be an immersion of a manifold M into a Riemannian manifold N with Riemannian metric g. The* ***pullback*** *of g is the metric h on M defined by*

$$h_p(u, v) \doteqdot g_{F(p)}(dF_p(u), dF_p(v))\,.$$

The pullback metric is often denoted by $h = F^*g$.

Definition 14.4. *Let (M, g) and (N, h) be Riemannian manifolds (of the same dimension). A smooth map $\Phi : M \to N$ is called a* ***local isometry***

if $g = \Phi^* h$. A local isometry which is also a diffeomorphism is called an ***isometry****. Two Riemannian manifolds are called (****locally****)* ***isometric*** *if there exists a (local) isometry between them.*

In particular, this endows any submanifold of $\mathbb{R}^n$ with a canonical Riemannian structure. An important example is the sphere S^{n-1}_r of radius $r > 0$.

An important example of a Riemannian manifold which does not arise as an isometrically embedded hypersurface[12] of Euclidean space is the *hyperbolic space*.

Example 14.5. ***Minkowski space*** *$\mathbb{R}^{n,1}$ is the linear space $\mathbb{R}^{n+1}$ equipped with the bilinear form $\eta : \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \to \mathbb{R}$ defined by*

$$\eta(x^0 e_0 + \vec{x}, y^0 e_0 + \vec{y}) \doteqdot -x^0 y^0 + \vec{x} \cdot \vec{y}\,,$$

where $\vec{x} \doteqdot \sum_{i=1}^n x^i e_i \in \mathbb{R}^n$ and $\cdot$ denotes the usual dot product on $\mathbb{R}^n$.

The n-dimensional ***hyperbolic space*** *H^n is the submanifold of unit* ***future directed timelike vectors***

$$\mathcal{J}_+ \doteqdot \{z \in \mathbb{R}^{n+1} : \eta(z,z) = -1 \ \text{ and } \ \eta(z, e_0) < 0\}$$

in Minkowski space equipped with the pullback metric

$$h(u,v) \doteqdot \eta(d\iota(u), d\iota(v))\,,$$

where $\iota : H^n \to \mathbb{R}^{n,1}$ is the inclusion map and, for each $p \in \mathbb{R}^{n+1}$, $T_p\mathbb{R}^{n+1}$ is identified with $\mathbb{R}^{n+1}$ in the usual way.

14.1. Length and distance.

In this section, we introduce the natural metric space structure on a Riemannian manifold determined by measuring the lengths of curves.

A mapping $F : M \to N$ between manifolds is deemed C^k if, given coordinate charts φ and η for M and N respectively, the map $\eta \circ F \circ \varphi^{-1}$ is C^k. If A is a subset of M, a mapping $F : A \to N$ is deemed C^k if there is an open set O of M containing A such that $F : O \to N$ is C^k. A curve $\omega : [a,b] \to M$ is **piecewise** C^k if there are points $a \doteqdot a_0 < a_1 < \cdots < a_l \doteqdot b$ such that $\omega_{[a_{i-1},a_i]}$ is C^k for each $i = 1, \ldots, l$. We denote by $D^k(M)$ the set of piecewise C^k curves from intervals $I \subset \mathbb{R}$ into M.

Definition 14.6. *Let (M,g) be a Riemannian manifold and $\omega : I \to M$ a piecewise C^1 curve. The* ***length*** *$L(\omega)$ of ω is defined as*

$$L(\gamma) \doteqdot \int_a^b |\omega'(t)| dt\,.$$

[12] A **hypersurface** is a submanifold of codimension 1.

Given two points p and q in M, we define the ***distance from*** p ***to*** q *by*

$$d(p,q) \doteqdot \inf_{\omega \in D^1(p,q)} L(\omega)\,,$$

where $D^1(p,q) \doteqdot \{\omega : [a,b] \to M \in D^1(M) : \omega(a) = p \text{ and } \omega(b) = q\}$. *We refer to the corresponding function*[13] $d : M \times M \to \mathbb{R}_+ \cup \{\infty\}$ *as the* ***Riemannian distance function*** *(**induced by** g).*

Lemma 14.7. *Let (M,g) be a Riemannian manifold with induced Riemannian distance function $d : M \times M \to \mathbb{R}$. If M is connected, then (M,d) is a metric space whose induced topology agrees with that of M.*

Proof. The symmetry of the distance function is immediate, as is its non-negativity. The triangle inequality is also easily established: For any curves $\omega_1 : [a_1,b_1] \to M, \omega_2 : [a_2,b_2] \to M \in D^1(M)$ with $\omega_1(b_1) = \omega_2(a_2)$, we can define the concatenation $\omega_1\#\omega_2 : [0, b_1 - a_1 + b_2 - a_2] \to M \in D^1(M)$ by

$$\omega_1\#\omega_2(t) \doteqdot \begin{cases} \omega_1(t+a_1) & \text{if } \ 0 \le t \le b_1 - a_1 \\ \omega_2(t + a_2 - (b_1 - a_1)) & \text{if } \ b_1 - a_1 \le t \le b_1 - a_1 + b_2 - a_2\,. \end{cases}$$

Observe that $\omega_1\#\omega_2$ is a curve of length $L(\omega_1) + L(\omega_2)$. By the definition of L, given points p, q and r in M and any $\varepsilon > 0$ we can choose curves ω_1 joining p to q and ω_2 joining q to r such that $L(\omega_1) < d(p,q) + \varepsilon/2$ and $L(\omega_2) < d(q,r) + \varepsilon/2$ and hence $d(p,r) \le L(\omega_1\#\omega_2) \le d(p,q) + d(q,r) + \varepsilon$. The claim follows.

Consider distinct points $p \neq q \in M$. We need to show that $d(p,q) > 0$. To do so, we will compare d to the Euclidean distance in charts (which will also prove that the induced topology is the correct one). So choose a chart $\varphi : U \to V$ around p. Then we can choose $\delta > 0$ and $\lambda > 0$ such that[14] $\varphi^{-1}(B_{2\delta}(\varphi(p))) \subset U$, $q \notin \varphi^{-1}(B_\delta(\varphi(p)))$ and $g(u,u) \ge \lambda |d\varphi(u)|^2$ on $B_{2\delta}(\varphi(p))$. We can now choose a curve $\omega \in D^1(p,q)$ such that

$$d(p,q) \ge \int |\omega'(t)|\, dt - \frac{\sqrt{\lambda}\delta}{2}\,.$$

If we denote by γ the component of ω in U which contains p, then, since $\phi(q) \notin B_\delta(\phi(p))$, the Euclidean curve $\phi \circ \gamma$ must have length at least δ. We

[13] Recall that the infimum of the empty set is ∞.

[14] The third condition is possible because the left and right hand sides are homogeneous in u: For each $p \in U$, $g(u,u)/|d\varphi(u)|^2$ has a lower bound on the (compact) unit sphere in T_pM. But this lower bound extends to all of T_pM since $g(ru,ru)/|d\varphi(ru)|^2 = g(u,u)/|d\varphi(u)|^2$. We may then take λ to be the minimum of these lower bounds over the compact set $\varphi^{-1}\left(\overline{B_{2\delta}(\varphi(p))}\right)$, where δ is chosen so that this set lies in U.

may therefore estimate

$$\begin{aligned} d(p,q) &\geq \int |\gamma'(t)|\,dt - \frac{\sqrt{\lambda}\delta}{2} \\ &\geq \sqrt{\lambda}\int |d\varphi(\gamma')|\,dt - \frac{\sqrt{\lambda}\delta}{2} \\ &= \sqrt{\lambda}\int |(\varphi\circ\gamma)'|\,dt - \frac{\sqrt{\lambda}\delta}{2} \\ &= \sqrt{\lambda}L(\varphi\circ\gamma) - \frac{\sqrt{\lambda}\delta}{2} \\ &\geq \frac{\sqrt{\lambda}\delta}{2} > 0\,. \end{aligned}$$

This proves that d is a metric. But it also proves that d is comparable from below by the Euclidean distance when viewed through charts. A similar argument shows that it is also comparable from above. Since charts are homeomorphisms, this proves the final claim. □

14.2. The Levi-Civita connection. A Riemannian metric on a manifold M is a section of the tensor bundle $T^{(2,0)}M$. Thus, given a connection on M, we can take the covariant derivative of g. A connection is called **metric compatible** if the covariant differential of g is the zero section. Such a tensor field is also called **parallel** or **covariantly constant**.

Definition 14.8. *A connection ∇ on a Riemannian manifold (M,g) is called **metric(-compatible)** (or g-**compatible**) if*

$$\nabla g = 0\,.$$

Equivalently,

$$X(g(Y,Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$$

for any vector fields X, Y and Z.

Metric compatibility is a natural generalization of the product rule for inner products on $\mathbb{R}^n$.

Theorem 14.9 (T. Levi-Civita (1929))**.** *Let (M,g) be a Riemannian manifold. There exists a unique symmetric, metric connection ∇ on M.*

Proof. Assuming the conditions, we will derive an explicit formula for ∇, proving the uniqueness. So consider, for any vector fields X, Y and Z,

$$\begin{aligned}
&Xg(Y,Z) + Yg(X,Z) - Zg(X,Y) \\
&\qquad = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) + g(\nabla_Y X, Z) + g(X, \nabla_Y Z) \\
&\qquad\quad - g(\nabla_Z X, Y) - g(X, \nabla_Z Y) \\
&\qquad = 2g(\nabla_X Y, Z) + g(Y, \nabla_X Z - \nabla_Z X) + g(\nabla_Y X - \nabla_X Y, Z) \\
&\qquad\quad + g(X, \nabla_Y Z - \nabla_Z Y) \\
&\qquad = 2g(\nabla_X Y, Z) + g(Y, [X,Z]) + g([Y,X], Z) + g(X, [Y,Z])\,.
\end{aligned}$$

That is,

$$\begin{aligned}
2g(\nabla_X Y, Z) &= Xg(Y,Z) + Yg(X,Z) - Zg(X,Y) \\
&\qquad + g([X,Y],Z) + g([Z,X],Y) - g([Y,Z],X)\,.
\end{aligned} \tag{14.1}$$

The right-hand side is independent of ∇, so this proves uniqueness.

Existence follows by checking that the formula (14.1) determines a symmetric, metric connection, which we leave as an exercise. □

The connection uniquely determined by (14.1) is called the **Levi-Civita connection** of (M, g).

If we define, with respect to some local field of bases $\{e_i\}_{i=1}^n$ for TM, the **connection coefficients**

$$\Gamma_{ijk} \doteqdot g(\nabla_i e_j, e_k)$$

and the **structure coefficients**

$$C_{ijk} \doteqdot g([e_i, e_j], e_k) = g(\mathcal{L}_i e_j, e_k)$$

then the Levi-Civita formula (14.1) becomes

$$2\Gamma_{ijk} = e_i g_{jk} + e_j g_{ik} - e_k g_{ij} + C_{ijk} - C_{jki} + C_{kij}\,.$$

In particular, with respect to a coordinate basis, the connection coefficients are determined by derivatives of the metric components,

$$2\Gamma_{ijk} = \partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}\,, \tag{14.2}$$

and, with respect to an orthonormal basis, the connection coefficients are determined by the commutation relations,

$$2\Gamma_{ijk} = C_{ijk} - C_{jki} + C_{kij}\,. \tag{14.3}$$

The connection is also uniquely determined by the **connection one-forms** $\{\omega_i{}^j\}_{i,j=1}^n$, which are determined by

$$\omega_i{}^j(e_k) \doteqdot -(\nabla_k \theta^j)(e_i)\,,$$

where $\{\theta^i\}_{i=1}^n$ is the local field of bases for T^*M dual to $\{e_i\}_{i=1}^n$. Indeed, since $\theta^j(e_i) \equiv \delta_i^j$,

$$\begin{aligned} {\omega_j}^k(e_i) &= -(\nabla_i\theta^k)(e_j) \\ &= -e_i(\theta^k(e_j)) + \theta^k(\nabla_i e_j) \\ &= {\Gamma_{ij}}^k . \end{aligned}$$

In Euclidean space, there is a (global) coordinate chart in which

$$g_{ij} \equiv \delta_{ij}$$

and hence

$$\Gamma_{ijk} = 0 .$$

For a general Riemannian manifold, no such orthonormal coordinates exist; however, it is always possible to choose coordinates so that these two identities hold *at a given point*.

Observe that the Levi-Civita connection has a total of n^3 linearly independent coefficients ${\Gamma_{ij}}^k$ with respect to a general local field of bases. This reduces to $\frac{n^2(n+1)}{2}$ if the basis is a coordinate basis, since, in that case, we have the symmetry

$${\Gamma_{ij}}^k = {\Gamma_{ji}}^k .$$

The number reduces to $\frac{n^2(n-1)}{2}$ if the basis is orthonormal, since, in that case, we have the skew-symmetry

$$\Gamma_{ijk} + \Gamma_{ikj} = 0 .$$

Similarly, with respect to a general basis, there are n^2 connection one-forms, while, with respect to an orthonormal basis, this reduces to $\frac{n(n-1)}{2}$. With respect to an coordinate basis, we still have n^2 connection one-forms, but their number of independent components is reduced from n^3 to $\frac{n^2(n+1)}{2}$.

Proposition 14.10 (Exponential normal coordinates)**.** *Let (M^n, g) be a smooth Riemannian manifold equipped with its Levi-Civita connection ∇. Given $p \in M$ and an orthonormal basis $\{e_i\}_{i=1}^n$ for T_pM we may define a chart φ on a neighborhood U of p by*

$$\varphi^{-1}(x^1, \ldots, x^n) = \exp_p x^i e_i .$$

With respect to this chart,

$$g_{ij}(p) = \delta_{ij} \quad \textit{and} \quad {\Gamma_{ij}}^k(p) = 0 .$$

The coordinates defined by φ are called **exponential normal coordinates**.

Proof of Proposition 14.10. By Theorem 12.5, there exists $U \subset M$ such that φ is a well-defined chart for M. Since the derivative of the exponential map at the origin is the identity map, we find that

$$\partial_i|_p \doteqdot d\varphi^{-1}|_{\varphi(p)} e_i = (d\exp_p)_0 e_i = e_i$$

and hence, in particular,

$$g_{ij}(p) = \delta_{ij}\,.$$

Next, observe that, given any constants a^i, $i = 1, \ldots, n$, the vector $a^i\partial_i$ is tangent to the geodesic $t \mapsto \exp_p(ta^i e_i)$ along γ. It follows that

$$\nabla_{a^i\partial_i}(a^j\partial_j) = 0$$

at points $\exp_p(ta^i e_i)$. In particular, this holds at p. But then, by symmetry of the connection, we find at p

$$\begin{aligned} 0 = \nabla_{\partial_i+\partial_j}(\partial_i + \partial_j) &= \nabla_i\partial_i + \nabla_i\partial_j + \nabla_j\partial_i + \nabla_j\partial_j \\ &= \nabla_i\partial_j + \nabla_j\partial_i \\ &= 2\nabla_i\partial_j\,. \end{aligned}$$

This completes the proof. $\square$

Exercises.

Exercise 14.1. *Given $u, v \in \mathbb{R}^{n,1}$ satisfying $\eta(u,u) < 0$ and $\eta(v,v) < 0$, show that*

$$|\eta(u,v)| \geq |u||v|\,,$$

where $|u| \doteqdot \sqrt{-\eta(u,u)}$ (and similarly for v).

Exercise 14.2. *Show that (H^n, h) is a Riemannian manifold. Hint: Let $\omega : I \to H^n$ be a smooth curve and use the defining relation for H^n to determine the form of $u \doteqdot \omega'(0)$. Then check that $h(u,u) > 0$ for $u \neq 0$.*

Exercise 14.3. *Consider the* ***hyperbolic stereographic projection*** $\Phi : H^n \to B_1(0) \subset \mathbb{R}^n = \{0\} \times \mathbb{R}^n \subset \mathbb{R}^{n,1}$ *defined as follows: Given a point $z \in H^n = \{(t,x) \in \mathbb{R} \times \mathbb{R}^n : t = \sqrt{1+|x|^2}\}$, where $|\cdot|$ is the norm in $\mathbb{R}^n$, let $\Phi(z)$ be the point of intersection of the line from z to the point $-e_0$ with the plane $\{0\} \times \mathbb{R}^n$.*

Show that

$$\Phi^{-1}(x) = \left(\frac{1+|x|^2}{1-|x|^2}, \frac{2x}{1-|x|^2}\right).$$

Show that the pullback metric g induced on the unit ball by Φ^{-1} is given (in standard rectilinear coordinates) by

$$g_{ij} = \frac{4}{(1-|x|^2)^2}\delta_{ij}\,.$$

The Riemannian manifold $(B_1(0), g)$ is called the ***Poincaré disk****.*

Exercise 14.4. *Let (M, g) be a Riemannian manifold. Show that the length of a smooth curve $\omega : I \to M$ is invariant under reparametrization; that is, if $r : J \to I$ is a diffeomorphism of the intervals I and J, then $L(\omega) = L(\omega \circ r)$.*

Exercise 14.5. *Let (M, g) be a Riemannian manifold with metric compatible connection ∇.*

(a) *Show that lengths and angles are preserved by parallel translation. That is,*
$$\partial_t \left[g(E(t), F(t)) \right] = 0$$
for any parallel vector fields $E, F \in \Gamma(\omega^* TM)$ along a smooth curve $\omega : I \to M$.

(b) *Show that any connection for which the lengths and angles between parallel vector fields are constant must be compatible with the metric.*

15. Convexity and completeness

15.1. The Gauss Lemma.

Let (M, g) be a Riemannian manifold of dimension n. Given $p \in M$, any orthonormal basis $\{e_i\}_{i=1}^n$ for T_pM induces an isomorphism from $\mathbb{R}^n$ to T_pM via the identification $(v^1, \ldots, v^n) \mapsto v^i e_i$. This in turn induces a map $\Phi : (\xi, r) \mapsto \exp_p(r\xi)$ from $S^{n-1} \times [0, r_0)$ into M for some $r_0 > 0$. We refer to the pair (ξ, r) as **geodesic polar coordinates** (**centered at** p) of the point $\exp_p r\xi \in M$.

Let ∂_r be the canonical vector field on $T(S^{n-1} \times (0, r_0)) = TS^n \oplus T(0, r_0)$. Observe that

$$d\Phi_{(\xi,r)}\partial_r = \gamma_\xi'(r)$$

where $\gamma_\xi(r) \doteqdot \exp_p(r\xi)$, and, given $\zeta \in T_\xi S^{n-1}$,

$$d\Phi_{(\xi,r)}\zeta = (d\exp_p)_{r\xi}\zeta\,.$$

Lemma 15.1 (Gauss (1825))**.** *Let (M, g) be a Riemannian manifold with Levi-Civita connection ∇. In geodesic polar coordinates,*

$$\Phi^* g(\partial_r, \partial_r) = 1$$

and

$$\Phi^* g(\partial_r, \zeta) = 0$$

for all $\zeta \in TS^{n-1}$.

In other words, the image under the derivative of the exponential map of the unit radial vector in T_pM is always a unit vector, and the image of a vector tangent to a sphere about the origin is always orthogonal to the image of the radial vector.

Proof of Lemma 15.1. The first identity is straightforward: for each $\xi \in S^{n-1}$ the curve $r \mapsto \gamma_\xi(r) \doteqdot \exp_p(r\xi)$ is a geodesic with initial length

$$|\gamma_\xi'(0)| = |\xi| = 1\,.$$

The claim follows since

$$\frac{d}{dr}\left[\Phi^* g(\partial_r, \partial_r)\right] = \frac{d}{dr} g(\gamma_\xi', \gamma_\xi') = 2g(\nabla_r \gamma_\xi', \gamma_\xi') = 0\,.$$

To prove the second claim, define a one-parameter family $\omega : [0, r_0) \times (-\varepsilon_0, \varepsilon_0) \to M$ of geodesics $\gamma(\cdot, \varepsilon)$ by $\omega(r, \varepsilon) \doteqdot \exp_p(r\xi(\varepsilon))$, where ξ is a curve $\xi : (-\varepsilon_0, \varepsilon_0) \to S^{n-1}$ with $(\xi(0), \xi'(0)) = (\xi_0, \zeta)$. Then

$$\begin{aligned}\frac{d}{dr} g(\partial_r \omega, \partial_\varepsilon \omega) &= g(\partial_r \omega, \nabla_r \partial_\varepsilon \omega) \\ &= g(\partial_r \omega, \nabla_\varepsilon \partial_r \omega) \\ &= \frac{1}{2} \partial_\varepsilon g(\partial_r \omega, \partial_r \omega) = 0\,.\end{aligned}$$

Thus, $g(\partial_r \omega, \partial_\varepsilon \omega)$ is constant in r. In particular,

$$\Phi^* g(\partial_r, \zeta) = g(\partial_r \omega, \partial_\varepsilon \omega)|_{(r,0)} = g(\partial_r \omega, \partial_\varepsilon \omega)|_{(0,0)}\,.$$

The claim follows since $\partial_\varepsilon \omega(0, \varepsilon) = 0$ for all ε. □

An important consequence of the Gauss Lemma is the fact that geodesics of the Levi-Civita connection, restricted to sufficiently short intervals, have smaller length than any other path between their endpoints. We say that geodesics *locally minimize distance*. Moreover, if the distance between two points is attained as the length of some curve, then this curve must be a geodesic.

Proposition 15.2. *Let (M, g) be a Riemannian manifold equipped with its Levi-Civita connection ∇ and let $\gamma : I \to M$ be a geodesic. For any $t_0 \in I$ there exists $\delta > 0$ such that*

$$L(\gamma|_{[t_0-\delta, t_0+\delta]}) = d(\gamma(t_0 - \delta), \gamma(t_0 + \delta))\,.$$

Conversely, let $\omega : [0,1] \to M$ be a piecewise smooth path for which $L(\omega) = d(\omega(0), \omega(1))$. Then $\omega = \gamma \circ r$, where $r : [0,1] \to [0,1]$ is monotone and γ is a geodesic.

Note that we cannot expect that geodesics minimize length on long intervals — consider the example of the sphere S^2: geodesics are great circles, and these achieve the distance between their endpoints on intervals of length no greater than π, but not on longer intervals.

Proof of Proposition 15.2. Let $\gamma : I \to M$ be a (without loss of generality, unit speed) geodesic. Fix $t_0 \in I$ and choose $\delta > 0$ sufficiently small that $[t_0 - \delta, t_0 + \delta] \subset I$ and $\exp_{\gamma(t_0-\delta)}$ is a diffeomorphism on a ball in $T_{\gamma(t_0-\delta)}M$ of radius $r > 2\delta$ about the origin. For convenience, we denote by p the point $\gamma(t_0 - \delta)$ and by q the point $\gamma(t_0 + \delta)$.

Now let $\omega : J \to M$ be any other curve joining the points p and q. We need to show that $L(\omega) \geq L(\gamma|_{[t_0-\delta, t_0+\delta]}) = 2\delta$. Suppose first that ω remains in the set $\exp_p(B_{2\delta}(0))$. Then we can write ω in geodesic polar coordinates based at p as $\omega(t) = \exp_p(r(t)\xi(t))$ for some function $r : J \to [0, 2\delta]$ (which maps the endpoints of J to the endpoints of $[0, 2\delta]$) and some $\xi : J \to S_pM$. Then, by the Gauss Lemma,

$$|\omega'(t)|^2 = |(d\exp_p)_{r\xi}(r'\xi + r\xi')|^2 = (r')^2 + r^2|(d\exp_p)_{r\xi}\xi'|^2 \geq (r')^2\,.$$

In particular,

$$L(\omega) = \int_J |\omega'(t)|dt \geq \int_J |r'(t)|\,dt \geq \int_0^{2\delta} dr = 2\delta = L(\gamma)\,. \tag{15.1}$$

This proves the first claim as long as ω does not leave the set $\exp_p(B_r(0))$. But should γ leave this set, the same argument applies on the portion of ω joining p to its boundary, giving $L(\omega) > 2\delta$.

To prove the second claim, first observe that ω achieves the distance between any pair of its points. Indeed, if there were some subinterval on which this were not true, then replacing ω by a shorter path on that subinterval would also yield a shorter path from $\omega(0)$ to $\omega(1)$. The claim now follows by observing that equality is attained in (15.1) only if $\xi' = 0$ and r is monotone, in which case ω is simply a monotone reparametrization of γ. $\square$

15.2. Convex neighborhoods.

Next, we show that the distance $d(p, q)$ between two points p and q is achieved by a length minimizing geodesic, so long as p and q are sufficiently close. In fact, we prove slightly more (cf. Corollary 12.6).

Proposition 15.3 (Convex neighbourhoods)**.** *Let (M, g) be a Riemannian manifold equipped with its Levi-Civita connection ∇. For any $p \in M$ there exist constants $0 < \varepsilon \leq \eta$ such that every pair of points q and r in the ball $B_\varepsilon(p) \doteqdot \{x \in M : d(p, x) < \varepsilon\}$ is joined by a unique geodesic $\gamma_{qr} : [0, 1] \to M$ of length $L(\gamma_{qr}) = d(q, r) < \eta$.*

Moreover, $d(p, \gamma_{qr}(t)) \leq \max\{d(p, q), d(p, r)\}$ for all $t \in [0, 1]$. In particular, $\gamma_{qr}([0, 1]) \subset B_\varepsilon(p)$.

Proof. By Theorem 12.5 the map $(\pi \times \exp) : \mathcal{T}M \to M \times M$ is a diffeomorphism from a neighborhood of $(p, 0)$ into $M \times M$. Thus, as in the proof of Corollary 12.6, we may choose η sufficiently small that $(\pi \times \exp)$ is a diffeomorphism on the set $\mathcal{O}_\eta \doteqdot \{(x, v) \in TM : d(x, p) < \eta,\ |v| < \eta\}$. Setting $B_\varepsilon(p) \doteqdot \{x \in M : d(x, p) < \varepsilon\}$, now choose ε so small that $B_\varepsilon(p) \times B_\varepsilon(p) \subset (\pi \times \exp)(\mathcal{O}_\eta)$. In particular, for any $q, r \in B_\varepsilon(p)$, there exists a unique geodesic γ_{qr} of length less than η joining q to r. Moreover, applying the Gauss Lemma as in the proof of Proposition 15.2, we find that γ_{qr} achieves the distance between its endpoints.

It remains to prove that $d(p, \gamma_{qr}(t)) \leq \max\{d(p,q), d(p,r)\}$ for all $t \in [0,1]$. Observe that, in exponential normal coordinates φ based at p,

$$d(p,x) = d(p, \exp_p(x^i e_i)) = |\gamma'(0)| = \sum_{i=1}^n (x^i)^2 \,,$$

where $\gamma(t) \doteqdot \exp(t x^i e_i)$. Thus, writing $\gamma_{qr}(t) = \exp_p(x^i(t) e_i)$ and using the geodesic equation,

$$\begin{aligned}
\frac{d^2}{dt^2} d(p, \gamma_{qr}(\cdot)) = \frac{d^2}{dt^2} \sum_{i=k}^n (x^k)^2 &= 2 \frac{d}{dt} \sum_{k=1}^n x^k \dot{x}^k \\
&= 2 \sum_{k=1}^n \left((\dot{x}^k)^2 + x^k \ddot{x}^k \right) \\
&= 2 \sum_{k=1}^n \left((\dot{x}^k)^2 - x^k \dot{x}^i \dot{x}^j {\Gamma_{ij}}^k \right) \,.
\end{aligned}$$

Now, in exponential normal coordinates, ${\Gamma_{ij}}^k(0) = 0$ and hence, for η sufficiently small, we can arrange that[15]

$$\left| \sum_{i,j,k} x^k {\Gamma_{ij}}^k \xi_i \xi_j \right| \leq \frac{1}{2} \sum_k (\xi^k)^2$$

whenever $\sum_k (x^k)^2 < 2\eta$. Note that, since γ_{qr} is of length less than η and since q and r are within distance $\varepsilon < \eta$ of p, the entire geodesic γ_{qr} stays within distance 2η of p. So we may estimate

$$\frac{d^2}{dt^2} d(p, \gamma_{qr}(\cdot)) \geq \sum_{k=1}^n (\dot{x}^k)^2 > 0 \,.$$

Thus, the function $t \mapsto d(p, \gamma_{qr}(t))$ is strictly convex; in particular, its maximum must occur at one of the endpoints. $\square$

15.3. The Hopf–Rinow Theorem.

A Riemannian manifold is called **geodesically complete** if $\mathcal{T}M = TM$. This means that the geodesic flow is defined for all $t \in \mathbb{R}$ or, in other words, that geodesics can be extended indefinitely. The Hopf–Rinow Theorem states that a Riemannian manifold is geodesically complete if and only if it is complete as a metric space when equipped with the Riemannian distance function.

[15] This follows again from the homogeneity of the right and left hand sides in ξ and compactness of the unit sphere.

Theorem 15.4 (H. Hopf and W. Rinow[16] (1931))**.** *Let (M, g) be a connected Riemannian manifold equipped with its Levi-Civita connection and its Riemannian distance function d. The following are equivalent:*

(1) (M, d) is a complete metric space;

(2) (M, g) is geodesically complete;

(3) There exists $p \in M$ for which $\mathcal{T}M \cap T_pM = T_pM$.

Furthermore, each of these conditions implies

(*) For every p and q in M, there exists a length minimizing geodesic joining p and q.

Proof. Observe that (2) immediately implies (3). We prove the following implications: (1) $\Longrightarrow$ (2), (3) $\Longrightarrow$ $(*_p)$ and $((3) + (*_p)) \Longrightarrow$ (1), where $(*_p)$ is the statement of $(*)$ but with the point p fixed and satisfying (3).

(1) $\Longrightarrow$ (2): Suppose that (M, d) is a complete metric space. If (2) does not hold, then there is a point $\xi = (p, u) \in TM$ with $|u| = 1$ such that the maximal interval I_ξ of existence of the geodesic $\gamma(t) \doteqdot \exp(t\xi)$ is not all of $\mathbb{R}$. Without loss of generality, we can assume that $T \doteqdot \sup I_\xi < \infty$. Observe that

$$d(\gamma(s), \gamma(t)) \leq L(\gamma\big|_{[s,t]}) = |t - s|$$

for any $s, t \in I_\xi$. That is, $\gamma(t)$ is Cauchy as $t \to T$ and hence converges to some limit $q \in M$. Extend the curve γ to T by setting $\gamma(T) = q$. We need to show that the tangent vector has a limit as $t \to T$.

Choose $\varepsilon > 0$ such that the ball $B_\varepsilon(q)$ is convex in the sense of Proposition 15.3. Then, for any $s, t \in (T - \varepsilon, T)$, $\gamma(s)$ and $\gamma(t)$ lie in $B_\varepsilon(q)$, and γ achieves the distance between them:

$$d(\gamma(s), \gamma(t)) = L(\gamma|_{[s,t]}) = |t - s| \text{ for all } s, t \in (t - \varepsilon, T).$$

Since d is continuous,

$$d(\gamma(s), q) = \lim_{t \to T} d(\gamma(s), \gamma(t)) = \lim_{t \to T} (t - s) = T - s = L(\gamma|_{[s,T]}).$$

So $\gamma|_{[s,T]}$ is a length minimizing path and hence a geodesic emanating from q:

$$\gamma|_{[s,T]}(T - t) = \exp_q(t\eta)$$

[16]Willi Rinow, who received his PhD in Berlin in 1932 under the direction of Heinz Hopf and Ludwig Bieberbach, joined the Nazi party in 1937 and worked as a cryptanalyst for the Nazi war effort. Hopf, whose father was born Jewish, was forced to file for Swiss citizenship in 1940 after his property was confiscated by the Nazis. (By 1933, Bieberbach had become a fervent Nazi, and was enthusiastically involved in the efforts to dismiss his Jewish colleagues and promote "Deutsche Mathematik", even facilitating the Gestapo arrest, and later execution, of Juliusz Schauder.)

for some $\eta \in S_qM$. Local existence of solutions to the geodesic equation now ensures that this geodesic exists for sufficiently small values of $t < 0$, which extends γ beyond T.

Next, we prove that (3) $\Longrightarrow$ ($*_p$): Suppose that $\mathcal{T}M \cap T_pM = T_pM$. Fix $q \neq p$. We need to show that there is a length minimizing geodesic joining p with q. Choose $\varepsilon > 0$ such that $\exp_p$ is a diffeomorphism on a set containing the closure of $B_\varepsilon(0_p) \subset T_pM$. If $q \in B_\varepsilon(p)$ then we are done, so assume not. Observe that $S_\varepsilon(p) \doteqdot \{q \in M : d(p,q) = \varepsilon\}$ is compact (it is the image of the compact set $S_\varepsilon(0_p) \doteqdot \{v \in T_pM : |v| = \varepsilon\}$ under the continuous map $\exp_p$). In particular, the function $d(\cdot, q)$ attains its minimum on this set. In other words, there exists $r = \exp_p(\varepsilon v)$ such that

$$d(r,q) = d(S_\varepsilon(p), q) \doteqdot \inf\{d(r', q) : r' \in S_\varepsilon(p)\}\,.$$

We claim that

$$d(p,q) = \varepsilon + d(r,q)\,. \tag{15.2}$$

Indeed, the inequality $d(p,q) \leq \varepsilon + d(r,q)$ follows from the triangle inequality since $d(p,r) = \varepsilon$, and the reverse inequality follows because any path from p to q must pass through $S_\varepsilon(p)$. Set $\gamma(t) \doteqdot \exp_p tv$ and let J be the set of times

$$J \doteqdot \{t \in [0, d(p,q)] : t + d(\gamma(t), q) = d(p,q)\}\,.$$

We need to show that J contains the point $d(p,q)$ — taking $t = d(p,q)$ will then imply that $d(\gamma(d(p,q)), q) = 0$ and hence $q = \gamma(d(p,q))$, which proves our claim.

First observe that J is non-empty (by (15.2)) and closed (because the distance function is continuous and $\gamma(t)$ exists for all $t \in \mathbb{R}$). Next, we claim that $[0,t] \subset J$ whenever $t \in J$. Indeed, if $s < t$, the triangle inequality and the fact that $\gamma|_{[0,t]}$ is a minimising path imply that

$$\begin{aligned}
d(p,q) &= t + d(\gamma(t), q)\\
&= s + (t-s) + d(\gamma(t), q)\\
&\geq s + d(\gamma(s), q)\\
&\geq d(p,q)\,.
\end{aligned}$$

Finally, we will prove that J is open relative to $[0, d(p,q)]$, which will prove that $J = [0, d(p,q)]$ and hence $d(p,q) \in J$ as desired.

So fix $\tau \in J$ and set $x = \gamma(\tau)$. Choose $\delta > 0$ such that $\exp_x$ is a diffeomorphism on $B_\delta(0)$. If $q \in B_\delta(x)$ then $q = \exp_x(d(x,q)w)$ for some $w \in S(0_x)$. Consider the curve

$$\omega(t) \doteqdot \begin{cases} \gamma(t) & \text{for } t \leq \tau \\ \exp_x(-(\tau - t)w) & \text{for } t \geq \tau\,. \end{cases} \tag{15.3}$$

Observe that ω is a curve joining p to q, and, by assumption,

$$d(p,q) = \tau + d(x,q) = L(\omega)\,.$$

Thus, ω is a minimising curve and hence a geodesic. Hence, by uniqueness, $\omega(t) = \exp_p(tv)$ for all $t \in [0, d(p,q)]$ and we are done.

If, instead, $q \notin B_\delta(x)$ then we can choose $y \in S_\delta(x)$ such that $d(y,q) = d(B_\delta(x), q)$. Then, as before,

$$d(x,q) = \delta + d(y,q)$$

and hence, by assumption

$$d(p,q) = d(p,x) + d(x,q) = \tau + \delta + d(y,q)\,.$$

Let ω be the unit speed curve given by following γ from p to x and then following the radial geodesic from x to y. Then $L(\omega) = d(p,x) + \delta$ and hence

$$d(p,q) = L(\omega) + d(y,q)\,.$$

It follows that $L(\omega) = d(p,y)$ since if there were any shorter path ω' from p to y we would have

$$d(p,q) \leq L(\omega') + d(y,q) < L(\omega) + d(y,q) = d(p,q)\,,$$

a contradiction. Therefore ω is a geodesic and hence $\omega(t) = \gamma(t)$. Thus, $[0, \tau+\delta] \subset J$, and hence J is open, as claimed, which completes the proof of the second implication.

Finally, we will prove that $((3)+(*_p)) \implies (1)$: By (3), we may define for each $k \in \mathbb{N}$ the set $M_k \doteqdot \exp_p(\overline{B}_k(0_p))$. This set is the image of a compact set under a continuous map, and is therefore compact. By $(*_p)$, $\cup_{k\in\mathbb{N}} M_k = M$. Now, if $\{q_i\}_{i\in\mathbb{N}}$ is a Cauchy sequence in M, then $\{q_i\}_{i\in\mathbb{N}}$ is contained in M_k for some k. But since M_k is compact, it follows that $\{q_i\}_{i\in\mathbb{N}}$ converges. This completes the proof of the Theorem. $\square$

16. Riemannian curvature

We have already been introduced to the curvature tensor of a general connection. In this section, we will be interested in the additional properties of the curvature tensor of the Levi-Civita connection on a Riemannian manifold. From here on, unless otherwise stated, a Riemannian manifold will always be equipped with its Levi-Civita connection.

As usual, we can decompose the $(3,1)$-tensor $X, Y, Z \mapsto \mathrm{Rm}(X,Y)Z$ into components with respect to some tangent basis $\{e_i\}_{i=1}^{\dim M}$:

$$\mathrm{Rm} = \mathrm{Rm}_{ijk}{}^{l} \theta^i \otimes \theta^j \otimes \theta^k \otimes e_l \,,$$

where $\{\theta^i\}_{i=1}^{\dim M}$ is the cotangent basis dual to $\{e_i\}_{i=1}^{\dim M}$.

Using the metric, we can identify Rm with a tensor of degree $(4,0)$ via

$$\mathrm{Rm}(W,X,Y,Z) \doteqdot g(\mathrm{Rm}(W,X)Y,Z) \,.$$

Or, in components,

$$\mathrm{Rm}_{ijkl} = g_{lm}\, \mathrm{Rm}_{ijk}{}^{m}.$$

Proposition 16.1. *Let (M,g) be a Riemannian manifold. The curvature tensor $\mathrm{Rm} \in \Gamma(T^{(4,0)}M)$ has the following symmetries:*

(i) $\mathrm{Rm}(W,X,Y,Z) + \mathrm{Rm}(X,W,Y,Z) = 0$*;*

(ii) $\mathrm{Rm}(W,X,Y,Z) + \mathrm{Rm}(X,Y,W,Z) + \mathrm{Rm}(Y,W,X,Z) = 0$*;*

(iii) $\mathrm{Rm}(W,X,Y,Z) + \mathrm{Rm}(W,X,Z,Y) = 0$*;*

(iv) $\mathrm{Rm}(W,X,Y,Z) = \mathrm{Rm}(Y,Z,W,X)\,.$

Proof. The first identity is obvious from the definition of the curvature operator and the second is just the first Bianchi identity. The third symmetry is a consequence of the compatibility of the connection with the metric, and the fact that the curvature operator is a derivation on the tensor algebra which commutes with contractions (Lemma 13.5) and vanishes on functions:

$$\begin{aligned} 0 = \mathrm{Rm}(W,X)(g(Y,Z)) &= g(\mathrm{Rm}(W,X)Y,Z) + g(Y,\mathrm{Rm}(W,X)Z) \\ &= \mathrm{Rm}(W,X,Y,Z) + \mathrm{Rm}(W,X,Z,Y)\,. \end{aligned}$$

The final identity is a consequence of the previous three:

$$\begin{aligned}
\mathrm{Rm}(W,X,Y,Z) &= -\mathrm{Rm}(X,Y,W,Z) - \mathrm{Rm}(Y,W,X,Z)\\
&= \mathrm{Rm}(X,Y,Z,W) + \mathrm{Rm}(Y,W,Z,X)\\
&= -\mathrm{Rm}(Y,Z,X,W) - \mathrm{Rm}(Z,X,Y,W)\\
&\quad - \mathrm{Rm}(W,Z,Y,X) - \mathrm{Rm}(Z,Y,W,X)\\
&= 2\,\mathrm{Rm}(Y,Z,W,X) + \mathrm{Rm}(Z,X,W,Y) + \mathrm{Rm}(W,Z,X,Y)\\
&= 2\,\mathrm{Rm}(Y,Z,W,X) - \mathrm{Rm}(X,W,Z,Y)\\
&= 2\,\mathrm{Rm}(Y,Z,W,X) - \mathrm{Rm}(W,X,Y,Z)\,.
\end{aligned}$$

Rearranging yields the the claim. □

Note that if M is a one-dimensional Riemannian manifold then its curvature is zero (since the curvature tensor is skew-symmetric). This reflects the fact that any one-dimensional Riemmannian manifold can be locally parametrised by arc length, and so is locally isometric to any other one-dimensional manifold.

Next consider the two-dimensional case: Any component of Rm in which the first two or the last two entries are the same must vanish, by the anti-symmetries in these components. There is therefore only one independent component of the curvature: If we take $\{e_1, e_2\}$ to be an orthonormal basis for T_pM, then we define the **Gaussian curvature** K of M at p as $K(p) = \mathrm{Rm}(e_1, e_2, e_1, e_2)$. This is independent of the choice of orthonormal basis: Any other one $\{e_1', e_2'\}$ is given by $e_1' = \cos\theta e_1 + \sin\theta e_2$ and $e_2' = -\sin\theta e_1 + \cos\theta e_2$ for some θ and hence

$$\begin{aligned}
\mathrm{Rm}(e_1', e_2', e_1', e_2') &= \mathrm{Rm}(\cos\theta e_1 + \sin\theta e_2, -\sin\theta e_1 + \cos\theta e_2, e_1', e_2')\\
&= \cos^2\theta\,\mathrm{Rm}(e_1, e_2, e_1', e_2') - \sin^2\theta\,\mathrm{Rm}(e_2, e_1, e_1', e_2')\\
&= \mathrm{Rm}(e_1, e_2, e_1', e_2')\,.
\end{aligned}$$

Using the pairwise symmetry (part (iv) of Proposition 16.1) and applying the same argument again, we conclude

$$\mathrm{Rm}(e_1', e_2', e_1', e_2') = \mathrm{Rm}(e_1, e_2, e_1, e_2)\,.$$

Definition 16.2. *Let (M, g) be a Riemannian manifold. Given $p \in M$ and a two-dimensional subspace $\Sigma_p \subset T_pM$, we define the **sectional curvature** $K(\Sigma_p)$ as*

$$K(\Sigma_p) \doteqdot \frac{\mathrm{Rm}(v_1, v_2, v_1, v_2)}{|v_1|^2|v_2|^2 - g(v_1, v_2)^2}\,,$$

where $\{v_1, v_2\}$ is any basis for Σ_p.

If u and v are linearly independent tangent vectors at some point, then we denote by $u \wedge v$ the (oriented) plane spanned by u and v.

Proposition 16.3. *The curvature tensor is completely determined by the sectional curvatures.*

Proof. Working in an orthonormal basis for T_pM at a point $p \in M$, it suffices to find an explicit expression for any component Rm_{ijkl} in terms of sectional curvatures.

Let $\{e_i\}_{i=1}^n$ be an orthonormal basis for T_pM. It will be convenient to refer to the oriented plane generated by two distinct basis vectors e_i and e_j using the notation $e_i \wedge e_j$. We first compute the sectional curvature of the plane $\frac{1}{2}(e_i + e_k) \wedge (e_j + e_l)$:

$$\begin{aligned}
K\Big(\frac{1}{2}(e_i + e_k) \wedge (e_j + e_l)\Big) \\
&= \frac{1}{4}\,\mathrm{Rm}(e_i + e_k, e_j + e_l, e_i + e_k, e_j + e_l) \\
&= \frac{1}{4}K(e_i \wedge e_j) + \frac{1}{4}K(e_i \wedge e_l) + \frac{1}{4}K(e_j \wedge e_k) + \frac{1}{4}K(e_j \wedge e_l) \\
&\quad + \frac{1}{2}\,\mathrm{Rm}_{ijil} + \frac{1}{2}\,\mathrm{Rm}_{ijkj} + \frac{1}{2}\,\mathrm{Rm}_{ilkl} + \frac{1}{2}\,\mathrm{Rm}_{klkj} \\
&\quad + \frac{1}{2}\,\mathrm{Rm}_{ijkl} + \frac{1}{2}\,\mathrm{Rm}_{kjil}\,.
\end{aligned}$$

Replacing e_k and e_l by $-e_k$ and $-e_l$ respectively yields the sectional curvature of the plane $\frac{1}{2}(e_i - e_k) \wedge (e_j - e_l)$, which we add to the previous expression to obtain:

$$\begin{aligned}
\mathrm{Rm}_{ijkl} + \mathrm{Rm}_{kjil} &= K\left(\frac{1}{2}(e_i + e_k) \wedge (e_j + e_l)\right) + K\left(\frac{1}{2}(e_i - e_k) \wedge (e_j - e_l)\right) \\
&\quad - \frac{1}{2}K(e_i \wedge e_j) - \frac{1}{2}K(e_i \wedge e_l) - \frac{1}{2}K(e_j \wedge e_k) - \frac{1}{2}K(e_k \wedge e_l)\,.
\end{aligned}$$

Finally, subtract the same expression with e_i and e_j interchanged and apply the Bianchi identity to obtain

$$\begin{aligned}
3\,\mathrm{Rm}_{ijkl} &= \mathrm{Rm}_{ijkl} + \mathrm{Rm}_{kjil} - \mathrm{Rm}_{jikl} - \mathrm{Rm}_{kijl} \\
&= K\left(\frac{1}{2}(e_i + e_k) \wedge (e_j + e_l)\right) + K\left(\frac{1}{2}(e_i - e_k) \wedge (e_j - e_l)\right) \\
&\quad - K\left(\frac{1}{2}(e_j + e_k) \wedge (e_i + e_l)\right) - K\left(\frac{1}{2}(e_j - e_k) \wedge (e_i - e_l)\right) \\
&\quad - \frac{1}{2}K(e_i \wedge e_l) - \frac{1}{2}K(e_j \wedge e_k) + \frac{1}{2}K(e_j \wedge e_l) + \frac{1}{2}K(e_i \wedge e_k)\,. \qquad \square
\end{aligned}$$

Due to its symmetries, the curvature tensor of a Riemannian manifold has only one linearly independent trace, which is called the **Ricci tensor**.

We denote the Ricci tensor by Rc and use the convention

$$\mathrm{Rc}(u,v) \doteqdot \vartheta^i(\mathrm{Rm}(u,e_i)v) = \mathrm{tr}(w \mapsto \mathrm{Rm}(u,w)v)$$

or, in components,

$$\mathrm{Rc}_{ij} = \mathrm{Rm}_{kij}{}^{k}.$$

By Proposition 16.1 (iv), Rc is a symmetric tensor:

$$\mathrm{Rc}(u,v) = \mathrm{Rc}(v,u).$$

We can relate the corresponding quadratic form, $v \mapsto \mathrm{Rc}(v,v)$ to the sectional curvatures: Given a unit vector v, choose an orthonormal basis for TM with $e_n = v$. Then

$$\mathrm{Rc}(v,v) = \sum_{i=1}^{n-1} \mathrm{Rm}(v,e_i,v,e_i) = \sum_{i=1}^{n-1} K(v \wedge e_i).$$

Thus, the Ricci quadratic form in a direction v is $(n-1)$ times the average of the sectional curvatures in 2-planes containing v.

The trace of the Ricci curvature is called the **scalar curvature**, which we denote by R. That is,

$$\mathrm{R} = \mathrm{Rc}_i{}^{i} = g^{ij}\,\mathrm{Rc}_{ij}\ .$$

Thus, $\mathrm{R}(p)$ is $n(n-1)$ times the average of the sectional curvatures over all 2-planes in T_pM.

The full algebraic structure of the curvature tensor is elucidated by constructing a vector space on which it acts as a bilinear form. To each point p of M is associated the linear space $\Lambda^2 T_pM$ of **oriented two-planes**. This space is simply the quotient of the tensor product $T_pM \otimes T_pM$ by the relation

$$u \otimes v \sim -v \otimes u.$$

The corresponding vector bundle $\Lambda^2 TM$ is called the **two-plane bundle** of M.

We define the bilinear map $\wedge : T_pM \times T_pM \to \Lambda^2 T_pM$, called the **wedge product**, by

$$u \wedge v = [u \otimes v].$$

We can identify $u \wedge v \in \Lambda^2 T_pM$ with the two dimensional oriented plane in T_pM spanned by u and v (although not everything in $\Lambda^2 T_pM$ can be interpreted as a plane in T_pM). Thus, the construction of $\Lambda^2 T_pM$ allows us to perform formal sums and scalar multiplication of oriented two-planes. We extend the metric to $\Lambda^2 TM$ by declaring $\{\frac{1}{2} e_i \wedge e_j : 1 \le i < j \le n\}$ an orthonormal basis for $\Lambda^2 T_pM$ whenever $\{e_1, \ldots, e_n\}$ is an orthonormal basis for T_pM. By Exercise 16.3,

$$g(A,B) = A^{ij} B^{kl} (g_{ik} g_{jl} - g_{il} g_{jk}),$$

A two-plane which can be expressed in the form $u \wedge v$ for some vectors u and v is called a **simple** two-plane (and corresponds to a subspace of T_pM).

The significance of the two-plane bundle is that the curvature tensor defines a *symmetric* bilinear form on it.

Proposition 16.4. *The curvature tensor defines a symmetric bilinear form on the space of two-planes $\Lambda^2 T_pM$ via*

$$\mathrm{Rm}(A, B) = \mathrm{Rm}(A^{ij} e_i \wedge e_j, B^{kl} e_k \wedge e_l) \doteqdot A^{ij} A^{kl} \, \mathrm{Rm}_{ijkl} \, ,$$

where the sum is over all i and j with $i < j$ and all k and l with $k < l$.

Since the curvature operator is symmetric, it can be diagonalized. It is important to note that the eigenvalues of the curvature operator need not be sectional curvatures! The sectional curvatures are the values of the curvature operator on simple two-planes, but there is no reason why the eigenvectors of the curvature operator should be simple two-planes. In particular, it is possible to have all the sectional curvatures positive (or negative) at a point, while not having all of the eigenvalues of the curvature operator positive (negative).

In the special case of three dimensions, however, every two-plane is simple, and so the eigenvalues of the curvature operator are sectional curvatures.

16.1. The Taylor expansion of the metric. Recall that in exponential normal coordinates centered at p the metric tensor is Euclidean to first order (at p). We will show that the curvature tensor is an obstruction to obtaining local coordinates in which the metric tensor is Euclidean to second order at a given point.

Theorem 16.5. *Let (M, g) be a Riemannian manifold of dimension n. In exponential normal coordinates about a point $p \in M$, the components g_{ij} of g are given by*

$$g_{ij} = \delta_{ij} - \frac{1}{3} \mathrm{Rm}_{ikjl} \, x^k x^l + O(\|x\|^3) \, .$$

Proof. Fix $p \in M$ and an orthonormal basis $\{e_i\}_{i=1}^n$ for T_pM and let $(x^1, \ldots, x^n) : U \to \mathbb{R}^n$ be the corresponding exponential normal coordinate chart. Given $q = \exp_p x^i e_i \in U$, let u be the unit vector $u \doteqdot x^i e_i / |x|$ and consider a geodesic variation

$$\begin{aligned} \omega : (-t_0, t_0) \times (-\varepsilon_0, \varepsilon_0) &\to M \\ (t, \varepsilon) &\mapsto \exp_p(t\xi(\varepsilon)) \, , \end{aligned}$$

where $\xi : (-\varepsilon_0, \varepsilon_0) \to S^{n-1} \subset T_pM$ has initial data $\xi(0) = u$, $\xi'(0) = v \in T_u S^{n-1}$. In exponential normal coordinates,

$$\partial_\varepsilon \omega = (d \exp_p)_{t\xi} t\xi' = t(\partial_\varepsilon \xi^i) \partial_i$$

and hence
$$\omega^* g(\partial_\varepsilon, \partial_\varepsilon) = t^2 \partial_\varepsilon \xi^i \partial_\varepsilon \xi^j g_{ij}\,,$$
where we are writing g_{ij} for $g_{ij} \circ \omega$. Differentiating both sides k times with respect to t yields
$$\sum_{l=0}^{k} \binom{k}{l} \omega^* g(\nabla_t^{k-l}\partial_\varepsilon, \nabla_t^l \partial_\varepsilon) = \partial_\varepsilon \xi^i \partial_\varepsilon \xi^j \left(k(k-1)\partial_t^{k-2} g_{ij} + 2kt\partial_t^{k-1} g_{ij} + t^2 \partial_t^k g_{ij}\right)$$
In particular, at $(t, \varepsilon) = (0, 0)$,
$$\sum_{l=0}^{k} \binom{k}{l} \omega^* g(\nabla_t^{k-l}\partial_\varepsilon, \nabla_t^l \partial_\varepsilon) = k(k-1) v^i v^j \partial_t^{k-2} g_{ij}\,.$$

Note that, at $(0, 0)$,
$$\partial_\varepsilon \omega = 0\,, \quad \partial_t \omega = u\,, \quad \nabla_t \partial_t = 0 \quad \text{and} \quad \nabla_t \partial_\varepsilon \omega = v\,.$$
Differentiating Jacobi's equation, we also have
$$\begin{aligned} 0 &= \nabla_t^3 \partial_\varepsilon \omega + \nabla_t \operatorname{Rm}(\partial_t \omega, \partial_\varepsilon \omega)\partial_t \omega + \operatorname{Rm}(\partial_t \omega, \nabla_t \partial_\varepsilon \omega)\partial_t \omega \\ &= \nabla_t^3 \partial_\varepsilon \omega + \operatorname{Rm}(u, v)u \quad \text{at} \quad (0,0)\,. \end{aligned}$$
Thus, setting $k = 4$,
$$\begin{aligned} 12 v^i v^j \partial_t^2 g_{ij} &= 2\omega^* g(\nabla_t^4 \partial_\varepsilon, \partial_\varepsilon) + 8\omega^* g(\nabla_t^3 \partial_\varepsilon, \nabla_t \partial_\varepsilon) + 6\omega^* g(\nabla_t^2 \partial_\varepsilon, \nabla_t^2 \partial_\varepsilon) \\ &= \, - 8 \operatorname{Rm}(u, v, u, v) \quad \text{at} \quad (0,0)\,. \end{aligned}$$
By Proposition 14.10, $g_{ij} = \delta_{ij}$ and $\partial_t g_{ij} = 0$ at $(0, 0)$ and hence
$$v^i v^j g_{ij}(\exp_p(tu)) = v^i v^j \delta_{ij} - \frac{t^2}{3} \operatorname{Rm}(u, v, u, v) + O(t^3)\,.$$
Setting $t = |x|$ then yields
$$g_{ij}(q) = \delta_{ij} - \frac{1}{3} \operatorname{Rm}_{ikjl} x^k x^l + O(\|x\|^3)\,. \qquad \square$$

Recall that the covariant differential of a tensor T of type (k, ℓ) is the tensor ∇T of type $(k + 1, \ell)$ defined by
$$\nabla T(U, \cdot) \doteqdot \nabla_U T(\cdot)\,.$$
The **covariant Hessian** $\nabla^2 T$ of T is the covariant differential of ∇T. That is,
$$\begin{aligned} \nabla^2 T(U, V) &\doteqdot \nabla_U(\nabla T)(V) \\ &= \nabla_U(\nabla_V T) - \nabla_{\nabla_U V} T\,. \end{aligned}$$
It is also common to write $\nabla^2_{U,V} T$ and $\nabla_U \nabla_V T$ for $\nabla^2 T(U, V)$. So is important to keep in mind that $\nabla_U \nabla_V T \neq \nabla_U(\nabla_V T)$!

Observe that, for a symmetric connection,

$$\begin{aligned}
\operatorname{Rm}(U,V)T &= \nabla_V(\nabla_U T) - \nabla_U(\nabla_V T) - \nabla_{[V,U]}T \\
&= \nabla_V(\nabla_U T) - \nabla_U(\nabla_V T) - \nabla_{\nabla_V U - \nabla_U V}T \\
&= \nabla_V(\nabla_U T) - \nabla_{\nabla_V U} - (\nabla_U(\nabla_V T) - \nabla_{\nabla_U V}T) \\
&= \nabla_V\nabla_U T - \nabla_U\nabla_V T\,.
\end{aligned}$$

In particular, for a symmetric connection, the covariant Hessian of a smooth function f is symmetric,

$$\nabla_U\nabla_V f = \nabla_V\nabla_U f\,.$$

The trace of ∇^2 is called the **Laplace–Beltrami** operator, or simply the (**covariant**) **Laplacian**. That is,

$$\Delta T = \operatorname{tr}(\nabla^2 T) = g^{ij}\nabla_i\nabla_j T\,.$$

Lemma 16.6. *Let (M^n, g) be a Riemannian manifold equipped with its Levi-Civita connection ∇. Then*

$$\nabla\Delta f = \Delta\nabla f - \operatorname{Rc}(\operatorname{grad} f) \tag{16.1}$$

for any smooth function $f \in C(M)$, where $\operatorname{grad} f$ *is the vector field dual to the one-form df and $u \mapsto \operatorname{Rc}(u) \in T^*M$ is the linear map dual to* Rc*; that is,*

$$\operatorname{Rc}(u)(v) \doteqdot \operatorname{Rc}(u,v)\,.$$

Proof. Since $\nabla^2 f$ is symmetric,

$$\begin{aligned}
-\operatorname{Rm}(U,V,U,\operatorname{grad} f) &= (\operatorname{Rm}(U,V)\nabla f)(U) \\
&= \nabla^3 f(V,U,U) - \nabla^3 f(U,V,U) \\
&= \nabla^3 f(V,U,U) - \nabla^3 f(U,U,V)\,.
\end{aligned}$$

The claim follows by taking the trace. □

A slightly more involved argument shows that

$$\nabla\Delta T = \Delta\nabla T + \nabla R * T + R * \nabla T \tag{16.2}$$

for any tensor field T, where, given tensors S and T, we denote by $S * T$ any tensor obtained from the summing of constant multiples (depending only on the ranks of S and T) of contractions (possibly using the metric g or its inverse g^{-1}) of $S \otimes T$.

Exercises.

Exercise 16.1. *Show that the algebraic symmetries of Proposition 16.1 reduce the number of algebraically independent components* Rm_{ijkl} *of the Riemann tensor from* n^4 *to* $\frac{n^2(n^2-1)}{12}$. *In low dimensions, these numbers are quite manageable:* $0, 1, 6, 20, 50, \dots$.

Exercise 16.2. *Let* (M, g) *be a two dimensional Riemannian manifold. Given a point* $p \in M$ *and a basis* $\{v_1, v_2\}$ *for* T_pM, *show that the Gauss curvature of* (M, g) *at* p *is given by*

$$K(p) = \frac{\mathrm{Rm}(v_1, v_2, v_1, v_2)}{g(v_1, v_2, v_1, v_2)}\,,$$

where

$$g(u, v, w, z) \doteqdot g(u, w)g(v, z) - g(u, v)g(w, z)\,.$$

Exercise 16.3. *Show that the inner product of two two-planes* $A, B \in \Lambda^2 T_pM$ *is given by*

$$g(A, B) = A^{ij}B^{kl}(g_{ik}g_{jl} - g_{il}g_{jk})\,,$$

where the sums are over the indices $i < j$ *and* $k < l$.

Exercise 16.4. *Show that every two-plane on an* n*-manifold is simple if* $n = 2$ *or* $n = 3$ *but not if* $n \geq 4$.

17. Spaces of constant sectional curvature

Let (M, g) be a Riemannian manifold of dimension $n \geq 2$. We want to compute the various curvatures of (M, g) for some specific examples. We will first do this in the most naïve way: By computing the components of the Riemann curvature tensor with respect to some coordinate chart.

Given a tangent basis $\{e_i\}_{i=1}^n$, the components of the curvature tensor are given by

$$\begin{aligned}
\mathrm{Rm}(e_i, e_j)e_k &= \nabla_j(\nabla_i e_k) - \nabla_i(\nabla_j e_k) - \nabla_{[e_j,e_i]}e_k \\
&= \nabla_j(\Gamma_{ik}{}^p e_p) - \nabla_i(\Gamma_{jk}{}^p e_p) - C_{ji}{}^p \nabla_{e_p} e_k \\
&= e_j\Gamma_{ik}{}^p e_p + \Gamma_{ik}{}^p \nabla_j e_p - e_i\Gamma_{jk}{}^p e_p - \Gamma_{jk}{}^p \nabla_i e_p - C_{ji}{}^p \Gamma_{pk}{}^q e_q \\
&= e_j\Gamma_{ik}{}^p e_p + \Gamma_{ik}{}^p \Gamma_{jp}{}^q e_q - e_i\Gamma_{jk}{}^p e_p - \Gamma_{jk}{}^p \Gamma_{ip}{}^q e_q - C_{ji}{}^p \Gamma_{pk}{}^q e_q \\
&= \left(e_j\Gamma_{ik}{}^p - e_i\Gamma_{jk}{}^p + \Gamma_{ik}{}^q \Gamma_{jq}{}^p - \Gamma_{jk}{}^q \Gamma_{iq}{}^p - C_{ji}{}^q \Gamma_{qk}{}^p\right) e_p\,.
\end{aligned}$$

In particular, in a coordinate basis $\{\partial_i\}_{i=1}^n$,

$$\mathrm{Rm}_{ijk}{}^l = \partial_j\Gamma_{ik}{}^l - \partial_i\Gamma_{jk}{}^l + \Gamma_{ik}{}^q\Gamma_{jq}{}^l - \Gamma_{jk}{}^q\Gamma_{iq}{}^l\,. \tag{17.1}$$

We say that (M, g) has **constant sectional curvature** if the sectional curvature $K : \Lambda^2 TM \to \mathbb{R}$ is a constant function. By Proposition 16.4, the curvature tensor is then of the form

$$\mathrm{Rm}(u, v)w = K(g(u, w)v - g(v, w)u)$$

where K is a constant.

17.1. Euclidean space. Equip Euclidean space $\mathbb{R}^n$ with its standard metric $\langle\cdot,\cdot\rangle$ by declaring the coordinate basis of the identity chart orthonormal. By Theorem 14.9, the Levi-Civita connection D is given by

$$D_u V = (uV^i)\partial_i\,.$$

It then follows from (17.1) that $\mathbb{R}^n$ has constant sectional curvature $K \equiv 0$. A space with constant vanishing sectional curvature is called **locally flat**.

The geodesics of D are the straight lines $t \mapsto p + tv$ and hence $\mathbb{R}^n$ is geodesically complete.

The Jacobi equation is

$$\nabla_t^2 J = 0\,.$$

With respect to a parallel orthonormal frame $\{E_i\}_{i=1}^n$ along γ with $E_n = \gamma'$, this becomes

$$(J^i)'' = 0\,.$$

The solutions are

$$J^i(t) = A_i t + B_i\,.$$

17.2. The sphere. Consider the sphere $S^n_r \doteqdot \{z \in \mathbb{R}^{n+1} : |z| = r\}$ of radius $r > 0$ equipped with the pullback metric $g = \iota^* \langle \cdot, \cdot \rangle$, where $\iota : S^n_r \hookrightarrow \mathbb{R}^{n+1}$ is the inclusion map and $\langle \cdot, \cdot \rangle$ the standard metric on $\mathbb{R}^{n+1}$. Recall that **stereographic projection** $\Phi : S^n_r \setminus \{e_{n+1}\} \to \mathbb{R}^n \cong \mathbb{R}^n \times \{0\} \subset \mathbb{R}^{n+1}$ maps a point $y \in S^n_r$ to the point of intersection $\Phi(y)$ of the line joining y to e_{n+1} with the plane $\mathbb{R}^n \times \{0\}$.

Writing $y = (\hat{y}, y^{n+1})$, we find that

$$\Phi(y) = \frac{r\hat{y}}{r - y^{n+1}}$$

and

$$\Phi^{-1}(x) = \left(\hat{y}, y^{n+1}\right) ,$$

where

$$\hat{y} = \frac{2r^2 x}{r^2 + |x|^2} \quad \text{and} \quad y^{n+1} = \frac{|x|^2 - r^2}{|x|^2 + r^2} r\,.$$

Similar calculations as in Exercise 14.3 yield, with respect to the coordinate basis determined by Φ,

$$g_{ij} = \frac{4\delta_{ij}}{(1 + |x|^2/r^2)^2}.$$

The components of the Levi-Civita connection are then given by (14.2):

$$\Gamma_{ij}{}^k = g^{kl}\Gamma_{ijl} = \frac{-2}{|x|^2 + r^2}\left(x^i\delta_{jk} + x^j\delta_{ik} - x^k\delta_{ij}\right). \tag{17.2}$$

We can then compute, for instance,

$$\begin{aligned}\partial_k\Gamma_{ij}{}^l = {} & \frac{4}{(|x|^2 + r^2)^2}\left(x^k x^i\delta_{jl} + x^k x^j\delta_{il} - x^k x^l\delta_{ij}\right) \\ & - \frac{2}{|x|^2 + r^2}\left(\delta_{ik}\delta_{jl} + \delta_{jk}\delta_{il} - \delta_{kl}\delta_{ij}\right)\end{aligned}$$

and hence

$$\begin{aligned}
\mathrm{Rm}_{kij}{}^{l} &= \partial_i \Gamma_{kj}{}^{l} - \partial_k \Gamma_{ij}{}^{l} + \Gamma_{kj}{}^{p}\Gamma_{ip}{}^{l} - \Gamma_{ij}{}^{p}\Gamma_{kp}{}^{l} \\
&= \frac{4}{(|x|^2+r^2)^2}\left(x^i x^j \delta_{kl} - x^k x^j \delta_{il} - x^i x^l \delta_{kj} + x^k x^l \delta_{ij}\right) \\
&\quad + \frac{4}{|x|^2+r^2}\left(\delta_{kj}\delta_{il} - \delta_{ij}\delta_{kl}\right) \\
&\quad + \frac{4}{(|x|^2+r^2)^2}\left(x^k \delta_{jp} + x^j \delta_{kp} - x^p \delta_{kj}\right)\left(x^i \delta_{pl} + x^p \delta_{il} - x^l \delta_{ip}\right) \\
&\quad - \frac{4}{(|x|^2+r^2)^2}\left(x^i \delta_{jp} + x^j \delta_{ip} - x^p \delta_{ij}\right)\left(x^k \delta_{pl} + x^p \delta_{kl} - x^l \delta_{kp}\right) \\
&= \frac{4}{|x|^2+r^2}\left(\delta_{kj}\delta_{il} - \delta_{ij}\delta_{kl}\right) - \frac{4|x|^2}{(|x|^2+r^2)^2}\left(\delta_{kj}\delta_{il} - \delta_{ij}\delta_{kl}\right) \\
&= \frac{4r^{-2}}{(1+|x|^2/r^2)^2}\left(\delta_{kj}\delta_{il} - \delta_{ij}\delta_{kl}\right) .
\end{aligned}$$

Thus,

$$\mathrm{Rm}_{kijl} = g_{lp}\,\mathrm{Rm}_{kij}{}^{p} = \frac{1}{r^2}(g_{il}g_{jk} - g_{ij}g_{kl}) .$$

We conclude that the curvature operator of S^n_r is a constant multiple of the metric,

$$\mathrm{Rm}(A,B) = \frac{1}{r^2} g(A,B) ,$$

where we recall from Exercise 16.3 that the metric on two-planes is defined by

$$g(A,B) = A^{ij}B^{kl}(g_{ik}g_{jl} - g_{il}g_{jk}) .$$

In particular, S^n has constant sectional curvature:

$$K(A) \doteqdot \frac{\mathrm{Rm}(A,A)}{g(A,A)} \equiv \frac{1}{r^2} . \tag{17.3}$$

By (17.2), the geodesic equation is

$$\begin{aligned}
(\gamma^k)'' &= \frac{2}{|\gamma|^2+r^2}(\gamma^i)'(\gamma^j)'\left(\gamma^i \delta_{jk} + \gamma^j \delta_{ik} - \gamma^k \delta_{ij}\right) \\
&= \frac{2}{|\gamma|^2+r^2}\left(2(\gamma' \cdot \gamma)(\gamma^k)' - \gamma^k |\gamma'|^2\right) \\
&= \frac{2}{|\gamma|^2+r^2}\left((|\gamma|^2)'(\gamma^k)' - \gamma^k |\gamma'|^2\right) , \qquad (17.4)
\end{aligned}$$

where $\cdot$ and $|\cdot|$ are the standard dot product and norm on $\mathbb{R}^n$ in the coordinate variables.

We claim that the solutions are the great circles[17]. Indeed, given $y \in S^n$ and a unit tangent vector $v \in T_yS^n$, the great circle through y in the direction v is

$$\gamma(t) \doteqdot \cos(t/r)y + r\sin(t/r)v\,.$$

Observe that

$$|\gamma|_{\mathbb{R}^{n+1}} \equiv r$$

and

$$(\gamma(0), \gamma'(0)) = (y, v)\,.$$

So, by uniqueness of solutions, it suffices to show that each γ is a solution. The image of γ under Φ is

$$\Phi(\gamma) = \frac{r(\cos(t/r)\hat{y} + r\sin(t/r)\hat{v})}{r - (\cos(t/r)y^{n+1} + r\sin(t/r)v^{n+1})}\,.$$

The situation is simplified when y is the south pole, $-re_{n+1}$, for in that case $\hat{y} = v^{n+1} = 0$, and hence

$$\Phi(\gamma) = \frac{r\sin(t/r)}{1+\cos(t/r)}v\,.$$

If $v = e_1$, say, then $\gamma^k = 0$ for $k = 2, \ldots, n$ (which clearly solve (17.4)), while

$$\gamma^1 = \frac{r\sin(t/r)}{1+\cos(t/r)}\,.$$

We leave it as an exercise to check that γ^1 also solves (17.4). The general case follows, since rotations are isometries of S^n (another exercise) and for every $y \in S^n_r$ and every $v \perp y$, there is a rotation of $\mathbb{R}^{n+1}$ which maps the pair $(-e_{n+1}, e_1)$ to (y, v).

By (17.3), the Jacobi equation becomes

$$\nabla_t^2 J + r^{-2}J^{\perp} = 0\,,$$

where $J^{\perp}$ is the projection of J onto $\{\gamma'\}^{\perp}$. With respect to a parallel orthonormal frame $\{E_i\}_{i=1}^n$ along γ with $E_n = \gamma'$, this becomes

$$(J^i)'' + r^{-2}J^i = 0$$

for $i = 1, \ldots, n-1$, and

$$(J^n)'' = 0\,.$$

The solutions are

$$J^i(t) = \begin{cases} At + B & \text{if } i = n \\ A_i\cos(t/r) + B_i\sin(t/r) & \text{if } i < n\,. \end{cases}$$

[17] Indeed, we have already seen this in Example 11.8 by different means.

17.3. Hyperbolic space. By similar calculations using the hyperbolic stereographic projection, we find that the hyperbolic space $H^n_r \doteqdot \{z \in \mathbb{R}^{n,1} : b(z,z) = -r^2\}$ of radius $r > 0$ (equipped with its pullback metric) has constant sectional curvature $K \equiv -r^{-2}$.

The Jacobi equation then becomes

$$\nabla_t^2 J - r^{-2} J^\perp = 0\,,$$

where $J^\perp$ is the projection of J onto $\{\gamma'\}^\perp$. With respect to a parallel orthonormal frame $\{E_i\}_{i=1}^n$ along γ with $E_n = \gamma'$, this becomes

$$(J^i)'' - r^{-2} J^i = 0$$

for $i = 1, \ldots, n-1$, and

$$(J^n)'' = 0\,.$$

The solutions are

$$J^i(t) = \begin{cases} At + B & \text{if } i = n \\ A_i \cosh(t/r) + B_i \sinh(t/r) & \text{if } i < n\,. \end{cases}$$

Exercises.

Exercise 17.1. *Given $r > 0$, set*

$$f(t) \doteqdot \frac{r \sin(t/r)}{1 - \cos(t/r)}\,.$$

Show that

$$f'' + \frac{2f(f')^2}{f^2 + r^2} = 0\,.$$

Exercise 17.2. *Denote by $\Phi : S^n \setminus \{re_{n+1}\} \to \mathbb{R}^n$ the stereographic projection map and consider the geodesic*

$$\gamma(t) \doteqdot -r\cos(t/r)e_{n+1} + r\sin(t/r)e_1\,.$$

(a) *Show that*

$$\frac{1}{1+\cos(t/r)} \partial_i \Phi^{-1}|_{\gamma(t)} = \begin{cases} \cos(t/r)e_1 + \sin(t/r)e_{n+1} & \textit{if } i = 1 \\ e_i & \textit{if } i \neq 1\,. \end{cases}$$

(b) *Show that the general solution to the system*

$$\frac{dE^k}{dt} = -E^i \frac{d\gamma^j}{dt} (\Gamma_{ij}{}^k \circ \gamma)$$

is given by

$$E^k(t) = \frac{2E^k(0)}{1+\cos(t/r)}\,.$$

(c) *Deduce that the parallel orthonormal frame* $\{E_i(t)\}_{i=1}^n$ *for* $T_{\gamma(t)}S^n \hookrightarrow T_{\gamma(t)}\mathbb{R}^{n+1}$ *along* γ *with initial conditions* $E_i(0) = e_i$ *for* $i = 1, \ldots, n$ *is given by*

$$E_i(t) \doteqdot \begin{cases} \cos(t/r)e_1 + \sin(t/r)e_{n+1} & \text{if } i = 1 \\ e_i & \text{if } i \neq 1 \,. \end{cases}$$

Exercise 17.3. *Two Riemannian metrics* g *and* $\overline{g}$ *on a differentiable manifold* M *are* ***conformally equivalent*** *if there exists a smooth function* u *on* M *such that*

$$g = \mathrm{e}^{2u}\overline{g}\,.$$

Show that the scalar curvature R *with respect to* g *is given by*

$$\mathrm{e}^{-2u}\,\mathrm{R} = \overline{\mathrm{R}} - 2(n-1)\Delta u - (n-1)(n-2)|\operatorname{grad} u|^2\,,$$

where $\overline{\mathrm{R}}$ *is the scalar curvature of* $\overline{g}$ *and* n *is the dimension of* M.

18. Riemannian submanifolds

Let M be an n-dimensional smooth manifold, N an $(n+k)$-dimensional smooth manifold and $X : M \to N$ a smooth **immersion**; that is, a smooth map whose derivative $dX : TM \to TN$ is everywhere injective. By the implicit function theorem, X is locally a diffeomorphism onto its image: about any point $p \in M^n$ we can find a neighbourhood $U \subset M^n$ such that $\mathcal{U} \doteqdot X(U)$ is a smooth n-dimensional submanifold of N and $X|_U : U \to \mathcal{U}$ is a diffeomorphism. For an embedded submanifold $\mathcal{M} \subset N$, we can take X to be the inclusion map, in which case its derivative dX_p is the inclusion of $T_p\mathcal{M}$ into T_pN for each p. In what follows, it is useful to keep this case in mind.

Recall that the pullback bundle X^*TN is the vector bundle obtained by equipping the set

$$X^*TN \doteqdot \{(p,v) \in M \times TN : X(p) = \pi(v)\}$$

with the submanifold differentiable structure and the the obvious projection $(p,v) \mapsto p$. Define the **pullback** X^*V of $V \in \Gamma(TN)$ by

$$X^*V_p \doteqdot (p, V_{X(p)}).$$

Pullbacks of sections V of TN are sections of X^*TN; however, in general, not every section of X^*TN is of this form. For example, given $V \in \Gamma(TM)$ the vector field $dX(V)$ defined by

$$dX(V)_p \doteqdot (X(p), dX_p(V_p))$$

is not, in general, globally of the form $X^*\tilde{V}$ for some $\tilde{V} \in \Gamma(TN)$ (consider what happens at points where X has self-intersections). On the other hand, when X is an embedding, the map

$$(p,v) \mapsto v$$

is a bundle isomorphism from X^*TN to TN, and it follows that every section is a pulled-back section in this case. Moreover, fo general immersions X, every section of X^*TN coincides *locally* with a pulled-back section of TN.

When X is an embedding, the map $(p,u) \mapsto (X(p), dX_p(u))$ is an embedding of TM into TN. More generally, dX induces an embedding of TM into X^*TN via

$$(p,u) \mapsto (p, dX(u))\,.$$

Abusing notation, we denote this vector subbundle of X^*TN (equipped with the obvious projection) by $dX(TM)$ and conflate the embedding map with dX.

If N is equipped with a (pseudo-)Riemannian metric $\langle\cdot,\cdot\rangle$, then we define the pullback metric $X^*\langle\cdot,\cdot\rangle$ on X^*TN by

$$X^*\langle(p,u),(p,v)\rangle \doteqdot \langle u,v\rangle\,.$$

The **normal space** to M at $p\in M$ is the orthogonal compliment

$$N_pM \doteqdot \{\nu\in T_{X(p)}N : \langle u,\nu\rangle = 0 \text{ for all } u\in dX_p(T_pM)\}$$

of $dX_p(T_pM)$ in $T_{X(p)}N$. When X is an embedding, the **normal bundle**

$$NM \doteqdot \bigsqcup_{p\in M} N_pM\,,$$

equipped with the obvious projection, is a vector subbundle of TN, and

$$TN = dX(TM)\oplus NM\,.$$

More generally, we define the **normal bundle** by equipping

$$NM \doteqdot \{\nu\in X^*TN : {}^X\langle u,\nu\rangle = 0 \text{ for all } u\in dX(TM)_{\pi(\nu)}\}$$

with the submanifold structure[18] and the induced projection. Then

$$X^*TN = dX(TM)\oplus NM\,.$$

We can also equip TM with a (pseudo-)Riemannian metric g: given $u,v\in T_pM$, set

$$g_p(u,v)\doteqdot \langle dX_p(u),dX_p(v)\rangle_{X(p)}\,.$$

If N is equipped with a connection D, then the **pullback connection** ${}^XD : TM\times\Gamma(X^*TN)\to X^*TN$ defines a connection on X^*TN via

$${}^XD_uX^*V \doteqdot (\pi(u), D_{dX(u)}V)$$

for any $u\in TM$ and $V\in\Gamma(TN)$. Note that ${}^XD$ is well-defined since every section of X^*TN coincides locally with a pulled-back section of TN.

If $V\in\Gamma(TM)$, then we define ∇_uV by

$$dX(\nabla_uV)\doteqdot\left({}^XD_uX^*\tilde V\right)^\top = \left(D_{dX(u)}\tilde V\right)^\top,$$

where $\tilde V$ is any vector field on N which is locally X-related to V near $\pi(u)$; that is, $\tilde V_{X(p)} = dX_p(V_p)$ for p in a neighbourhood of $\pi(u)$. We can also define a connection $\nabla^\perp : TM\times\Gamma(NM)\to NM$ on the normal bundle NM via

$$\nabla^\perp_uN \doteqdot ({}^XD_uN)^\perp.$$

It is a straightforward exercise to check that ∇ and $\nabla^\perp$ define connections on TM and NM, respectively. Indeed, if D is the Levi-Civita connection of $\langle\cdot,\cdot\rangle$, then ∇ is the Levi-Civita connection of g.

[18] That NM is a submanifold follows from the implicit function theorem (surjective version).

Proposition 18.1. *Let $(N, \langle\cdot,\cdot\rangle)$ be a Riemannian manifold equipped with its Levi-Civita connection and let $X : M \to N$ be an immersion. If g and ∇ are the induced metric and connection, then ∇ is the Levi-Civita connection of g.*

Proof. We need to check that ∇ is symmetric and g-compatible. Given vector fields U, V, W on M with $U_p = u$, $V_p = v$ and $W_p = w$, choose vector fields $\tilde{U}$, $\tilde{V}$ and $\tilde{W}$ on N such that $dX(U) = \tilde{U}$, $dX(V) = \tilde{V}$ and $dX(W) = \tilde{W}$ on a neighbourhood of p. Then, by Proposition 7.6,

$$\begin{aligned} dX_p(\nabla_u V - \nabla_v U) &= (D_{\tilde{U}}\tilde{V} - D_{\tilde{V}}\tilde{U})|_{X(p)} \\ &= [\tilde{U}, \tilde{V}]_{X(p)} \\ &= dX_p[U, V]_p\,. \end{aligned}$$

Similarly,

$$\begin{aligned} ug(V, W) &= u\left(\left\langle \tilde{V}, \tilde{W}\right\rangle \circ X\right) \\ &= dX(u)\left\langle \tilde{V}, \tilde{W}\right\rangle \\ &= \left\langle D_{dX(u)}\tilde{V}, \tilde{W}_{X(p)}\right\rangle + \left\langle \tilde{V}_{X(p)}, D_{dX(u)}\tilde{W}\right\rangle \\ &= g(\nabla_u V, w) + g(v, \nabla_u W)\,. \end{aligned}$$ □

The normal (tangential) component of ${}^X D$ acting on tangent (normal) vector fields induces a normal bundle valued symmetric bilinear form on TM (a conormal bundle valued self-adjoint endomorphism of TM).

Proposition 18.2. *There are tensor-fields*

$$\mathrm{A} \in \Gamma(T^*M \otimes T^*M \otimes NM) \cong \mathrm{Hom}(\Gamma(TM), \Gamma(TM); \Gamma(NM))$$

and

$$\mathrm{W} \in \Gamma(T^*M \otimes TM \otimes N^*M) \cong \mathrm{Hom}(\Gamma(NM), \Gamma(TM); \Gamma(TM))$$

such that

(1) $\mathrm{A}(u, v) = \left({}^X D_u\, [dX(V)]\,\right)^\perp$,
(2) $dX\big(\mathrm{W}^\nu(u)\big) = ({}^X D_u N)^\top$,
(3) $\mathrm{A}(u, v) = \mathrm{A}(v, u)$, *and*
(4) $\mathrm{A}^\nu(u, v) = g(\mathrm{W}^\nu(u), v)$

for any $u, v \in T_pM$ and $\nu \in N_pM$, any extension $V \in \Gamma(TM)$ of v, and any $N \in \Gamma(NM)$ such that $N_p = \nu$, where

$$\mathrm{A}^\nu(u, v) \doteqdot -\left\langle \mathrm{A}(u, v), \nu\right\rangle\,.$$

Proof. To prove the first claim, it suffices to check that the map

$$V \mapsto \big({}^X D_u\,[dX(V)]\,\big)^{\perp}$$

is $C(M)$-linear. Given $V \in \Gamma(TM)$ and $f \in C(M)$,

$$\begin{aligned}
\big({}^X D_u\,[dX(fV)]\,\big)^{\perp} &= \big({}^X D_u\,[f\,dX(V)]\,\big)^{\perp}\\
&= \big(uf dX_p(V_p) + f\,{}^X D_u\,[dX(V)]\,\big)^{\perp}\\
&= f\big({}^X D_u\,[dX(V)]\,\big)^{\perp}
\end{aligned}$$

since $dX(V)$ is tangential.

The second claim is proved similarly (see Exercise 18.1).

The third claim follows from the symmetry of D similarly as in Proposition 18.1: given $u, v \in T_pM$, extensions $U, V \in \Gamma(TM)$, and vector fields $\tilde U, \tilde V \in \Gamma(TN)$ X-related to U and V, respectively, near p, Proposition 7.6 yields

$$\begin{aligned}
{}^X D_u\big[dX(V)\big] - {}^X D_v\big[dX(U)\big] &= D_{\tilde U_{X(p)}}\tilde V - D_{\tilde V_{X(p)}}\tilde U\\
&= [\tilde U, \tilde V]_{X(p)}\\
&= dX_p([U,V]_p)\,.
\end{aligned}$$

Taking the normal projection yields the claim.

To obtain the final claim, choose a vector field $\tilde N \in \Gamma(TN)$ which is locally X-related to N in a neighbourhood of p. Then, since $\langle \tilde V, \tilde N\rangle \equiv 0$,

$$\begin{aligned}
0 &\equiv u\Big(\big\langle \tilde V, \tilde N\big\rangle \circ X\Big)\\
&= dX_p(u)\big\langle \tilde V, \tilde N\big\rangle\\
&= \Big\langle D_{dX_p(u)}\tilde V, \nu\Big\rangle + \Big\langle dX_p(V), D_{dX_p(u)}\tilde N\Big\rangle\\
&= \Big\langle (D_{dX_p(u)}\tilde V)^{\perp}, \nu\Big\rangle + \Big\langle dX_p(V), (D_{dX_p(u)}\tilde N)^{\top}\Big\rangle\\
&= -\,\mathrm{A}^{\nu}_p(u,v) + g(\mathrm{W}^{\nu}_p(u), v)\,.
\end{aligned}$$

□

The tensor field A in Lemma 18.2 is called the **second fundamental form** of the immersion. The tensor field W is called the **Weingarten tensor**.

Given $u \in TM$ and $V \in \Gamma(TM)$, resolving ${}^X D_u\tilde V$ into tangential and normal components yields the (first) **Weingarten equation**:

$${}^X D_u\tilde V = dX(\nabla_u V) + \mathrm{A}(u,v)\,.$$

Playing the same game with second derivatives (the curvature) yields the Gauss–Codazzi equations: given $u, v, w, z \in T_pM$, choose extension vector fields $U, V, W, Z \in \Gamma(TM)$, respectively, and vector fields $\tilde U, \tilde V, \tilde W, \tilde Z \in$

$\Gamma(TN)$ which are X-related to U, V, W, Z, respectively, in a neighbourhood of p. Then

$$D_{\tilde{U}}(D_{\tilde{V}}\tilde{W}) = D_{\tilde{U}}\left(\widetilde{\nabla_V W} + \mathrm{A}(V,W)\right)$$

where $\widetilde{\nabla_V W} \doteqdot (D_{\tilde{V}}\tilde{W})^\top$ (which is X-related to $\nabla_V W$ in a neighbourhood of p). Since $\mathrm{A}(V,W)$ is a normal vector field, we obtain

$$\begin{aligned} D_{dX(u)}(D_{\tilde{V}}\tilde{W}) &= D_{dX(u)}\left(\widetilde{\nabla_V W} + \mathrm{A}(V,W)\right) \\ &= dX\left(\nabla_u\big(\widetilde{\nabla_V W}\big)\right) + \mathrm{A}(u, \nabla_v W) \\ &\quad + \mathrm{W}^{\mathrm{A}(v,w)}(u) + \nabla^\perp_u\left(\mathrm{A}(V,W)\right). \end{aligned}$$

The tangential terms yield the **Gauss equation**:

$$\big({}^X\mathrm{Rm}(u,v)w\big)^\top = dX\left(\mathrm{Rm}^\top(u,v)w + \mathrm{W}^{\mathrm{A}(u,w)}(v) - \mathrm{W}^{\mathrm{A}(v,w)}(w)\right),$$

where ${}^X\,\mathrm{Rm}$ is the pullback of the curvature operator of ${}^XD$ and $\mathrm{Rm}^\top$ is the curvature operator of ∇. Equivalently, using Proposition 18.2,

$${}^X\mathrm{Rm}(u,v,w,z) = \mathrm{Rm}^\top(u,v,w,z) + g\big(\mathrm{A}(u,z),\mathrm{A}(v,w)\big) - g\big(\mathrm{A}(u,w),\mathrm{A}(v,z)\big).$$

The normal terms yield the **Codazzi–Mainardi equation**:

$$\big({}^X\mathrm{Rm}(u,v)w\big)^\perp = (\nabla_v\,\mathrm{A})(u,w) - (\nabla_u\,\mathrm{A})(v,w)\,,$$

where the tensor $\nabla\mathrm{A}$ is defined, as usual, by asserting the Leibniz rule:

$$\nabla_u\,\mathrm{A}(v,w) \doteqdot \nabla^\perp_u(\mathrm{A}(V,W)) - \mathrm{A}(\nabla_u V, w) - \mathrm{A}(v, \nabla_u W)\,.$$

Now consider a normal vector $\nu \in N_pM$, an extension $N \in \Gamma(NM)$ and a vector field $\tilde{N} \in \Gamma(TN)$ with $X^*\tilde{N} = N$ in a neighbourhood of p. Resolving tangential and normal components of ${}^XD_uN$ yields the (second) **Weingarten equation**:

$${}^XD_vN = \nabla^\perp_v N + dX(\mathrm{W}^\nu(v))\,.$$

The second derivative is

$$\begin{aligned} D_{\tilde{U}}(D_{\tilde{V}}\tilde{N}) &= \nabla^\perp_{\tilde{U}}\big(\nabla^\perp_{\tilde{V}}N\big) + dX\big(\mathrm{W}^{\nabla^\perp_V N}(U)\big) \\ &\quad + dX\big(\nabla_U(\mathrm{W}^N(V))\big) + \mathrm{A}(U, \mathrm{W}^N(V))\,. \end{aligned} \tag{18.1}$$

Resolving the normal components yields the **Ricci equation**

$$\big({}^X\mathrm{Rm}(u,v)\nu\big)^\perp = \mathrm{Rm}^\perp(u,v)\nu + \mathrm{A}(v, \mathrm{W}^\nu(u)) - \mathrm{A}(u, \mathrm{W}^\nu(v))\,,$$

where $\mathrm{Rm}^\perp$ is the curvature operator of $\nabla^\perp$. Equivalently,

$$\begin{aligned} {}^X\mathrm{Rm}(u,v,\nu,\mu) &= \mathrm{Rm}^\perp(u,v,\nu,\mu) + \mathrm{A}^\mu(u, \mathrm{W}^\nu(v)) - \mathrm{A}^\mu(v, \mathrm{W}^\nu(u)) \\ &= \mathrm{Rm}^\perp(u,v,\nu,\mu) + g(\mathrm{W}^\mu(u), \mathrm{W}^\nu(v)) - g(\mathrm{W}^\mu(v), \mathrm{W}^\nu(u))\,. \end{aligned}$$

It turns out that the tangential components just produce the Codazzi equation again.

18.1. Exercises.

Exercise 18.1. *Show that the map*

$$N \mapsto \mathrm{W}^N(u) \doteqdot ({}^X D_u N)^\top$$

for $N \in \Gamma(NM)$, $u \in T_pM$ is $C(M)$-linear.

Exercise 18.2. *Prove the Codazzi identity*

$$\left({}^X\mathrm{Rm}(u,v)\nu\right)^\top = (\nabla_v \mathrm{W})^\nu(u) - (\nabla_u \mathrm{W})^\nu(v)$$

by resolving the tangential components of (18.1)*, where*

$$(\nabla_u \mathrm{W})^\nu(v) \doteqdot \nabla_u(\mathrm{W}^N(V)) - \mathrm{W}^{\nabla^\perp_u N}(v) - \mathrm{W}^\nu(\nabla_u V)\,.$$

Exercise 18.3. *Assume that $X : M^n \to (N^{n+1}, \langle\cdot,\cdot\rangle)$ admits a global unit normal field N. Prove that*

$${}^X\mathrm{Rm}(u,v,w,N_p) = (\nabla_u A)(v,w) - (\nabla_v A)(u,w)\,,$$

where

$$A(u,v) \doteqdot -\langle \mathrm{A}(u,v), N_p\rangle$$

Exercise 18.4. *Use the Gauss equation to compute the curvature tensor of*

(1) *the sphere,*

$$S^n \doteqdot \{x \in \mathbb{R}^{n+1} : \langle x,x\rangle = 1\}\,.$$

(2) *the hyperbolic space,*

$$H^n \doteqdot \{x \in \mathbb{R}^{n,1} : \eta(x,x) = -1,\ \eta(x,e_0) < 0\}\,.$$

(3) *the Clifford torus,*

$$\mathrm{Cliff}^n \doteqdot \{(z_1,\ldots,z_n) \in S^{2n-1} \subset \mathbb{C}^n : |z_1|^2 = \cdots = |z_n|^2 = \tfrac{1}{n}\}\,.$$

19. First and second variations of arc-length

In this section, we study the variational properties of geodesics.

Definition 19.1. *Let (M,g) be a Riemannian manifold and $\gamma : [a,b] \to M$ a continuous curve. A **variation of** γ is a continuous map $\omega : [a,b] \times (-\varepsilon_0, \varepsilon_0) \to M$ such that $\omega(t,0) = \gamma(t)$ for all $t \in [a,b]$. We say that ω*

- *is **piecewise** C^k if there are points $a \doteqdot a_0 < a_1 < \cdots < a_l \doteqdot b$ such that $\omega_{[a_{i-1},a_i]\times(-\varepsilon_0,\varepsilon_0)}$ is C^k for each $i = 1, \ldots, l$. We refer to the points $\{a_1, \ldots, a_{l-1}\}$ as the **singular points** of ω.*
- *is a **homotopy of** γ or it **has fixed endpoints** if $\omega(a,\varepsilon) \equiv \gamma(a)$ and $\omega(b,\varepsilon) \equiv \gamma(b)$ for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$.*
- *is a **geodesic variation** if $t \mapsto \omega(t,\varepsilon)$ is a geodesic for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$.*

Lemma 19.2 (First variation of arc-length)**.** *Let (M,g) be a Riemannian manifold and $\gamma : [a,b] \to M$ a C^2 curve parametrized by arc-length. Consider a C^2 variation $\omega : [a,b] \times (-\varepsilon_0, \varepsilon_0) \to M$ of γ and set $J \doteqdot \partial_\varepsilon \omega|_{\varepsilon=0}$. Then*

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} L(\omega(\cdot,\varepsilon)) = g(J,\gamma')\big|_a^b - \int_a^b g(\nabla_t \gamma', J)\, dt\,. \tag{19.1}$$

Proof.

$$\begin{aligned}
\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} L(\omega(\cdot,\varepsilon)) &= \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \int_a^b |\partial_t \omega|\, dt \\
&= \int_a^b g\left(\frac{\partial_t\omega}{|\partial_t\omega|}, \nabla_\varepsilon \partial_t \omega\right)\, dt\Big|_{\varepsilon=0} \\
&= \int_a^b g\left(\frac{\partial_t\omega}{|\partial_t\omega|}, \nabla_t \partial_\varepsilon \omega\right)\, dt\Big|_{\varepsilon=0} \qquad (19.2)\\
&= \left[g\left(\frac{\partial_t\omega}{|\partial_t\omega|}, \partial_\varepsilon\omega\right)\Big|_a^b - \int_a^b g\left(\nabla_t \frac{\partial_t\omega}{|\partial_t\omega|}, \partial_\varepsilon\omega\right)\, dt\right]\Big|_{\varepsilon=0} \\
&= g(J,\gamma')\big|_a^b - \int_a^b g(\nabla_t\gamma', J)\, dt\,.
\end{aligned}$$

□

Corollary 19.3. *A piecewise C^2 path $\gamma : [a,b] \to M$ parametrized by arc-length is geodesic if and only if*

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} L(\omega(\cdot,\varepsilon)) = 0$$

for every piecewise C^2 homotopy ω.

Proof. Let $\omega : [a,b]\times(-\varepsilon_0,\varepsilon_0)\to M$ be a piecewise C^2 variation of γ with singular points $a=t_0<t_1<\cdots<t_{k+1}=b$. Then, by (19.1),

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0}L(\omega(\cdot,\varepsilon))=\sum_{i=1}^{k}g(J,\Delta\gamma')\big|_{t_{i-1}}^{t_i}-\int_a^b g(\nabla_t\gamma',J)\,dt\,, \tag{19.3}$$

where

$$\Delta\gamma'(t)\doteqdot\lim_{s\searrow t}\gamma'(s)-\lim_{s\nearrow t}\gamma'(s)\,.$$

The right hand side vanishes when is γ is a geodesic.

To prove the converse, suppose that the first variation vanishes for all piecewise C^2 variations ω of γ. In particular, given any any piecewise C^1 vector field J along γ with $J(a)=J(b)=0$, set

$$\omega(t,\varepsilon)\doteqdot\exp\varepsilon J(t) \tag{19.4}$$

for $(t,\varepsilon)\in[a,b]\times(-\varepsilon_0,\varepsilon_0)$ with ε_0 sufficiently small that $\pm\varepsilon_0 J(t)\in\mathcal{T}M$ for all $t\in[a,b]$. Then

$$0=\frac{d}{d\varepsilon}\Big|_{\varepsilon=0}L(\omega(\cdot,\varepsilon))=\sum_{i=1}^{k}g(J,\Delta\gamma')\big|_{t_{i-1}}^{t_i}-\int_a^b g(\nabla_t\gamma',J)\,dt\,.$$

Let $a=t_1<\cdots<t_k=b$ be the discontinuities of ω' and suppose there exists $t_0\in[a,b]\setminus\{t_1,\dots,t_k\}$ such that $U_0\doteqdot\nabla_t\omega'(t_0)\neq0$. Extend U_0 to a parallel vector field $U:(t_0-\delta,t_0+\delta)\to M$ such that $(t_0-\delta,t_0+\delta)\subset[a,b]\setminus\{t_1,\dots,t_k\}$. By continuity, we can also assume that $\langle U,\nabla_t\omega'\rangle>0$ so long as δ is sufficiently small. Now pick a cut-off function $\varphi\in C^\infty((t_0-\delta,t_0+\delta))$ with $\varphi(t_0)=1$ and $\varphi|_{(t_0-\delta,t_0+\delta)\setminus(t_0-\delta/2,t_0+\delta/2)}\equiv0$. But then, by (19.3) the variation $\omega^\varphi:[a,b]\times(-\varepsilon_0,\varepsilon_0)\to M$ of γ given by

$$\omega^\varphi(t,\varepsilon)\doteqdot\exp\varepsilon\varphi(t)U(t)$$

has negative first variation, a contradiction. We conclude that $\nabla_t\omega'=0$. It then follows from (19.3) (which must hold for any piecewise continuous vector field J by the formula (19.4)) that γ' is continuous and hence, by Theorem 7.1, γ is C^∞. This completes the proof. $\square$

So geodesics are critical points of the length functional. Note, however, that they are not always (local) minima (consider the equator on a sphere: any small perturbation in the direction of due north decreases its length).

The stability properties of geodesics can be analysed via the *second* variation of arc-length.

Lemma 19.4 (Second variation of arc-length)**.** *Let $\gamma:[a,b]\to M$ be an arc-length parametrized C^2 curve in a Riemannian manifold (M,g). For* *any C^3 variation $\omega : [a,b] \times (-\varepsilon_0, \varepsilon_0) \to M$ of γ,*

$$\frac{d^2}{d\varepsilon^2}\Big|_{\varepsilon=0} L(\omega(\cdot,\varepsilon)) = g(\nabla_\varepsilon \partial_\varepsilon \omega|_{\varepsilon=0}, \gamma')\big|_a^b + \int_a^b \Big[|\nabla_t J|^2 - R(\gamma', J, \gamma', J) \\ - g(\nabla_t J, \gamma')^2 - g(\nabla_t \gamma', \nabla_\varepsilon \partial_\varepsilon \omega|_{\varepsilon=0})\Big] dt\,,$$

where $J \doteqdot \partial_\varepsilon \omega|_{\varepsilon=0}$.

Proof. We continue from (19.2) in the computation of the first variation:

$$\begin{aligned}
&\frac{d^2}{d\varepsilon^2}\Big|_{\varepsilon=0} L(\omega(\cdot,\varepsilon)) \\
&\quad = \frac{d}{d\varepsilon} \int_a^b g\left(\frac{\partial_t \omega}{|\partial_t \omega|}, \nabla_t \partial_\varepsilon \omega\right) dt \,\Big|_{\varepsilon=0} \\
&\quad = \int_a^b \left[g\left(\nabla_\varepsilon \frac{\partial_t \omega}{|\partial_t \omega|}, \nabla_t \partial_\varepsilon \omega\right) + g\left(\frac{\partial_t \omega}{|\partial_t \omega|}, \nabla_\varepsilon (\nabla_t \partial_\varepsilon \omega)\right)\right] dt \,\Big|_{\varepsilon=0} \\
&\quad = \int_a^b \left[g\left(\frac{\nabla_t \partial_\varepsilon \omega}{|\partial_t \omega|} - g\left(\frac{\partial_t \omega}{|\partial_t \omega|}, \nabla_t \partial_\varepsilon \omega\right) \frac{\partial_t \omega}{|\partial_t \omega|^2}, \nabla_t \partial_\varepsilon \omega\right) \right. \\
&\qquad \left. +\ g\left(\frac{\partial_t \omega}{|\partial_t \omega|}, \nabla_t (\nabla_\varepsilon \partial_\varepsilon \omega) + R(\partial_t \omega, \partial_\varepsilon \omega) \partial_\varepsilon \omega\right)\right] dt \,\Big|_{\varepsilon=0} \\
&\quad = g(\nabla_\varepsilon \partial_\varepsilon \omega|_{\varepsilon=0}, \gamma')\big|_a^b + \int_a^b \Big[|\nabla_t J|^2 - R(\gamma', J, \gamma', J) \\
&\qquad\qquad - g(\nabla_t J, \gamma')^2 - g(\nabla_t \gamma', \nabla_\varepsilon \partial_\varepsilon \omega|_{\varepsilon=0})\Big] dt\,. \qquad \square
\end{aligned}$$

Setting

$$J^\perp \doteqdot J - g(J, \gamma')\gamma'\,,$$

we find that

$$\nabla_t J^\perp \doteqdot \nabla_t J - g(\nabla_t J, \gamma')\gamma' = (\nabla_t J)^\perp$$

and hence

$$|\nabla_t J^\perp|^2 = |\nabla_t J|^2 - g(\nabla_t J, \gamma')^2\,.$$

Corollary 19.5. *If $\gamma : [a,b] \to M$ is an arc-length parametrised geodesic in (M,g) and $\omega : [a,b] \times (-\varepsilon, \varepsilon) \to M$ any smooth variation of γ, then*

$$\frac{d^2}{d\varepsilon^2}\Big|_{\varepsilon=0} L(\omega(\cdot,\varepsilon)) = g(\nabla_\varepsilon \partial_\varepsilon \omega|_{\varepsilon=0}, \gamma')\big|_a^b \\ + \int_a^b \big[|\nabla_t J^\perp|^2 - R(\gamma', J^\perp, \gamma', J^\perp)\big] dt\,.$$

In particular, if ω is a homotopy, then

$$\frac{d^2}{d\varepsilon^2}\Big|_{\varepsilon=0} L(\omega(\cdot,\varepsilon)) = \int_a^b \big[|\nabla_t J^\perp|^2 - R(\gamma', J^\perp, \gamma', J^\perp)\big] dt\,.$$

Definition 19.6. *Given a geodesic $\gamma : [a, b] \to M$, denote by Υ^k the linear space of piecewise C^k vector fields along γ and by Υ^k_0 the subspace of $U \in \Upsilon^k$ satisfying $U(a) = U(b) = 0$. The* ***index form*** *of γ is the symmetric bilinear form $I : \Upsilon^k_0 \times \Upsilon^k_0 \to \mathbb{R}$ defined by*

$$I(U, V) \doteqdot \int_a^b \left[g(\nabla_t U, \nabla_t V) - R(\gamma', U, \gamma', V)\right] .$$

Integrating by parts, we find, for $U, V \in \Upsilon^{k+2}_0$, $k \geq 0$, that

$$I(U, V) = -\int_a^b g\left(\nabla_t^2 U + R(\gamma', U)\gamma'), V\right) \, dt$$

and, similarly,

$$I(U, V) = -\int_a^b g\left(U, \nabla_t^2 V + R(\gamma', V)\gamma')\right) \, dt \, .$$

That is, the linear operator $\mathcal{L} : \Gamma^{k+2}(\gamma^* TM) \to \Gamma^k(\gamma^* TM)$ defined by

$$-\mathcal{L}U \doteqdot \nabla_t^2 U + R(\gamma', U)\gamma'$$

is self-adjoint with respect to the inner product

$$\langle U, V\rangle \doteqdot \int_a^b g(U, V) \, dt$$

and satisfies

$$I(U, V) = \langle \mathcal{L}U, V\rangle \, .$$

Definition 19.7. *A vector field $J \in \Upsilon^1$ satisfying*

$$0 = -\mathcal{L}J \doteqdot \nabla_t^2 J + R(\gamma', J)\gamma' \tag{19.5}$$

in the sense that

$$I(U, J) = \int_a^b \left[g(\nabla_t U, \nabla_t J) + R(\gamma', U, \gamma', J)\right] dt = 0$$

for all $U \in \Upsilon^1_0$ is called a ***Jacobi field*** *along γ. The operator $\mathcal{L}$ is called the* ***Jacobi operator*** *and equation (19.5) is called* ***Jacobi's equation****. We denote by $\mathcal{J}$ the set of all Jacobi fields along γ.*

Theorem 19.8. *Let (M^n, g) be a Riemannian manifold. Given a geodesic $\gamma : [a, b] \to M$, the set $\mathcal{J}$ of Jacobi fields along γ is an $\mathbb{R}$-linear space of dimension $2n$. Indeed, given $t_0 \in [a, b]$ and $u, v \in T_{\gamma(t_0)}M$ there exists a unique Jacobi field J along γ satisfying $J(t_0) = u$ and $\nabla_t J(t_0) = v$. In particular, if $J(t_0) \neq 0$ then*

$$|J|^2 + |\nabla_t J|^2 > 0 \quad in \quad [a, b] \, .$$

Moreover, if $J(t_0) = 0$ then either $J \equiv 0$ or there is a neighbourhood I of t_0 in $[a, b]$ such that $|J(t)| > 0$ for all $t \in I \setminus \{t_0\}$.

Proof. Let $\{e_i\}_{i=1}^n$ be an orthonormal basis for $T_{\gamma(t_0)}M$ and let $\{E_i(t)\}_{i=1}^n$ be the parallel vector fields along γ satisfying $E_i(t_0) = e_i$. Then $\{E_i(t)\}_{i=1}^n$ is an orthonormal basis for $T_{\gamma(t)}M$ for all $t \in [a,b]$ and we can write

$$J(t) = J^i(t)E_i(t)\,.$$

Set

$$B_{ij}(t) \doteqdot R(\gamma', E_i, \gamma', E_j)(t)\,.$$

Note that B_{ij} is symmetric for every t. Jacobi's equation then reads

$$(Y^k)'' + \sum_{i=1}^n Y^i B_{ik} = 0\,.$$

The first claim now follows from the existence and uniqueness of solutions to linear second order ODE with prescribed initial data. The remaining claims follow from uniqueness of solutions (in particular, the zero Jacobi field is the unique Jacobi field along γ satisfying $J(t_0) = \nabla_t J(t_0) = 0$). $\square$

Lemma 19.9. *Let (M,g) be a Riemannian manifold and $\gamma : [a,b] \to M$ a geodesic. For every $U, V \in \mathcal{J}$,*

$$g(\nabla_t U, V) - g(U, \nabla_t V) \quad \textit{is constant.}$$

In particular, for any $J \in \mathcal{J}$ there are constants $a, b \in \mathbb{R}$ such that

$$g(J, \gamma')(t) = at + b$$

and hence $\mathcal{J}^\perp \doteqdot \{J \in \mathcal{J} : g(J,\gamma') \equiv 0\}$ is a linear subspace of $\mathcal{J}$ of codimension 2.

Proof. Observe that

$$\begin{aligned}\partial_t\left(g(\nabla_t U, V) - g(U, \nabla_t V)\right) &= g(\nabla_t^2 U, V) - g(U, \nabla_t^2 V)\\ &= R(\gamma', U, \gamma', V) - R(\gamma', V, \gamma', U)\\ &= 0\,.\end{aligned}$$

This proves the first claim. The second follows because $\gamma' \in \mathcal{J}$ and hence, by the first claim,

$$\partial_t^2 g(J,\gamma') = \partial_t g(\nabla_t J, \gamma') = \partial_t g(J, \nabla_t \gamma') = 0\,.$$

It follows that $\mathcal{J}^\perp$ is the kernel of the linear map $L : \mathcal{J}^\perp \to \mathbb{R}^2$ defined by

$$L(J) = (a,b)\,,$$

where a and b are the constants satisfying $g(J,\gamma')(t) = at + b$. This yields the final claim. $\square$

Definition 19.10. *Let (M, g) be a Riemannian manifold and $\gamma : [a, b] \to M$ a geodesic. Given $t_0, t_1 \in [a, b]$, the point $\gamma(t_1)$ is **conjugate to** $\gamma(t_0)$ **along** γ if there exists a non identically zero Jacobi field $J \in \mathcal{J}$ such that $J(t_0) = J(t_1) = 0$.*

Note that such a Jacobi field is necessarily in $\mathcal{J}^\perp$.

Proposition 19.11. *Let (M^n, g) be a Riemannian manifold and $\gamma : [a, b] \to M$ a geodesic. Given $t_0 \in (a, b]$, suppose that $\gamma(t_0)$ is not conjugate to $\gamma(a)$. Then for every $u \in \{\gamma'(t_0)\}^\perp$ there exists a unique $J \in \mathcal{J}^\perp$ such that $J(a) = 0$ and $J(t_0) = u$.*

Proof. Let $\mathcal{J}_0^\perp$ be the subspace of Jacobi fields $J \in \mathcal{J}^\perp$ satisfying $J(a) = 0$. That is, $\mathcal{J}_0^\perp$ is the kernel of the surjective linear map

$$\begin{aligned} \mathcal{J}^\perp &\to \{\gamma'(a)\}^\perp \\ J &\mapsto J(a)\,. \end{aligned}$$

In particular, $\dim \mathcal{J}_0^\perp = n - 1$. The claim follows since, by assumption, the kernel of the linear map

$$\begin{aligned} \mathcal{J}_0^\perp &\to \{\gamma'(t_0)\}^\perp \\ J &\mapsto J(t_0) \end{aligned}$$

is trivial. □

Example 19.12. *Let (M^n, g) be a Riemannian manifold and $\gamma : [a, b] \to M$ a unit speed geodesic. Suppose that M has constant sectional curvature $K \equiv \kappa \in \mathbb{R}$ along γ; that is,*

$$R(\gamma'(t), u, \gamma'(t), u) = \kappa$$

for all unit vectors $u \in \{\gamma'(t)\}^\perp$ for all $t \in [a, b]$.

Then Jacobi's equation becomes

$$\nabla_t^2 J + \kappa J = 0\,.$$

Since γ' is parallel along γ we can choose a parallel orthonormal frame $\{E_i\}_{i=1}^n$ along γ with $E_n = \gamma'$. Then, writing $J = J^k E_k$,

$$(J^i)'' + \kappa J^i = 0\,.$$

If $J \in \mathcal{J}^\perp$, we conclude that

$$J(t) = C_\kappa(t)A(t) + S_\kappa(t)B(t)\,,$$

where $A(t)$ and $B(t)$ are parallel vector fields along γ which are pointwise orthogonal to γ' and the functions C_κ and S_κ are the solutions of the ordinary differential equation

$$\psi'' + \kappa\psi = 0$$

with respective initial conditions

$$(C_\kappa(t_0), C'_\kappa(t_0)) = (1,0) \quad \textit{and} \quad (S_\kappa(t_0), S'_\kappa(t_0)) = (0,1)$$

for some $t_0 \in [a,b]$. *More explicitly,*

$$C_\kappa(t) \doteqdot \begin{cases} \cos\sqrt{\kappa}(t-t_0) & \textit{if} \quad \kappa > 0 \\ 1 & \textit{if} \quad \kappa = 0 \\ \cosh\sqrt{-\kappa}(t-t_0) & \textit{if} \quad \kappa < 0 \end{cases} \tag{19.6}$$

and

$$S_\kappa(t) \doteqdot \begin{cases} \frac{1}{\sqrt{\kappa}}\sin\sqrt{\kappa}(t-t_0) & \textit{if} \quad \kappa > 0 \\ t-t_0 & \textit{if} \quad \kappa = 0 \\ \frac{1}{\sqrt{-\kappa}}\sinh\sqrt{-\kappa}(t-t_0) & \textit{if} \quad \kappa < 0\,. \end{cases} \tag{19.7}$$

In particular, if $\kappa \le 0$ *then* $\gamma(t_0)$ *has no conjugate points along* γ. *On the other hand, if* $\kappa > 0$, *then* $\gamma(t)$ *is conjugate to* $\gamma(t_0)$ *along* γ *if and only if* $t = t_0 + \frac{l\pi}{\sqrt{\kappa}}$ *for some integer* l.

Theorem 19.13 (C. F. Jacobi (1836))**.** *Let* (M^n, g) *be a Riemannian manifold and* $\gamma : [a,b] \to M$ *a geodesic.*

- *If* $\gamma(a)$ *has no conjugate points along* $\gamma_{(a,b]}$ *then the index form is positive definite on* Υ^1_0.
- *If* $\gamma(a)$ *has no conjugate points along* $\gamma|_{(a,b)}$, *then the index form is non-negative definite on* Υ^1_0 *and* $I(U,U) = 0$ *if and only if* U *is a Jacobi field satisfying* $J(a) = J(b) = 0$ *(that is,* $U \in \mathcal{J}^\perp \cap \Upsilon^1_0$*).*

Proof. Let $\{J_i\}_{i=1}^{n-1}$ be a basis for $\mathcal{J}_0^\perp$ (the space of orthogonal Jacobi fields which vanish at a). Since $\gamma(a)$ has no conjugate points along $\gamma_{(a,b)}$, the fields $\{J_i\}_{i=1}^{n-1}$ are pointwise linearly independent for every $t \in (a,b)$. Thus, for any $U \in \Upsilon^1_0$, we can write

$$U(t) = U^j(t) J_j(t)\,.$$

We claim that

$$|\nabla_t U|^2 - R(\gamma', U, \gamma', U) = |(U^i)' J_i|^2 + \partial_t g(U^i \nabla_t J_i, U)$$

Indeed, (since $J_i(a) = 0$ for each i)

$$g(\nabla_t J_i, J_j) - g(J_i, \nabla_t J_j) \equiv 0$$

for every i and j and hence

$$\begin{aligned}|\nabla_t U|^2 - R(\gamma', U, \gamma', U) &= |(U^i)' J_i|^2 + 2(U^i)' U^j g(J_i, \nabla_t J_j) \\ &\quad + U^i U^j \left(g(\nabla_t J_i, \nabla_t J_j) - R(\gamma', J_i, \gamma', J_j)\right) \\ &= |(U^i)' J_i|^2 + (U^i U^j)' g(J_i, \nabla_t J_j) \\ &\quad + U^i U^j \left(g(\nabla_t J_i, \nabla_t J_j) + g(\nabla_t^2 J_i, J_j)\right) \\ &= |(U^i)' J_i|^2 + [g(U^i \nabla_t J_i, U)]' .\end{aligned}$$

Thus,

$$\begin{aligned}I(U,U) &= \lim_{\varepsilon \searrow 0} \int_{a+\varepsilon}^{b-\varepsilon} \left(|\nabla_t U|^2 - R(\gamma', U, \gamma', U)\right) dt \\ &= \lim_{\varepsilon \searrow 0} \left(U^i g(\nabla_t J_i, U)\Big|_{a+\varepsilon}^{b-\varepsilon} + \int_{a+\varepsilon}^{b-\varepsilon} |(U^i)' J_i|^2 \, dt \right) .\end{aligned}$$

If U^i is bounded on (a,b) for each i, then

$$I(U,U) = \lim_{\varepsilon \searrow 0} \int_{a+\varepsilon}^{b-\varepsilon} |(U^i)' J_i|^2 \, dt \geq 0$$

with equality only if the coefficients U^i are constant for each i, in which case $U \in \mathcal{J}^\perp \cap \Upsilon_0^1$. The claims follow. It remains to show that the coefficients are bounded.

Set $l = \dim \Upsilon_0 \cap \mathcal{J}^\perp$ and let $\{J_1, \ldots, J_l\}$ be a basis for $\Upsilon_0 \cap \mathcal{J}^\perp$. Next, set $e_i \doteqdot \nabla_t J_i(a)$ for each $i = 1, \ldots, l$ and extend $\{e_i\}_{i=1}^l$ to a basis $\{e_i\}_{i=1}^{n-1}$ for $\{\gamma'(a)\}^\perp$. Finally, for each $j = l+1, \ldots, n-1$, we let J_j be the Jacobi field along γ with $J_j(a) = 0$ and $\nabla_t J_j(a) = e_j$. Since $J_i(a) = 0$ for each $i = 1, \ldots, l$, we find that $\{J_j(b)\}_{j=l+1}^{n-1}$ and $\{\nabla_t J_j(b)\}_{j=1}^{l}$ are linearly independent. We also have

$$\{\nabla_t J_i(b)\}_{i=1}^l \perp \{J_j(b)\}_{j=l+1}^{n-1}$$

and hence $\{\nabla_t J_i(b)\}_{i=1}^l \cup \{J_j(b)\}_{j=l+1}^{n-1}$ is a basis for $\{\gamma'(b)\}^\perp$.

Next, we use Taylor's Theorem to write, for any $U \in \Gamma(\gamma^* TM)$,

$$\begin{aligned}\tau_{t,t_0} U(t) &= U(t_0) + (t - t_0) \frac{d}{dt}\Big|_{t=t_0} \tau_{t,t_0} U(t) + o(t - t_0) \\ &= U(t_0) + (t - t_0) \nabla_t U(t_0) + o(t - t_0) ,\end{aligned}$$

where $\tau_{t,t_0} : T_{\gamma(t)} M \to T_{\gamma(t_0)} M$ denotes parallel translation along γ, and hence

$$U(t) = \tau_{t_0,t} \left[U(t_0) + (t - t_0) \nabla_t U(t_0)\right] + o(t - t_0) . \tag{19.8}$$

Given $U \in \Upsilon_0$, choose $\{u^i\}_{i=1}^{n-1}$ so that

$$\nabla_t U(b) = \sum_{i=1}^{l} u^i \nabla_t J_i(b) + \sum_{j=l+1}^{n-1} u^j J_j(b)\,.$$

Then, applying (19.8), we obtain for $t \in [a,b]$ close to b

$$\begin{aligned} U(t) &= (t-b)\tau_{b,t}\nabla_t U(b) + o(b-t) \\ &= (t-b)\tau_{b,t}\left(\sum_{i=1}^{l} u^i \nabla_t J_i(b) + \sum_{j=l+1}^{n-1} u^j J_j(b)\right) + o(b-t)\,. \end{aligned}$$

Applying (19.8) again then yields

$$= \sum_{i=1}^{l} u^i J_i(t) + (t-b)\sum_{j=l+1}^{n-1} u^j J_j(t) + o(b-t)\,.$$

Taking $t \to b$, we conclude

$$\lim_{t \nearrow b} U(t) = \sum_{j=1}^{l} u^j J_j(b)\,.$$

That is,

$$\lim_{t \nearrow b} U^i(t) = \begin{cases} u^i & \text{for} \quad i = 1, \dots, l \\ 0 & \text{for} \quad i = l+1, \dots, n-1\,. \end{cases}$$

This completes the proof. $\square$

Theorem 19.13 says that if $\gamma : [0,b] \to M$ is a unit speed geodesic with no points in $\gamma((0,b])$ conjugate to $\gamma(0)$ along γ, then γ is a *strict local minimum* of the distance function among curves joining $\gamma(0)$ and $\gamma(b)$ in the sense that $\partial_\varepsilon|_{\varepsilon=0} L(\omega(\cdot,\varepsilon)) = 0$ and $\partial^2_\varepsilon|_{\varepsilon=0} L(\omega(\cdot,\varepsilon)) > 0$ for any homotopy ω of γ.

Theorem 19.14 (C. F. Jacobi (1836))**.** *Let (M,g) be a Riemannian manifold and $\gamma : [a,b] \to M$ a geodesic. Suppose that $\gamma(t_0)$ is conjugate to $\gamma(a)$ for some $t_0 \in (a,b)$. Then there exists $U \in \Upsilon_0$ such that $I(U,U) < 0$. Thus, γ cannot minimize distance past t_0.*

Proof. Let $J \in \mathcal{J}^\perp$ be a Jacobi field which vanishes at a and t_0 and set

$$U(t) \doteqdot \begin{cases} J(t) & \text{for} \quad t \in [a, t_0] \\ 0 & \text{for} \quad t \in (t_0, b]\,. \end{cases}$$

Then $U \in \Upsilon_0^1$ and $I(U,U) = 0$. We will perturb U to a vector field U_λ along γ satisfying $I(U_\lambda, U_\lambda) < 0$. Indeed, set $V_0 = -\nabla_t J(t_0)$ and let V be the corresponding parallel vector field along γ. Now choose a smooth function $\varphi : [a,b] \to \mathbb{R}$ satisfying $\varphi(t_0) = 1$ and $\varphi(a) = \varphi(b) = 0$ and set

$$U_\lambda(t) \doteqdot U(t) + \lambda\varphi(t)V(t)\,.$$

Then

$$\begin{aligned}
I(U_\lambda, U_\lambda) &= I(U,U) + 2\lambda\, I(U, \varphi V) + O(\lambda^2)\\
&= 2\lambda\, I(U,\varphi V) + O(\lambda^2)\\
&= 2\lambda \int_a^b \big[g(\nabla_t U, \nabla_t(\varphi V)) - R(\gamma', U, \gamma', \varphi Z)\big]\, dt + O(\lambda^2)\\
&= 2\lambda \int_a^{t_0} \big[g(\nabla_t J, \nabla_t(\varphi V)) - R(\gamma', J, \gamma', \varphi Z)\big]\, dt + O(\lambda^2)\\
&= 2\lambda g(\nabla_t J, \varphi V)\big|_a^{t_0} - 2\lambda \int_a^{t_0} g(\nabla_t^2 J + R(\gamma', J)\gamma', \varphi Z) dt + O(\lambda^2)\\
&= \,-2\lambda |\nabla_t J(t_0)|^2 + O(\lambda^2)\,.
\end{aligned}$$

For λ sufficiently small, this can be made negative. $\square$

19.1. Exercises.

Exercise 19.1. *Let (M^n, g) be a Riemannian manifold and $\gamma : [0,b] \to M$ a geodesic with no points in $\gamma((0,b])$ conjugate to $\gamma(0)$ along γ and let $\omega : [0,b] \times (-\varepsilon_0, \varepsilon_0) \to M$ be a homotopy of γ. Show that there is some $\varepsilon_1 \in (0, \varepsilon_0)$ such that $L(\omega(\cdot, \varepsilon)) > L(\gamma)$ for all $\varepsilon \in (-\varepsilon_1, \varepsilon_1) \setminus \{0\}$.*

Theorem 19.14 says that any geodesic $\gamma[a,b] \to M$ which contains an interior point conjugate to $\gamma(a)$ admits small perturbations (with fixed endpoints) which decrease length, so that γ does not locally minimize L amongst curves joining its endpoints.

Note that even a geodesic which locally minimizes the length between its endpoints may not globally minimize this length: consider the cylinder $S^1 \times \mathbb{R}$; the curves $\gamma_z : t \mapsto (\mathrm{e}^{it}, z)$ for fixed z locally minimize L on intervals of length less than 2π, but do not globally minimize for intervals of length greater than π. These curves do minimize L within their homotopy classes, however. On the other hand, if we close up the ends of the cylinder at $z = \pm L$, $L > 10$, by smoothly attaching almost spherical caps, then the curve $\gamma_0|_{[a,b]}$ is homotopic to its "complement", $t \mapsto \gamma_0|_{[0,2\pi-(b-a)]}(2\pi+a-t)$, and hence, for $b-a \in (\pi, 2\pi)$, does not even minimize L within its homotopy class.

20. Elementary comparison theorems

Our next goal is to show, using the preliminary results of the previous section, how control on the curvature of a Riemannian manifold implies control on other geometric and topological aspects.

Our first result shows the intuitive notion that a positively curved manifold must "close up" on itself.

Theorem 20.1 (O. Bonnet[19] (1855), S. B. Myers (1941))**.** *Let (M^n, g) be a Riemannian manifold and $\gamma : [0, b] \to M$ a unit speed geodesic along which*

$$\mathrm{Ric}(\gamma', \gamma') \geq (n-1)\kappa$$

for some $\kappa > 0$. If $b \geq \frac{\pi}{\sqrt{\kappa}}$ then $\gamma((0, b])$ contains a point conjugate to $\gamma(0)$ along γ.

In particular, if (M, g) is a complete Riemannian manifold of dimension $n \geq 2$ satisfying

$$\mathrm{Ric}(u, u) \geq (n-1)\kappa |u|^2 \quad \textit{for all} \quad u \in TM$$

then M is compact and has diameter $\mathrm{diam}(M) \leq \frac{\pi}{\sqrt{\kappa}}$.

Proof. Let $\{e_i\}_{i=1}^n$ be a basis for $\gamma(0)$ with $e_n = \gamma'(0)$ and let $\{E_i\}_{i=1}^n$ be the corresponding parallel frame along γ. Set, for each $i = 1, \ldots, n-1$,

$$U_i(t) \doteqdot \sin\left(\frac{\pi t}{b}\right) E_i(t)\,.$$

Then $U_i \in \Upsilon_0$ and

$$\nabla_t U_i(t) = \frac{\pi}{b} \cos\left(\frac{\pi t}{b}\right) E_i(t)$$

and hence

$$\begin{aligned}
\sum_{i=1}^{n-1} I(U_i, U_i) &= \sum_{i=1}^{n-1} \int_0^b \left[\frac{\pi^2}{b^2} \cos^2\left(\frac{\pi t}{b}\right) - R(E_i, \gamma', E_i, \gamma') \sin^2\left(\frac{\pi t}{b}\right)\right] dt \\
&= \int_0^b \left[(n-1)\frac{\pi^2}{b^2} \cos^2\left(\frac{\pi t}{b}\right) - \mathrm{Ric}(\gamma', \gamma') \sin^2\left(\frac{\pi t}{b}\right)\right] dt \\
&\leq (n-1) \int_0^b \left[\frac{\pi^2}{b^2} \cos^2\left(\frac{\pi t}{b}\right) - \kappa \sin^2\left(\frac{\pi t}{b}\right)\right] dt \\
&= (n-1)\frac{b}{2}\left(\frac{\pi^2}{b^2} - \kappa\right).
\end{aligned}$$

If $b \geq \frac{\pi}{\sqrt{\kappa}}$ the right hand side is non-positive. The first claim now follows from Theorem 19.13.

[19] The conclusions were proved first by Bonnet under the stronger assumption that the sectional curvatures are bounded from below. The more general statement, in which only the Ricci curvature is bounded from below, was shown by Myers.

To prove the second claim, note that given any $p, q \in M$ there exists (by the completeness assumption) a unit speed geodesic $\gamma : [0, b] \to M$ such that $\gamma(0) = p$, $\gamma(b) = q$ and $b = L(\gamma) = d(p, q)$. Since $\gamma|_{(0,b]}$ is a minimizing geodesic, its index form is positive semidefinite, which implies

$$d(p, q) = b \leq \frac{\pi}{\sqrt{\kappa}} \, .$$

The second claim follows. □

Another simple consequence of Jacobi's investigations is that negatively curved manifolds must "open up" in the sense that they cannot have conjugate points.

Theorem 20.2 (J. Hadamard (1898), E. Cartan (1946))**.** *Let (M^n, g) be a Riemannian manifold. A unit speed geodesic $\gamma : [0, b] \to M$ in M along which*

$$K \leq 0$$

Then contains no point conjugate to $\gamma(0)$ in $\gamma((0, b])$. In particular, if M is complete with nonpositive sectional curvature, then M has no conjugate points.

Proof. Given $U \in \Upsilon_0$,

$$I(U, U) = \int_0^b \left[|\nabla_t U|^2 - R(\gamma', U, \gamma', U) \right] \, dt \geq 0 \, .$$

The claim now follows from Theorem 19.14. □

We note that this does not necessarily mean that all geodesic segments are global minimizing (consider the flat torus).

The next result provides a more quantitative conclusion.

Theorem 20.3 (M. Morse (1930), I. J. Schönberg (1932))**.** *Let (M^n, g) be a Riemannian manifold and $\gamma : [0, b] \to M$ a unit speed geodesic in M along which*

$$K \leq \delta$$

for some $\delta > 0$. If $\gamma(b)$ is conjugate to $\gamma(0)$ along γ then $b \geq \frac{\pi}{\sqrt{\delta}}$.

Proof. By assumption, there exists a non-trivial Jacobi field $J \in \mathcal{J}^\perp \cap \Upsilon_0$ along γ. By Wirtinger's inequality and the Cauchy–Schwarz inequality,

$$\frac{\pi^2}{b^2} \int_0^b |J|^2 \, dt \leq \int_0^b |\nabla_t J|^2 \, dt \, .$$

Thus,

$$\begin{aligned} 0 = I(J,J) &= \int_0^b \left[|\nabla_t J|^2 - R(\gamma', J, \gamma', J) \right] \, dt \\ &\geq \int_0^b \left[|\nabla_t J|^2 - \delta |J|^2 \right] \, dt \\ &\geq \left(\frac{\pi^2}{b^2} - \delta \right) \int_0^b |J|^2 \, dt \,. \end{aligned}$$

The claim follows. ☐

The following lemma allows us to compute the derivative of the exponential map in terms of Jacobi fields.

Lemma 20.4. *Let (M^n, g) be a Riemannian manifold of dimension $n \geq 2$. Given $\xi \doteqdot (p, u) \in \mathcal{T}M$ and $v \in T_pM$, let $J \in \mathcal{J}$ be the Jacobi field along $\gamma_\xi(t) \doteqdot \exp_p tu$ satisfying*

$$J(0) = 0 \quad \text{and} \quad \nabla_t J(0) = v \,.$$

Then, identifying $T_{tu}(T_pM) \cong T_pM$ in the usual way,

$$(d \exp_p)|_{tu}(v) = t^{-1} J(t)$$

for all $t \in I_\xi$, the domain of γ_ξ.

In particular, if $(x^1, \ldots, x^n) : U \to \mathbb{R}^n$ are exponential normal coordinates about p with respect to an orthonormal basis $\{e_i\}_{i=1}^n$ for T_pM then

$$\partial_i|_{\exp_p tu} = t^{-1} J_i(t) \,,$$

where $J_i \in \mathcal{J}$ is the Jacobi field along $t \mapsto \exp_p tu$ satisfying

$$J_i(0) = u \quad \text{and} \quad \nabla_t J_i(0) = e_i \,.$$

Proof. Let $\omega : I_\xi \times (-\varepsilon_0, \varepsilon_0) \to M$ be the geodesic variation of γ_ξ given by

$$\omega(t, \varepsilon) \doteqdot \exp_p(t(u + \varepsilon v)) \,.$$

Then $J(t) \doteqdot \partial_\varepsilon \omega(t, 0)$ is the Jacobi field along γ satisfying $J(0) = u$ and $\nabla_t J(0) = v$ (recall that the derivative of $\exp_p$ at the zero vector is the identity map). Thus,

$$(D_{tu} \exp_p)(tv) = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \exp_p(t(u + \varepsilon v)) = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \omega(t, \varepsilon) = J \,. \qquad ☐$$

Corollary 20.5. *Let (M^n, g) be a Riemannian manifold of dimension $n \geq 2$. Given $\xi \doteqdot (p, u) \in \mathcal{T}M$, the null space of $D_u \exp_p$ is isomorphic to the subspace of Jacobi fields along the geodesic $t \mapsto \exp_p tu$ which vanish at p and $\exp_p u$.*

The Theorem of Hadamard and Cartan yields the following.

Corollary 20.6 (J. Hadamard (1898), E. Cartan (1946))**.** *If (M, g) is a complete Riemannian manifold with non-positive sectional curvature, then* exp *is of maximal rank on all of TM.*

Corollary 20.7. *Let (M^n, g) be a Riemannian manifold of constant sectional curvature $\kappa \in \mathbb{R}$. Given $p \in M$, let $\{e_i\}_{i=1}^n$ be an orthonormal basis for T_pM. With respect to the corresponding exponential normal coordinates $\varphi : (x^1, \ldots, x^n) \mapsto \exp_p x^k e_k$,*

$$g_{ij}(\varphi(x)) = \frac{x^i x^j}{|x|^2} + \frac{S_\kappa^2(|x|)}{|x|^2}\left(\delta_{ij} - \frac{x^i x^j}{|x|^2}\right).$$

In geodesic polar coordinates $\varphi : (r, \xi) \mapsto \exp_p r\xi$ based at $p \in M$,

$$\varphi^* g\big|_{(r,\xi)} = dr \otimes dr\big|_{(r,\xi)} + S_\kappa^2(r) g_{n-1}(\xi)$$

where g_{n-1} is the standard metric on the unit sphere $S^{n-1} \subset T_pM$ and S_κ is the generalized sine function defined in (19.7)*.*

In particular, any two Riemannian manifolds of the same dimension $n \geq 2$ and the same constant sectional curvature $\kappa \in \mathbb{R}$ are locally isometric.

Proof. By Lemma 20.4,

$$\partial_i\big|_{\exp_p x^k e_k} \doteqdot D_{x^k e_k} \exp_p e_i = |x|^{-1} J_i(|x|),$$

where J_i is the Jacobi field along the geodesic $t \mapsto \exp_p t\frac{x^k e_k}{|x|}$ satisfying $J(0) = 0$ and $\nabla_t J(0) = e_i$. Solving Jacobi's equation, we obtain

$$J_i(t) = (a_i t + b_i)\gamma'(t) + S_\kappa(t) A_i(t) + C_\kappa(t) B_i(t),$$

where A_i and B_i are parallel vector fields along γ satisfying which are pointwise orthogonal to γ'. The initial condition $J_i(0) = 0$ implies $b = 0$ and $B \equiv 0$ and the initial condition $\nabla_t J_i(0) = e_i$ implies

$$\left(a_i \frac{x^k}{|x|} + A_i{}^k\right) e_k = e_i$$
$$\implies a_i = \frac{x^i}{|x|} \quad \text{and} \quad A_i^k = \delta_{ik} - \frac{x^i x^k}{|x|^2}.$$

Thus,

$$
\begin{aligned}
g_{ij}(x) &= g(\partial_i\big|_{\exp_p x^k e_k}, \partial_j\big|_{\exp_p x^k e_k}) \\
&= g\left(\frac{J_i(|x|)}{|x|}, \frac{J_j(|x|)}{|x|}\right) \\
&= \frac{x^i x^j}{|x|^2} + \frac{S_\kappa^2(|x|)}{|x|^2}\delta_{kl}\left(\delta_i^k - \frac{x^i x^k}{|x|^2}\right)\left(\delta_j^l - \frac{x^j x^l}{|x|^2}\right) \\
&= \frac{x^i x^j}{|x|^2} + \frac{S_\kappa^2(|x|)}{|x|^2}\left(\delta_{ij} - \frac{x^i x^j}{|x|^2}\right).
\end{aligned}
$$

This proves the first claim. The second claim is proved similarly, making use also of the Gauss Lemma (Lemma 15.1). □

The next two results (Rauch's comparison theorems) allow us to control the lengths of Jacobi fields, and hence the "geodesic deviation" of our Reimannian manifold. They are key tools in the proofs of the Hessian comparison theorem (Theorem 20.11) and Toponogov's distance comparison theorem (Theorem 22.3) below.

Theorem 20.8 (H. E. Rauch (1951))**.** *Let (M^n, g) be a Riemannian manifold and $\gamma : [0, b] \to M$ a unit speed geodesic in M along which*

$$K \le \delta$$

for some $\delta \in \mathbb{R}$. Let $J \in \mathcal{J}^\perp$ be a Jacobi field along γ. Then

$$|J|'' + \delta|J| \ge 0 \quad on \quad (0, b]\,.$$

Let $\psi \in C^\infty([0, b])$ be the solution of

$$\psi'' + \delta\psi = 0\,, \quad \psi(0) = |J|(0)\,, \quad \psi'(0) = |J|'(0)\,.$$

If ψ is non-zero on $(0, b)$ then

$$\left(\frac{|J|}{\psi}\right)' \ge 0 \quad and \quad \frac{|J|}{\psi} \ge 1 \quad on \quad (0, b)\,.$$

Moreover, $\left(\frac{|J|}{\psi}\right)'$ reaches zero at some point $t_0 \in (0, b)$ if and only if $R(\gamma', J)\gamma' \equiv \delta J$ in $[0, t_0]$ and there exists a parallel unit vector field E along γ for which

$$J(t) = \psi(t)E(t) \quad on \quad [0, t_0]\,.$$

Proof. To prove the first claim, we calculate

$$|J|' = g\left(\nabla_t J, \frac{J}{|J|}\right)$$

and, using the Cauchy–Schwarz inequality,

$$
\begin{aligned}
|J|'' &= g\left(\nabla_t\nabla_t J, \frac{J}{|J|}\right) + g\left(\nabla_t J, \frac{\nabla_t J}{|J|} - \frac{1}{|J|^2} g\left(\nabla_t J, \frac{J}{|J|}\right)\right) \\
&= -\frac{1}{|J|}R(\gamma', J, \gamma', J) + \frac{|\nabla_t J|^2}{|J|} - \frac{1}{|J|^3} g\left(\nabla_t J, J\right)^2 \\
&\geq -\delta|J| + \frac{1}{|J|^3}\left(|\nabla_t J|^2|J|^2 - g\left(\nabla_t J, J\right)^2\right) \\
&\geq -\delta|J|\,. \qquad (20.1)
\end{aligned}
$$

Next, consider

$$F \doteqdot \psi^2\left(\frac{|J|}{\psi}\right)' = |J|'\psi - |J|\psi'\,.$$

Then

$$F' = |J|''\psi - |J|\psi'' \geq 0\,.$$

That is, F is monotone non-decreasing. Since $F(0) = 0$, we conclude that $F \geq 0$ in $(0, b]$. Indeed, if $F(t_0) = 0$ for some $t_0 \in (0, b)$ then $F \equiv 0$, and hence $|J| \equiv \psi$, in $[0, t_0]$. Set $E \doteqdot J/|J|$ so that $J = \psi E$. Then

$$\nabla_t J = \psi' E + \psi \nabla_t E\,.$$

We claim that $\nabla_t J = \psi' E$ and hence $\nabla_t E = 0$. Indeed, since equality is attained in (20.1), it must hold in the Cauchy–Schwarz inequality, and hence $\nabla_t J$ and $J = \psi E$ are linearly dependent. But $\nabla_t E$ is orthogonal to E since the latter is of unit length. This implies the claim. □

Theorem 20.9 (H. E. Rauch (1951))**.** *Let (M^n, g) be a Riemannian manifold and $\gamma : [0, b] \to M$ a unit speed geodesic in M along which*

$$K \geq \kappa$$

for some $\kappa \in \mathbb{R}$. Suppose that $\gamma((0, b))$ contains no points conjugate to $\gamma(0)$ along γ. Let $J \in \mathcal{J}$ be a Jacobi field along γ satisfying $J(0) = 0$. Then

$$\frac{|J|'}{|J|} \leq \frac{C_\kappa}{S_\kappa} \quad \text{on} \quad (0, b)$$

and hence

$$\left(\frac{|J|}{S_\kappa}\right)' \leq 0 \quad \text{and} \quad \frac{|J|}{S_\kappa} \leq |\nabla_t J|(0) \quad \text{on} \quad (0, b)\,.$$

Moreover, $\left(\frac{|J|}{S_\kappa}\right)'$ reaches zero at $t_0 \in (0, b)$ if and only if $R(\gamma', J)\gamma' \equiv \kappa J$ in $[0, t_0]$ and there exists a parallel unit vector field E along γ for which

$$J(t) = S_\kappa(t)E(t) \quad \text{on} \quad [0, t_0]\,.$$

Proof. First note that $S_\kappa > 0$ on $(0, b)$ since the Bonnet–Myers Theorem (Theorem 20.1) implies $b \leq \frac{\pi}{\sqrt{\kappa}}$.

Fix $t \in (0, b)$. Then the index form of $\gamma|_{[0,t]}$ is positive definite on Υ_0. Let J be a Jacobi field along γ satisfying $J(0) = 0$ and let U be any piecewise C^1 vector field along γ which satisfies $U(0) = 0$ and $U(t) = J(t)$. Then

$$\begin{aligned} I_t(J, U) &\doteqdot \int_0^t \left[g(\nabla_t J, \nabla_t U) - R(\gamma', J, \gamma', U)\right] dt \\ &= \int_0^t \partial_t g(\nabla_t J, U)\, dt = g(\nabla_t J, U)(t)\,. \end{aligned}$$

and similarly,

$$I_t(J, J) = g(\nabla_t J, J)(t)$$

so that

$$\begin{aligned} 0 \leq I_t(J - U, J - U) &= I_t(J, J) - 2I_i(J, U) + I_t(U, U) \\ &= I_t(U, U) - g(\nabla_t J, J)(t)\,. \end{aligned}$$

Now set

$$U(s) \doteqdot \frac{S_\kappa(s)}{S_\kappa(t)} E(s)$$

for all $s \in [0, t]$, where E is the parallel vector field along γ satisfying $E(t) = J(t)$.

$$\begin{aligned} g(\nabla_t J, J)(t) &\leq \int_0^t \left[|\nabla_t U|^2 - R(\gamma', U, \gamma', U)\right] ds \\ &\leq \int_0^t \left[|\nabla_t U|^2 - \kappa |U|^2\right] ds \\ &= \int_0^t \left[|\nabla_t U|^2 - \kappa |U|^2\right] ds \\ &= \frac{|J(t)|^2}{S_\kappa^2(t)} \int_0^t \left(C_\kappa^2(s) - \kappa S_\kappa^2(s)\right) ds = |J(t)|^2 \frac{C_\kappa(t)}{S_\kappa(t)}\,. \end{aligned}$$

We conclude that

$$\frac{|J|'}{|J|} \leq \frac{C_\kappa}{S_\kappa} \quad \text{on} \quad (0, b)\,.$$

The remaining claims follow as in Theorem 20.8. ∎

Our first consequence of the Rauch comparison theorems is an estimate for the derivative of the exponential map, and hence the metric in exponential normal coordinates.

Corollary 20.10. *Let (M^n, g) be a Riemannian manifold. Given $\xi = (p, u) \in TM$, let $\gamma_\xi : I_\xi \to M$ be the geodesic satisfying $\gamma_\xi(0) = p$ and $\gamma_\xi'(0) = 0$. That is,$\gamma_\xi(t) \doteqdot \exp_p tu$. Suppose that*

$$\kappa \leq K \leq \delta$$

along γ_ξ and that γ_ξ has no points conjugate to p along $\gamma|_{(0,t]}$. Then

$$\frac{S_\delta(t)}{t} \leq \frac{|D_{tu} \exp_p(v)|}{|v|} \leq \frac{S_\kappa(t)}{t}$$

for all $v \in \{u\}^\perp$.

Thus, in exponential normal coordinates $\varphi : (x^1, \ldots, x^n) \mapsto \exp_p x^k e_k$ based at $p \in M$,

$$\frac{S_\delta(|x|)}{|x|}\delta_{ij} \leq g_{ij}(\varphi(x)) \leq \frac{S_\kappa(|x|)}{|x|}\delta_{ij} \,.$$

Note that the estimates in the corollary are sharp (in two different ways): First, equality clearly holds on the spaces of constant sectional curvature and, second, equality is attained when $t \to 0$ since $S_\kappa(r)/r \to 1$ as $r \to 0$ and, by Theorem 12.5 all the way back in Section 12, the derivative of the exponential map at the zero vector is the identity.

Proof of Corollary 20.10. Let J be the Jacobi field along γ_ξ satisfying $J(0) = 0$ and $\nabla_t J(0) = v$. Then

$$|J|' = g\left(\nabla_t J, \frac{J}{|J|}\right) = g\left(\nabla_t J, \frac{t^{-1}J}{t^{-1}|J|}\right) \,.$$

Taking $t \searrow 0$ yields

$$|J|'(0) = |v| \,.$$

Now combine Lemma 20.4 with Theorems 20.8 and 20.9. □

Theorem 20.11 (Hessian Comparison Theorem (local version))**.** *Let (M^n, g) be a complete Riemannian manifold and let $(r, \xi) : U \to \mathbb{R}_+ \times S^n$ be geodesic polar coordinates defined near $p \in U \subset M$. Then, for any $q \in U$,*

$$\operatorname{grad} r|_q = \gamma'(r(q)) \quad \textit{and} \quad \frac{\operatorname{Hess} r|_q(u,u)}{g(u,u)} = \frac{|J|'}{|J|}(r(q)) \,,$$

where J is the Jacobi field along the geodesic

$$\begin{aligned} \gamma : [0, r(q)] &\to M \\ t &\mapsto \exp_p t\xi(q) \end{aligned}$$

satisfying

$$J(0) = 0 \quad \textit{and} \quad J(r(q)) = u^\perp \,.$$

In particular, if the sectional curvatures of (M^n, g) *are bounded by*

$$\kappa \le K \le \delta$$

for some constants $\kappa,\ \delta \in \mathbb{R}$ *then*

$$\frac{C_\delta(r)}{S_\delta(r)} g^\perp \le \operatorname{Hess} r \le \frac{C_\kappa(r)}{S_\kappa(r)} g^\perp\,,$$

where $g^\perp \doteqdot g|_{\{\operatorname{grad} r\}^\perp}$.

Proof. Fix $q \in U$ and let $\gamma_{\xi_q} : [0, r_q] \to U$ be the unique unit speed geodesic joining $p = \gamma_{\xi_q}(0)$ and $q = \gamma_{\xi_q}(r_q)$, where $r_q \doteqdot r(q)$ and $\xi_q \doteqdot \xi(q)$. Consider a geodesic variation $\omega(t, \varepsilon) \doteqdot \exp_p t(\xi_q + \varepsilon\eta)$. Then the Jacobi field $J(t) \doteqdot \partial_\varepsilon \omega(t, 0)$ satisfies

$$J(0) = 0 \quad \text{and} \quad J(r_q) = D_{\xi_q} \exp_p \eta \doteqdot u\,.$$

Thus, by the first variation formula (Lemma 19.2),

$$g(\operatorname{grad} r|_q, u) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} r(\omega(r_q, \varepsilon)) = g(u, \gamma'(r_q))\,. \tag{20.2}$$

Since, by assumption, $\exp_p$ is a diffeomorphism on $\exp_p^{-1}(U)$, we can arrange that $J(r_q) = u$ for any $u \in T_qM$ we like. Thus, (20.2) holds for any $u \in T_qM$ and we deduce that

$$\operatorname{grad} r|_q = \gamma'(r_q)\,.$$

In particular,

$$\nabla_{\operatorname{grad} r} \operatorname{grad} r = 0\,.$$

So assume that $u \in \{\operatorname{grad} r|_q\}^\perp$. Then, applying the second variation formula (Corollary 19.5), we obtain

$$\begin{aligned}
\operatorname{Hess} r|_q(u, u) &+ g(\nabla_\varepsilon \partial_\varepsilon \omega|_{(r_q,0)}, \operatorname{grad} r|_q) \\
&= \frac{d^2}{d\varepsilon^2}\Big|_{\varepsilon=0} r(\omega(r_q, \varepsilon)) \\
&= g(\nabla_\varepsilon \partial_\varepsilon \omega|_{(r_q,0)}, \gamma'(r_q)) + g(u, \nabla_t J(r_q)) \\
\Longrightarrow \frac{\operatorname{Hess} r|_q(u, u)}{g(u, u)} &= \frac{1}{|J(r_q)|^2} g(J(r_q), \nabla_t J(r_q)) = \frac{|J|'}{|J|}(r_q)\,.
\end{aligned}$$

The remaining claims then follow from the Rauch Theorems 20.8 and 20.9. $\square$

21. The cut locus and the injectivity radius

Denote by

$$SM \doteqdot \{(p,u) \in TM : |u|_p = 1\}$$

the unit tangent bundle. Given $\xi = (p,u) \in SM$, consider the geodesic $\gamma_\xi(t) \doteqdot \exp_p(tu)$. By Lemma 20.4, the point $\gamma_\xi(t)$ is conjugate to p along γ_ξ if and only if $\operatorname{rank}(d\exp_p|_{tu}) < n$. So the **conjugate locus** of $p \in M$ (the set of points conjugate to p along geodesics emanating from p) coincides with the image under $\exp_p$ of its critical points.

We have seen that a geodesic cannot minimize distance past its first conjugate point; however, the possibility remains that it stops minimizing distance at an earlier point. The set of points at which the geodesics emanating from a point $p \in M$ cease minimizing distance is called the **cut locus** of p.

Definition 21.1. *Let (M,g) be a Riemannian manifold. Denote by $SM \doteqdot \{(p,u) \in TM : |u| = 1\}$ the unit tangent bundle of M. The* ***cut distance*** *is the function $c : SM \to \mathbb{R}$ defined by*

$$c(p,u) \doteqdot \sup\{t > 0 : (p,tu) \in \mathcal{T}M\,,\ d(p, \exp_p tu) = t\}\,.$$

The ***tangential cut locus*** *of a point $p \in M$ is the set*

$$\mathrm{C}_p \doteqdot \{c(p,u)u : c(p,u) < \infty\,,\ |u| = 1\} \cap \mathcal{T}M$$

The ***cut locus*** *of a point $p \in M$ is the set*

$$C_p \doteqdot \exp_p \mathrm{C}_p\,.$$

We also define the sets

$$\mathrm{D}_p \doteqdot \{tu : |u| = 1\,,\ 0 \le t < c(p,u)\} \quad \textit{and} \quad D_p \doteqdot \exp_p \mathrm{D}_p\,.$$

In the following, we will use the following basic properties of the cut distance and cut locii.

Lemma 21.2. *Let (M,g) be a connected Riemannian manifold. Given $\xi = (p,u) \in TM$ denote by $\gamma_\xi : I_\xi \to M$ the maximal geodesic through $p = \gamma_\xi(0)$ in the direction $u = \gamma_\xi'(0)$.*

(1) *γ_ξ minimizes distance between p and $\gamma_\xi(t)$ for all $t < c(\xi)$ and for no $t > c(\xi)$.*
(2) *If $t < c(\xi)$, then γ_ξ is the unique minimizing geodesic joining p and $\gamma_\xi(t)$.*

Suppose now that (M,g) is complete.

(4) *If $c(\xi) < \infty$, and $\gamma_\xi(c(\xi))$ is not conjugate to $\gamma_\xi(0)$, then there exists at least two distinct minimizing geodesics joining p and $\gamma_\xi(c(\xi))$.*

(5) *If* $c(\xi) < \infty$*, then* $c(-\gamma'_\xi(c(\xi))) = c(\xi)$*. Thus,* $q \in C_p$ *if and only if* $p \in C_q$*.*

(6) *If* γ_ξ *minimizes the distance between* p *and* $\gamma_\xi(t)$ *for all* $t > 0$*, then* $c(\xi) = \infty$*.*

(7) D_p *is the largest starshaped open subset of* T_pM *on which* $\exp_p$ *is a diffeomorphism.*

(8) $D_p = M \setminus C_p$*.*

Proof. (1) If γ_ξ minimizes distance between $\gamma_\xi(0)$ and $\gamma_\xi(t_0)$, then it minimizes distance between $\gamma_\xi(0)$ and $\gamma_\xi(t)$ for all $t \in [0, t_0]$. (This fact, which we have already seen, is a simple consequence of the triangle inequality).

(2) If ω is a minimizing geodesic joining p and $\gamma_\xi(t)$, then, by (1), the concatenation $\omega * \gamma_\xi|_{[t,c(\xi)]}$ of ω and $\gamma_\xi|_{[t,c(\xi)]}$ has length $L(\omega) + L(\gamma_\xi|_{[t,c_\xi(p)]}) = d(p, \gamma_\xi(t)) + d(\gamma_\xi(t), c_\xi(p)) = d(p, c_\xi(p))$. So $\omega * \gamma_\xi|_{[t,c(\xi)]}$ is minimizing, and hence, by Corollary 19.3, a (smooth) geodesic. In particular, the tangent vectors to ω and γ_ξ coincide at $\gamma_\xi(t)$, so we conclude that $\omega * \gamma_\xi|_{[t,c(\xi)]} = \gamma_\xi$.

(3) Consider a strictly decreasing sequence of times $t_j \in I_\xi$ with $t_j \to c_\xi(p)$ as $j \to \infty$. By the Hopf–Rinow Theorem (Theorem 15.4) we can find a minimizing geodesic γ_{η_j} joining p and $\gamma_\xi(t_j)$. Since S_pM is compact, some subsequence of the sequence of vectors v_j converges to a limit $v \in S_pM$. If $v = u$, then $\exp_p$ is not one-to-one in any neighbourhood of $c(\xi)u$, and hence, by the contrapositive of the inverse function theorem, $(d\exp_p)|_{c(\xi)u}$ cannot have maximal rank, and hence, by Lemma 20.4, $\gamma_\xi(c(\xi))$ is conjugate to p along γ_ξ. So, in fact, $v \neq u$, which implies the claim.

(4) This follows from (3).

(5) This follows from (1).

(6) This follows from (3) and Lemma 20.4.

(7) Due to the Hopf–Rinow theorem, every point $q \in M$ is reached by a minimizing geodesic starting from p, with initial unit tangent u, say. Thus, $d(p, q) \leq c(p, u)$. □

Proposition 21.3. *Let* (M, g) *be a Riemannian manifold. If* (M, g) *is complete then the cut distance* $c : SM \to \mathbb{R}$ *is continuous.*

Proof. Let $\xi_k = (p_k, u_k)$ be a sequence of points in SM which converge to $\xi \in SM$ as $k \to \infty$ and set $d_k \doteqdot c(\xi_k)$. We first prove upper semi-continuity. That is,

$$\limsup_{k\to\infty} d_k \leq c(\xi)\,.$$

We first deal with the case $\limsup_{k\to\infty} d_k = \infty$. Then for every $T > 0$ there is some element d_k of the sequence satisfying $d_k > T$. By completeness,

$$\lim_{k\to\infty} \gamma_{\xi_k}(T) = \gamma_\xi(T)$$

and hence

$$d(p, \gamma_\xi(T)) = \lim_{k\to\infty} d(p_k, \gamma_{\xi_k}(T)) = T\,.$$

It follows that $c(\xi) = \infty$.

If, instead, $\limsup_{k\to\infty} d_k < \infty$ then the sequence is bounded and hence has a convergent subsequence $d_k \to d$. Then, for any $\varepsilon > 0$,

$$d(p, \gamma_\xi(d-\varepsilon)) = \lim_{k\to\infty} d(p_k, \gamma_{\xi_k}(d_k - \varepsilon)) = \lim_{k\to\infty} (d_k - \varepsilon) = d - \varepsilon\,.$$

The claim follows.

It remains to prove lower semi-continuity; i.e.

$$\liminf_{k\to\infty} d_k \geq c(\xi)\,.$$

If $\liminf_{k\to\infty} d_k = \infty$, we are done. Otherwise, we pass to a convergent subsequence $d_k \to d \doteqdot \liminf_{k\to\infty} d_k$. Passing to a further subsequence, we can assume (from part (3) of Lemma 21.2) that either

(i) $\gamma_{\xi_k}(d_k)$ is conjugate to p_k along γ_{ξ_k} for all k or

(ii) there exists a sequence of unit vectors $\eta_k = (p_k, v_k) \in SM$ with $\eta_k \neq \xi_k$ and $\gamma_{\xi_k}(d_k) = \gamma_{\eta_k}(d_k)$ for all k.

Recalling Lemma 20.4, we find in the first case that $D_{d_k u_k} \exp_{p_k}$ has non-trivial kernel for each k. It follows that there is no neighbourhood of $(d\xi, d\xi) \in TM$ on which $\pi \times \exp$ is a diffeomorphism. By the inverse function theorem, $D_{du} \exp_p$ must also have non-trivial kernel and we conclude that $\gamma_\xi(d)$ is conjugate to p. Jacobi's Theorem 19.14 and part (1) of Lemma 21.2 then imply that $c(\xi) \leq d$.

We are left with case (ii). If in the limit $\xi \neq \eta$ then we have two distinct geodesics joining p with $\gamma_\xi(d)$. It follows that $c(\xi) \leq d$: If not, for sufficiently small $\varepsilon > 0$, the (non-smooth) curve which traverses γ_1 from p to $\gamma_1(d)$ and then γ_2 to $\gamma_2(d+\varepsilon)$ would be length minimizing and hence a geodesic, a contradiction. So suppose that $\xi = \eta$. Then we are again in the situation that $\pi \times \exp$ cannot be a diffeomorphism on any neighbourhood of $(d\xi, d\xi) \in TM$. $\square$

Definition 21.4. *Let (M, g) be a Riemannian manifold. Given $p \in M$, the **injectivity radius of** p is*

$$\mathrm{inj}(p) \doteqdot \inf_{u \in S_pM} c(p, u)\,.$$

*The **injectivity radius of** M is*

$$\operatorname{inj}(M) \doteqdot \inf_{p\in M} \operatorname{inj}(p)\,.$$

Observe that $\operatorname{inj}(p)$ is positive for all $p \in M$ and hence $\operatorname{inj}(M)$ is always non-negative.

Proposition 21.5. *Let (M, g) be a complete Riemannian manifold. Then the function $p \mapsto \operatorname{inj}(p)$ is continuous.*

Proof. The claim follows from compactness of S_pM and continuity of the cut distance. We leave the proof as an exercise. $\square$

Lemma 21.6 (Klingenberg's Lemma (1959))**.** *Let (M, g) be a complete Riemannian manifold. Given $p \in M$, suppose that $q \in C_p$ satisfies*

$$d(p, C_p) = d(p, q)\,.$$

If q is not conjugate to p along a minimizing geodesic joining p to q then q is the midpoint of a geodesic loop starting and ending at p.

In particular, if M is compact with sectional curvature satisfying

$$K \le \delta$$

then

$$\operatorname{inj}(M) \ge \min\left\{\frac{\pi}{\sqrt{\delta}}, \frac{\ell(M)}{2}\right\},$$

where $\ell(M)$ is the length of the shortest simple closed geodesic in M.

Proof. By part (3) of Lemma 21.2, if q is not conjugate to p then there exist two distinct unit speed geodesics γ_1 and γ_2 joining p and q and neither contain any points conjugate to p. Set $L \doteqdot d(p, q)$, let U_1 and U_2 be neighbourhoods in S_pM of $\gamma_1'(0)$ and $\gamma_2'(0)$ respectively and consider the hypersurfaces

$$U_1 \doteqdot \{\exp_p Lu : u \in \mathrm{U}_1\} \quad \text{and} \quad U_2 \doteqdot \{\exp_p Lu : u \in \mathrm{U}_2\}\,.$$

Note that the vectors $\gamma_1'(L)$ and $\gamma_2'(L)$ are normal to U_1 and U_2 respectively. If $\gamma_1'(L) \ne \gamma_2'(L)$, then for sufficiently small $\varepsilon > 0$ the two hypersurfaces

$$U_1^\varepsilon \doteqdot \{\exp_p(L-\varepsilon)u : u \in \mathrm{U}_1\} \quad \text{and} \quad U_2^\varepsilon \doteqdot \{\exp_p(L-\varepsilon)u : u \in \mathrm{U}_2\}$$

intersect. This contradicts the assumption that q is the closest point of C_p to p. This proves the first claim.

The second claim now follows from the Morse–Schöneberg Theorem, which gives the lower bound $\frac{\pi}{\sqrt{\delta}}$ for the distance between conjugate points. $\square$

Exercise 21.1. *Prove Lemma 21.2.*

22. Distance comparison

The ultimate goal of this section is to prove the Toponogov distance comparison theorem, which can be thought of as an integrated or nonlinear version of Rauch's comparison theorems for Jacobi fields. The ideas trace back to Alexandrov, who proved the theorem for convex surfaces. Toponogov's proof was technical and contained some difficulties which were later resolved by Gromov, Klingenberg and Meyer and since then further proofs have arisen, notably Karcher's proof using the Hessian comparison theorem, in the following "global" version

Theorem 22.1 (Hessian Comparison Theorem (global version))**.** *Let (M^n, g) be a complete Riemannian manifold with sectional curvatures bounded by*

$$K \geq \delta$$

for $\delta \in \mathbb{R}$. Given $p \in M$, the distance function $r(x) \doteqdot d(p,x)$ satisfies

$$\operatorname{Hess} r \leq \frac{C_\delta(r)}{S_\delta(r)} g^\perp$$

in the ***sense of supports****; that is, for each $x \in M$ and each $\varepsilon > 0$ there exists $r > 0$ and $\psi \in C^\infty(B_r(x))$ satisfying*

$$\psi(y) \geq r(y) \ \ \forall \ \ y \in B_r(x)\,, \quad \psi(x) = r(x) \tag{22.1}$$

and

$$\operatorname{Hess} \psi \leq \left(\frac{C_\delta(r)}{S_\delta(r)} + o(1) \right) g^\perp \quad at \quad x$$

as $\varepsilon \to 0$.

Proof. Fix $x \in M$ and a unit speed length minimizing geodesic γ joining p and x. If r is smooth at x, the claim follows from the local version. So suppose that this is not the case. We claim that, for ε sufficiently small, the function

$$r_\varepsilon(y) \doteqdot \varepsilon + d(y, \gamma(\varepsilon))$$

supports r from above at x in the sense of (22.1). Indeed,

$$r_\varepsilon(x) = \varepsilon + d(x, \gamma(\varepsilon)) = \varepsilon + d(x,p) - \varepsilon = r(x)$$

and, by the triangle inequality,

$$r(y) = d(p,y) \leq d(p, \gamma(\varepsilon)) + d(y, \gamma(\varepsilon)) = \varepsilon + d(y, \gamma(\varepsilon)) = r_\varepsilon(y)\,.$$

We claim that r_ε is smooth at x for ε small. Indeed, suppose that this is not the case. Then the point $\gamma(\varepsilon)$ is a cut point of x and hence either $\gamma(\varepsilon)$ is a conjugate point of x or there are two distinct minimizing geodesics joining $\gamma(\varepsilon)$ and x. The second case cannot occur since the path from x to p is minimizing. Thus, $\gamma(\varepsilon)$ is a conjugate point of x. But this cannot occur either since γ is a minimizing geodesic from p to x and geodesics cannot be minimizing past conjugate points.

Thus,

$$\begin{aligned}\operatorname{Hess} r_\varepsilon|_x &\leq \frac{C_\delta(r_\varepsilon(x))}{S_\delta(r_\varepsilon(x))} g_x^\perp \\ &= \frac{C_\delta(r(x)-\varepsilon)}{S_\delta(r(x)-\varepsilon)} g_x^\perp = \left(\frac{C_\delta(r(x))}{S_\delta(r(x))} + o(1)\right) g_x^\perp\,.\end{aligned}$$ □

Definition 22.2. *Let (M,g) be a Riemannian manifold. A* ***geodesic juncture*** *in (M,g) is a triple $(\gamma_1,\gamma_2,\alpha)$ consisting of unit speed, length minimizing geodesics $\gamma_i : [0,b_i] \to M$, $i = 1,2$ which meet at $\gamma_1(0) = \gamma_2(0)$ and their opening angle $\alpha = \arccos g(\gamma_1'(0), \gamma_2'(0))$.*

A ***geodesic triangle*** *in (M,g) is a triple $(\gamma_1,\gamma_2,\gamma_3)$ consisting of unit speed length minimizing geodesics $\gamma_i : [0,b_i] \to \mathbb{R}$ satisfying*

$$\gamma_i(b_i) = \gamma_{i+1}(0)$$

where the indices are taken modulo 3. The points $p_i \doteqdot \gamma_{i+2}(0)$ are called the ***vertices*** *and the geodesics γ_i the* ***edges*** *of $(\gamma_1,\gamma_2,\gamma_3)$. The* ***corresponding angles*** *are denoted by $\alpha_i \doteqdot \arccos g(\gamma_i'(0))$.*

We denote by $(\overline{M}{}^n_\kappa, \overline{g})$ the n-dimensional model space of constant sectional curvature κ. That is, the sphere S^n_κ of radius $r = 1/\sqrt{\kappa}$ when $\kappa > 0$, Euclidean space $\mathbb{R}^n$ when $\kappa = 0$ and the hyperbolic space H^n_κ of radius $r = 1/\sqrt{-\kappa}$ when $\kappa < 0$.

Theorem 22.3 (Distance comparison theorem, Toponogov (1959))**.** *Let (M^n,g) be a complete Riemannian manifold with sectional curvature satisfying $K \geq \delta$ and let $(\gamma_1,\gamma_2,\alpha)$ be a geodesic juncture in (M^n,g). Then*

$$d(\gamma_1(s),\gamma_2(t)) \leq \overline{d}(\overline{\gamma}_1(s),\overline{\gamma}_2(t)) \quad \textit{for all} \quad (s,t) \in [0,b_1]\times[0,b_2]$$

for any geodesic juncture $(\overline{\gamma}_1,\overline{\gamma}_2,\alpha)$ in $(\overline{M}{}^n_\kappa, \overline{g})$.

Proof. Fix $s_0 \in (0,b]$, set $p_0 \doteqdot \gamma_1(s_0)$ and $\overline{p}_0 \doteqdot \overline{\gamma}_1(s_0)$ and, for convenience, write $\gamma \doteqdot \gamma_2$. Consider the squared distance functions

$$\rho(t) \doteqdot f_\kappa(r(\gamma(t))) \quad \text{and} \quad \overline{\rho}(t) \doteqdot f_\kappa(\overline{r}(\overline{\gamma}(t)))\,,$$

where r and $\overline{r}$ are the distance functions to p_0 and $\overline{p}_0$ respectively and the function f_κ is the antiderivative of S_κ satisfying $f_\kappa(0) = 0$; that is,

$$f_\kappa(r) \doteqdot \begin{cases} \dfrac{1}{\kappa}\left(1-\cos\sqrt{\kappa}\,r\right) \text{if} & \kappa > 0 \\ \dfrac{1}{2}r^2 \text{if} & \kappa = 0 \\ \dfrac{1}{\kappa}\left(1-\cosh\sqrt{-\kappa}\,r\right) \text{if} & \kappa < 0\,. \end{cases}$$

Then, by the Hessian comparison theorem,

$$\begin{aligned}\rho'' &= \left(f_\kappa''(r\circ\gamma)dr|_\gamma \otimes dr|_\gamma + f_\kappa'(r\circ\gamma)\operatorname{Hess} r|_\gamma\right)(\gamma',\gamma')\\ &\leq C_\kappa(r\circ\gamma)\left(dr\otimes dr + g^\perp\right)|_\gamma(\gamma',\gamma')\\ &= C_\kappa(r\circ\gamma)\end{aligned}$$

if r is smooth at $\gamma(t)$ (else, we need to replace r by the function r_ε as in Theorem 22.1; we leave this technical detail as an exercise). Similarly,

$$\overline{\rho}'' = C_\kappa(\overline{r}\circ\overline{\gamma})\,.$$

It follows that

$$\psi'' + \kappa\psi \geq 0\,,$$

where $\psi \doteqdot \overline{\rho} - \rho$ (in the sense of supports at points of γ where r is not smooth). Noting that

$$\psi(0) = 0 \quad \text{and} \quad \psi'(0) = 0$$

the claim follows as in the proof of Theorem 20.9, at least when ψ is smooth. We leave it as a (non-trivial) exercise to check that a similar argument can be carried out when ψ is non-smooth. $\square$

For completeness we mention, without proof, the following corollary.

Corollary 22.4 (Angle comparison Theorem, Toponogov (1959))**.** *Let (M^n, g) be a complete Riemannian manifold with sectional curvature satisfying $K \geq \delta$ and let $(\gamma_1, \gamma_2, \gamma_3)$ be a geodesic triangle in (M^n, g) (with corresponding angles α_i) and $(\overline{\gamma}_1, \overline{\gamma}_2, \overline{\gamma}_3)$ any geodesic triangle in M^n_κ (with corresponding angles $\overline{\alpha}_i$) satisfying $L(\gamma_i) = \overline{L}(\overline{\gamma}_i)$ for each i. Then*

$$\alpha_i \geq \overline{\alpha}_i$$

for each i.

Proof. This follows from a Riemannian analogue of the Cosine Rule. See, for example, Peterson (2006). $\square$

23. Integration on Riemannian manifolds

Let Ω_1 and Ω_2 be domains in Euclidean space $\mathbb{R}^n$ and $\varphi : \Omega_1 \to \Omega_2$ a C^1 diffeomorphism. Given any $f \in L^1(\Omega_2)$, the change of variables law for the Lebesgue measure asserts that

$$\int_{\Omega_1} (f \circ \varphi) |\det D\varphi| \, dV = \int_{\Omega_2} f \, dV \,,$$

where V is the Lebesgue measure.

Now let (M^n, g) be a Riemannian manifold and $\varphi : U \to \mathbb{R}^n$ a chart for M. For eaxh $p \in U$ consider the matrix $g^\varphi(x)$ with components

$$g^\varphi(p)_{ij} \doteqdot g_{ij}(p) = g_p(\partial_i^\varphi|_p, \partial_j^\varphi|_p) \,.$$

where ∂_i^φ are the coordinate basis vectors corresponding to the chart φ; that is, for any $f \in C^\infty(M)$,

$$\partial_i|_p f \doteqdot D_{\varphi(p)}(f \circ \varphi^{-1})(e_i) \,.$$

If $\psi : V \to \mathbb{R}^n$ is a second chart with non-trivial intersection $U \cap V \neq \emptyset$, then

$$\partial_i^\psi|_p = D_{\psi(p)}(\varphi \circ \psi^{-1})(\partial_i^\varphi|_p) \doteqdot G_p(\partial_i^\varphi|_p)$$

and hence

$$g^\psi = G^T g^\varphi G$$

from which we conclude

$$\sqrt{\det g^\psi} = \det G \sqrt{\det g^\varphi} \,.$$

Thus, for any domain $\Omega \subset U$, the integral

$$\int_\Omega f \, d\mu \doteqdot \int_{\varphi(\Omega)} (f \sqrt{\det g^\varphi}) \circ \varphi^{-1} \, dV$$

is well-defined. We can then extend this definition to arbitrary domains $\Omega \subset M$ by introducing a partition of unity. Indeed, choose an atlas $\{\varphi_\alpha : U_\alpha \to \mathbb{R}^n \,, \alpha \in A\}$ and a subordinate partition of unity $\{\rho_\alpha : U_\alpha \to [0, 1] : \alpha \in A\}$ and set

$$\int_\Omega f \, d\mu \doteqdot \sum_{\alpha \in A} \int_{\varphi_\alpha(\Omega \cap U_\alpha)} (f \, \rho_\alpha \sqrt{\det g^{\varphi_\alpha}}) \circ \varphi_\alpha^{-1} \, dV \,.$$

We refer to μ as the **Riemannian measure** of (M, g). It is clear that a function $f : M \to \mathbb{R}$ is μ-measurable if and only if $f \circ \varphi^{-1}$ is Lebesgue measurable for any chart φ. We denote the set of μ-measurable functions by $L^1(M)$.

On an oriented *Riemannian* manifold (M^n, g), we can use integration of differential forms to define a notion of integration for functions by picking out a special nowhere zero n-form, $\sigma \in \Omega^n(M)$, and defining the value of the integral of a function f as the value of the integral of the n-form $f\sigma$.

Definition 23.1. *Let (M^n, g) be an oriented Riemannian manifold. The* ***Riemannian volume form*** *$\sigma \in \Omega^n(M)$ is the n-form defined locally in any chart $\varphi : U \to \mathbb{R}^n$ by*

$$\sigma \doteqdot \operatorname{sign}(\varphi)\sqrt{\det g^{\varphi}}\, dx^1 \wedge \cdots \wedge dx^n \,,$$

where $\operatorname{sign}(\varphi)$ *is equal to* 1 *if $dx^1 \wedge \cdots \wedge dx^n$ is positively oriented and* -1 *otherwise.*

Then the corresponding integral of any function $f \in L^1(M)$ is well-defined and

$$\int f\sigma = \int f\, d\mu \,.$$

We can now use Stokes' theorem to generalize the divergence theorem from multivariable calculus. Recall that the divergence $\operatorname{div} X$ of a vector field X is defined by

$$\operatorname{div} X \doteqdot \operatorname{tr}(\nabla X) \,.$$

We will need the following lemma:

Lemma 23.2. *Let (M, g) be an oriented Riemannian manifold with volume form σ and $X \in \Gamma(TM)$ a vector field. Then*

$$\mathcal{L}_X \sigma = (\operatorname{div} X)\sigma \,.$$

Let (M, g) be an oriented Riemannian manifold with boundary with volume form σ and $X \in \Gamma(TM)$ a vector field. Then

$$\iota_X \sigma\big|_{\partial M} = g(X, N)\tau \,,$$

where τ is the induced volume form on ∂M and N is the inward pointing unit normal field.

Proof. We will make use of the following identity for the Lie derivative of a differential form:

Exercise 23.1. *Given a differentiable manifold M, let $\omega \in \Omega^k(M)$ be a k-form and $X \in \Gamma(TM)$ a vector field on M. Then*

$$\mathcal{L}_X \omega = d\iota_X \omega + \iota_X d\omega$$

Hint: Check that the right hand side is a derivation on the exterior algebra which distributes over the wedge product and agrees with the Lie derivative on 0- *and* 1-*forms.*

In particular, since σ is an n-form,

$$\mathcal{L}_X \sigma = d\iota_X \sigma$$

Locally, we compute

$$\begin{aligned}
\mathcal{L}_X\sigma &= d\left(\iota_X\sqrt{\det g^x}\,dx^1\wedge\cdots\wedge dx^n\right)\\
&= d\left(\sum_{j=1}^n(-1)^{j-1}X^j\sqrt{\det g^x}\,dx^1\wedge\cdots\wedge\widehat{dx^j}\wedge\cdots\wedge dx^n\right)\\
&= \left(\sum_{j=1}^n(-1)^{j-1}d(X^j\sqrt{\det g^x})\wedge dx^1\wedge\cdots\wedge\widehat{dx^j}\wedge\cdots\wedge dx^n\right)\\
&= \left(\sum_{j=1}^n\frac{\partial(X^j\sqrt{\det g^x})}{\partial x^j}\wedge dx^1\wedge\cdots\wedge dx^n\right)\\
&= \frac{1}{\sqrt{\det g^x}}\sum_{j=1}^n\frac{\partial(X^j\sqrt{\det g^x})}{\partial x^j}\,\sigma\,.
\end{aligned}$$

This proves the first claim by Exercise 23.2.

To prove the second, observe that

$$\sigma=\omega^1\wedge\cdots\wedge\omega^n$$

with respect to a local orthonormal basis $\{\omega_i\}_{i=1}^n$ for $\Omega^1(M)$. At boundary points, we can arrange that ω^1 is dual to the inward normal vector N. Then, at boundary points,

$$\sigma=\omega^1\wedge\tau$$

and hence

$$\iota_X\sigma=g(X,N)\tau-\omega^1\wedge\iota_X\tau\,.$$

The second claim follows. □

Theorem 23.3 (Divergence Theorem)**.** *Let (M,g) be an oriented Riemannian manifold with boundary. Then*

$$\int_M \operatorname{div} X\,d\mu=\int_{\partial M}g(X,N)\,d\nu\,,$$

where N is the inward pointing normal to ∂M and $d\nu$ the induced Riemannian measure.

Proof. Using Lemma 23.2 and Cartan's formula

$$\mathcal{L}_X\omega=d\iota_X\omega+\iota_X d\omega$$

for the Lie derivative of a k-form ω, we compute

$$\begin{aligned}\int_M \operatorname{div} X \, d\mu = \int_M (\operatorname{div} X)\sigma &= \int_M \mathcal{L}_X \sigma \\ &= \int_M (d\iota_X \sigma + \iota_X d\sigma) \\ &= \int_M d\iota_X \sigma \\ &= \int_{\partial M} \iota_X \sigma \\ &= \int_{\partial M} g(X,N)\tau = \int_{\partial M} g(X,N)\, d\nu \,. \qquad \square\end{aligned}$$

23.1. Exercises.

Exercise 23.2. *Show that the divergence operator may be expressed as*

$$\operatorname{div} X \doteqdot \operatorname{tr}(\nabla X) = \frac{1}{\sqrt{\det g^x}} \sum_{j=1}^{n} \frac{\partial (X^j \sqrt{\det g^x})}{\partial x^j}$$

within the chart $x : U \to \mathbb{R}^n$*.*

Bibliography

[1] Benn, I. M., and Tucker, R. W. *An introduction to spinors and geometry with applications in physics*. Adam Hilger, Ltd., Bristol, 1989. Reprint of the 1987 original.

[2] Chavel, I. *Riemannian geometry*, second ed., vol. 98 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006. A modern introduction.

[3] do Carmo, M. P. a. *Riemannian geometry*. Mathematics: Theory & Applications. Birkhäuser Boston, Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty.

[4] Lee, J. M. *Riemannian manifolds*, vol. 176 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1997. An introduction to curvature.

[5] Lee, J. M. *Introduction to smooth manifolds*, first, second ed., vol. 218 of *Graduate Texts in Mathematics*. Springer, New York, 2003, 2013.

[6] Milnor, J. W. *Topology from the differentiable viewpoint*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997. Based on notes by David W. Weaver, Revised reprint of the 1965 original.

[7] Petersen, P. *Riemannian geometry*, third ed., vol. 171 of *Graduate Texts in Mathematics*. Springer, Cham, 2016.

[8] Warner, F. W. *Foundations of differentiable manifolds and Lie groups*, vol. 94 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1983. Corrected reprint of the 1971 edition.